Robust Sparse Bayesian Learning Scheme for DOA Estimation with Non-Circular Sources

Linlu Jian ¹, Xianpeng Wang ¹*, Jinmei Shi ² and Xiang Lan ¹

1 State Key Laboratory of Marine Resource Utilization in South China Sea, School of Information and Communication Engineering, Hainan University, Haikou 570228, China; 20081002110014@hainanu.edu.cn (L.J.); xlan@hainanu.edu.cn (X.L.)
2 College of Information Engineering, Hainan Vocational University of Science and Technology, Haikou 571158, China; shijinmei1003@hvust.edu.cn

* Correspondence: 993559@hainanu.edu.cn

Abstract: In this paper, a robust DOA estimation scheme based on sparse Bayesian learning (SBL) for non-circular signals in impulse noise and mutual coupling (MC) is proposed. Firstly, the Toeplitz property of the MC matrix is used to eliminate the effect of array MC, and the array aperture is extended by using the properties of the non-circular signal. To eliminate the effect of impulse noise, the outlier part of the impulse noise is reconstructed together with the original signal in the signal matrix, and the DOA coarse estimation is obtained by balancing the accuracy and efficiency of parameter estimation using the alternating SBL update algorithm. Finally, a one-dimensional search is used in the vicinity of the searched spectral peaks to achieve a high-precision DOA estimation. The effectiveness and robustness of the algorithm for dealing with the above errors are demonstrated by extensive simulations.

Keywords: sparse Bayesian learning; non-circular signal; impulsive noise; mutual coupling; DOA estimation

1. Introduction

In recent years, with the development of array signal processing, direction-of-arrival (DOA) estimation has gradually become a research hotspot [1], and its development has a significant impact on the fields of radar, sonar, and mobile communications [1,2]. With the innovation and development of technology, DOA estimation techniques are widely used in target localization and navigation systems [3–5]. The DOA estimation technique originated with Pisarenko’s discovery that DOA information could be obtained from the second order statistics of the receiver matrix. Soon after, many traditional direction finding algorithms were proposed. Subspace class DOA estimation algorithms represented by multiple signal classification (MUSIC) based methods [6,7] and rotational invariant techniques (ESPRT) based methods [8,9] were then proposed, which mainly exploit the relationship between the signal subspace and the noise subspace to achieve DOA estimation. Under ideal conditions, subspace class methods can perfectly approximate the Cramer–Rao lower bound (CRLB) [7]. However, the development of science and technology has made the electromagnetic wave environment more and more complicated. Non-ideal conditions, such as the correlation between signal sources, low signal-to-noise ratio, and small snapshots, make subspace class methods have great limitations in practical applications. Then, subspace fitting algorithms represented by weighted subspace fitting (WSF) [10,11] and maximum likelihood (ML) algorithm [12,13] are proposed. Although these types of algorithms solve the problem of poor performance of subspace algorithms under non-ideal conditions to
a certain extent, it is restricted by computational efficiency, and it is difficult to exert its performance advantages.

At a time when the development of traditional high-resolution DOA estimation algorithms is limited by environmental or efficiency constraints, the rise of compressive sensing (CS) techniques offers a new solution to target vectoring. CS technology that has emerged in recent years has improved the large amount of redundancy and waste caused by traditional sampling rules (i.e., the Nyquist sampling theorem) [14]. It directly samples useful information, which greatly saves the cost of data sampling, and can recover the original signal from a small number of measured values through the numerical optimization algorithm. From then on, a large number of DOA estimation methods based on sparse representation (SR) have been proposed, such as sparse Bayesian learning (SBL) [15,16] and convex relaxation class methods [17–19]. These methods can still effectively and accurately estimate DOAs in non-ideal situations with small snapshots, low SNR, and high source correlation. Although convex optimization class algorithms indirectly solve the NP-hard problem of $l_0$-parametric optimization, they suffer from high computational complexity and difficulty in ensuring global convergence, which is well solved by SBL-based algorithms [15,16]. However, the advantages of SBL-based vectoring algorithms often rely on ideal array manifolds and Gaussian noise environments.

In the actual array signal processing, compared with a separate antenna, the antenna array is not only affected by its own magnetic field, but also by the electromagnetic field of adjacent elements. If the elements of array are not fully calibrated, it will directly affect the array manifold. Nevertheless, the accuracy of DOA estimation depends on the perfect array manifold, and the existence of mutual coupling (MC) effect will make most high-resolution DOA estimation methods lose estimation accuracy to a certain extent. To solve this problem, a number of relevant solutions have been proposed [20,21], and an iterative mutual coupling calibration method is proposed in [20], which is able to complete the array calibration under the condition that the prior information of the flow pattern matrix is unknown, but it has high computational burden associated with iteration. Ye et al. proposed a non-iterative algorithm in literature [21], which can achieve the Toeplitz structure of the MC matrix only without any calibration source. This approach was subsequently extended and applied to various algorithms for DOA estimation by Dai et al. [22,23]. On the other hand, non-Gaussian noise with significant spike characteristics, called impulse noise, exists in the real environment. Under conventional Gaussian noise models, these large spikes can be considered as outliers. To obtain ideal DOA information under impulsive noise, many SR-based methods model impulsive noise as a combination of Gaussian noise and outlier matrices [24–26]. Dai et al. then rewrite the signal matrix with outlier matrices to avoid the effect on DOA estimation in SBL iterations [24]. Zheng et al. refer to the signal model in [24] and propose a variational SBL iterative estimation algorithm. This algorithm identifies and corrects outliers during the iterative process [25]. This iterative estimation algorithm thus becomes an efficient means of dealing with impulse noise.

Whether it is the subspace based methods or SR based methods, the performance of DOA estimation mostly depends on the array aperture. Using the properties of the signal itself to extend the array aperture is an effective way to improve the performance of DOA estimation. It has been shown that the non-zero elliptic covariance property of non-circular signals can not only distinguish non-circular signals from circular signals in the source, but also bring better performance to the traditional DOA estimation methods. Due to this property, non-circular signals represented by BPSK and ASK are widely used in digital signal processing [27–31]. For example, Zheng et al. used the property of non-zero elliptic covariance of non-circular signals to construct partial joint sparsity of non-circular sources in the covariance vector and pseudo-covariance vector [31]. The non-circular sources are then identified by setting a set of latent variables, and finally DOA estimation under the condition of co-existence of non-circular and circular sources is achieved. On the other hand, array expansion is accomplished by exploiting the property of non-zero elliptic covariance of non-circular signals, which can effectively expand the array aperture. For
example, Zhang et al. exploited the properties of non-circular signals to process the received data and extended the array aperture, which improved the efficiency of the algorithm to some extent [30].

However, the impulsive noise environment was considered in the literature [30] without the array MC. This paper proposes a robust algorithm based on SBL for non-circular signals in impulsive noise and MC. The proposed method uses the selection matrix to eliminate the influence of the mutual coupling of the array, and uses the characteristics of the non-circular signal to expand the array and expand the array aperture. Then, the outlier matrix of impulse noise is put into the signal model, and the alternating update SBL algorithm is used to estimate the parameters of signal and noise, respectively. This procedure solves the problem of too high complexity of using traditional SBL to deal with impulse noise, and better balances the accuracy and efficiency of DOA estimation. Finally, after the rough estimate of DOA is obtained, an improved one-dimensional search method is used according to the reconstruction result to obtain a high-precision DOA estimate. A large number of simulation experiments show that, under the influence of impulse noise and array MC, the proposed algorithm can mitigate the impact of off-grid errors to a certain extent and has good DOA estimation performance.

Notations: Lowercase bold letters and uppercase bold letters, respectively, denote vectors and matrices. Inverse, conjugate, transpose, and conjugate transpose operations on a matrix are denoted by \((\cdot)^{-1}\), \((\cdot)^*\), \((\cdot)^T\), and \((\cdot)^H\), respectively. \(\text{tr}(\cdot)\) denotes the trace operation on a matrix. The diagonalization and vectorization operations are represented by \(\text{diag}(\cdot)\) and \(\text{vec}(\cdot)\), respectively. \(E(A)\) is to find the expectation of \(A\) with respect to \(B\). \(\Re\{\cdot\}\) is the real part operator.

2. Data Model

We consider that, on a uniform linear array (ULA) composed of \(N\) identical sensors, \(K\) uncorrelated narrowband far-field signals are incident, where \(K \leq N - 1\). In addition, the distance between adjacent elements in a ULA is \(d = \nu/2\), where \(\nu\) represents the wavelength of the incident signal. Therefore, for strictly non-circular signals, under conditions of array mutual coupling and coexistence of impulse noise, the data receiving matrix is rewritten as

\[
X = CAH + N,
\]

where \(X = [x(t_1), x(t_2), \ldots, x(t_T)]^T\) and \(T\) is the number of snapshots.

In the data model shown in Equation (1), the presence of MC allows the array manifold under ideal conditions, which is represented as \(A = [a(\theta_1), a(\theta_2), \ldots, a(\theta_K)]^T\) to change into \(\hat{A}\), where \(a(\theta_k) = [1, e^{-j\pi \sin\theta_k}, \ldots, e^{-j\pi(N-1)\sin\theta_k}]^T\) is the steering vector for the \(k\)-th signal source, and band-symmetric Toeplitz matrix \(C\) is often used to describe MC effects in ULA, and it is indicated as

\[
C = \text{Toeplitz}([c, 0_{1 \times (N-K-1)})],
\]

where \(c\) is the MC coefficient.

The non-circular signal \(\hat{S} = [\hat{s}(t_1), \hat{s}(t_2), \ldots, \hat{s}(t_T)]\) is split and modeled as a non-circular phase matrix \(\Psi = \text{diag}[e^{j\psi_1/2}, e^{j\psi_2/2}, \ldots, e^{j\psi_K/2}]\) and a real-valued signal \(S\), where \(\psi_k\) is a non-circular phase and, when it is zero, \(\Psi\) is an identity matrix and \(\hat{S}\) is a strictly non-circular signal satisfying \(\hat{S}' = \hat{S}\).

Similarly, the spike characteristics of impulse noise will cause errors in DOA estimation. To avoid the impact of its non-Gaussian part on the DOA estimation performance, the noise matrix is usually modeled as the sum of the outlier matrix \(H = [h(t_1), h(t_2), \ldots, h(t_T)]\) and the Gaussian noise matrix \(N = [n(t_1), n(t_2), \ldots, n(t_T)]\) according to the impulsive characteristics.
3. The Proposed Method for NC Sources

As the proposed method requires the use of sparse recovery techniques to recover the signal, constructing a sparse-complete dictionary relies on a perfect array manifold. We use the banded complex symmetric Toeplitz structure of the MC matrix to construct the selection matrix $J$ to remove the MC:

$$J = \begin{bmatrix} 0_{(N-2P) \times P}, I_{(N-2P)}, 0_{(N-2P) \times P} \end{bmatrix},$$

where $P$ is the number of MC coefficients, speaking of the selection matrix acting on the steering vector. The principle of eliminating mutual coupling is clearly expressed as

$$\bar{a}(\theta_k) = J \tilde{a}(\theta_k) = \tau_k \check{a}(\theta_k)$$

where $\tau_k = 1 + \sum_{k=1}^{K} 2c_k \cos(k \pi \sin \theta_k)$. When the left and right ends of Equation (1) are respectively multiplied by the selection matrix $J$, the data receiving matrix after eliminating MC can be obtained

$$Y = JX = J \tilde{A} \tilde{S} + JH + JN$$

where $\tilde{A}$ is the ideal array manifold, and it is composed of the middle $N_1$ line of $\tilde{A}$ with $N_1 = N - 2P$, which is illustrated in Figure 1. In addition, $\tilde{S} = \Xi \tilde{S}$, where $\Xi = \text{diag}[\tau_1, \tau_2, \ldots, \tau_k]$, when the angle is determined, this parameter $\tau_k$ is a constant.

As shown in the array model in Figure 1, the decoupling process results in a loss of array aperture. However, the properties of non-circular signals can be exploited to extend the data reception matrix by constructing a virtual array. This allows for DOA estimation under more targets. At the same time, all possible source angles $\theta_k = [\theta_1, \theta_2, \ldots, \theta_\hat{K}]$ are assumed to exist in the discrete space $[-\pi/2, \pi/2]$, where $\hat{K}$ is the number of discrete grids and satisfies $\hat{K} \gg N > K$. The extended data reception matrix space domain discretization is denoted as
\[
Z = \begin{bmatrix} \mathbf{D} \mathbf{Y}^* \\ \mathbf{Y} \end{bmatrix} = \begin{bmatrix} \mathbf{D} \mathbf{A}^* \mathbf{S}^* + \mathbf{D} \mathbf{H}^* + \mathbf{D} \mathbf{N}^* \\ \mathbf{A} \mathbf{S} + \mathbf{H} + \mathbf{N} \end{bmatrix} = \begin{bmatrix} \tilde{\mathbf{A}} & \mathbf{I}_{2 \times N_1} \\ \mathbf{B} & \mathbf{S} \end{bmatrix} + \mathbf{N},
\]

(6)

where \( \mathbf{D} \) is the permutation matrix with the anti-diagonal elements being 1, and the remaining elements being 0. \( \tilde{\mathbf{A}} = \begin{bmatrix} \mathbf{D} \mathbf{A}^* & 0 \\ 0 & \mathbf{A} \end{bmatrix} \) is the completion matrix where \( \mathbf{A} = [\hat{a}(\theta_1), \hat{a}(\theta_2), \ldots, \hat{a}(\theta_\hat{K})] \) is an \( N_1 \times \hat{K} \) dimensional matrix.

The signal reconstruction process based on SBL techniques uses the observation matrix and prior probabilities to recover the signal matrix. The corresponding array manifolds are further obtained, and the final DOA estimates are achieved. Since the signal matrix is the object to be estimated in this technique, the method of reconstructing the outlier matrix into the signal matrix directly eliminates the detrimental effects on signal recovery. Therefore, \( \mathbf{S} \) is the signal after re-modeling with the outlier matrix in Equation (6) where \( \hat{\mathbf{S}} = [\hat{\mathbf{S}}^* \ \mathbf{S}]^T \), and \( \hat{\mathbf{S}} \) is the signal matrix after discretization of the null domain, with each of its non-zero rows corresponding to a real signal. Similarly, \( \hat{\mathbf{N}} = [\mathbf{D} \mathbf{N}^* \ \mathbf{N}]^T \).

4. Alternating SBL Updates for DOA Estimation

4.1. A Priori Setting of Parameters

According to the received data model described in Equation (6), we can obtain the prior probability of \( Z \) as

\[
p(Z|\tilde{\mathbf{S}}, \beta) = \prod_{l=1}^{L} \mathcal{CN}(z_l|B \tilde{s}_l, \beta^{-1} I_{2N_1}),
\]

(7)

where \( \beta = \sigma^2 \) is the power of Gaussian noise \( \mathbf{N} \), which is usually unknown, so it will be assumed to satisfy the gamma distribution:

\[
p(\beta) = \text{Gamma}(\beta|\iota, \varpi),
\]

(8)

where the hyperparameter \( \iota \) is usually set to 1 and \( \varpi \rightarrow 0 \) to obtain a broad hyper-prior. Similarly, the signal \( \tilde{\mathbf{S}} \) is assumed to satisfy a Gaussian distribution with variance \( \mathbf{\Lambda} = [\lambda_1, \lambda_2, \ldots, \lambda_{2\hat{K}}] \), where \( \lambda_i \) corresponds to the signal variance of the \( i \)-th row. Its prior probability is given by

\[
p(\tilde{\mathbf{S}}|\lambda) = \prod_{l=0}^{L} \mathcal{CN}(s_l|0_{2\hat{K} \times 1}, \mathbf{\Lambda}),
\]

(9)

where \( \mathbf{\Lambda} = \text{diag}(\lambda) \), and the hyperparameter \( \lambda_i \) in the formula is further modeled as a gamma distribution

\[
p(\lambda) = \prod_{n=1}^{2\hat{K}} \text{Gamma}(\lambda_n|1, \varrho),
\]

(10)

where \( \varrho \) is usually set to \( \varrho = 0.01 \).

The irregular characteristics of the outlier matrix result in sparse elements, so each element of \( \hat{\mathbf{H}} \) needs to be assigned a Gaussian prior distribution with different variances:

\[
p(\hat{\mathbf{H}}|\Omega) = \prod_{n=1}^{2N_1} \prod_{l=1}^{L} \mathcal{CN}(h_{n,l}|0, \omega_{n,l}),
\]

(11)
where \(\omega_{n,l}\) is used to represent the variance of the element in the \(n\)-th row and \(l\)-th column of \(H\), and \([H]_{n,l} = h_{n,l}\) \([\Omega]_{n,l} = \omega_{n,l}\) Since \(\Omega\) is unknown, it is modeled as a gamma distribution

\[
p(\Omega) = \frac{2N_1}{\prod_{n=1}^{L} \prod_{l=1}^{N_1} \text{Gamma}(\omega_{n,l}|1, \varrho)}.
\] (12)

If the prior of \(H\) and \(S\) were constructed in the form of \(\hat{S}\), according to the Equation (6), it is expressed as

\[
p(\hat{S}|\Gamma) = \prod_{l=1}^{L} \text{CN}(\hat{s}_l|0_2(k+N_1)\times 1, \Gamma_l)
\] (13)

where \(\Gamma_l = \text{diag}(\lambda; \omega_l)\) and \(\omega_l\) represents the \(l\)-th column of \(\Omega\). According to the derivation process of SBL, the power \(\Gamma\) is updated in the last iteration. However, the sparsity of the elements of \(\text{barhmH}\) determines that its variance is three-dimensional, and each iteration is accompanied by two matrix inversion processes, which greatly affects the rate of the algorithm.

In contrast to the joint estimation method proposed in Equation (20), this paper makes use of the alternating SBL update algorithm proposed in [24], which aims to repair \(\hat{S}\) while updating the matrix \(\hat{H}\). This method reduces the dimensionality of the inverse matrix and thus the computing time to some extent. In this case, the \(\hat{S}\) and \(\hat{H}\) components of \(\hat{S}\) are estimated separately, and the changed data models are given separately

\[
\frac{QZ - Q\hat{H}}{Z} = QA\hat{S} + QN
\] (14)

\[
\frac{QZ - Q\hat{A}\hat{S}}{Z} = Q\hat{H} + QN
\] (15)

among them, \(Q\) is an orthogonal matrix whose characteristics are used to reduce the complexity of subsequent matrix operations.

4.2. Estimate Signal \(\hat{S}\)

According to the a priori hypothesis process of Equation (8) to Equation (10), and based on an Equation (14) model, the prior probability of \(Z\) when only the signal is considered is given as

\[
p(Z|\hat{S}, \beta) = \prod_{l=1}^{L} \text{CN}(z_l|A\hat{s}_l, \beta^{-1}I_{2N_1}).
\] (16)

The classical SBL-based DOA estimation algorithm [16] uses Bayesian formula to estimate the parameters of unknown signal matrix \(\hat{S}\) based on known observation data \(Z\) and reasonable priori assumptions. However, the estimation of \(\hat{S}\) is related to the estimation of the parameters \(\{\lambda, \beta\}\), which can be achieved by finding the corresponding parameter that makes \(P(\lambda, \beta|Z)\) maximal. This process can be simply expressed as

\[
\{\lambda^{\text{new}}, \beta^{\text{new}}\} = \arg \max_{\lambda, \beta} P(\lambda, \beta|Z)
\]

\[
= \arg \max_{\lambda, \beta} P(Z|\lambda, \beta).
\] (17)

The reason for the above transformation is that \(p(Z)\) is constant when the observation matrix \(Z\) is known, and the equation \(P(\lambda, \beta|Z)p(Z) = P(Z|\lambda, \beta)\) can be given by the conditional probability formula. Then, the posterior probability of \(\hat{S}\) can be obtained by using the Bayesian formula, i.e.,
\begin{align*}
p(S|Z; \lambda, \beta) &= \frac{p(Z(S, \beta))p(S|\lambda)}{\int p(Z(S, \beta))p(S|\lambda)dS} \\
&= \prod_{i=1}^{L} \mathcal{N}(s_i|\mu_i, \Sigma),
\end{align*}
(18)

where
\begin{align*}
\mu_i &= \beta \Sigma A^H z_i, \quad l = 1, 2, \ldots, L,
\end{align*}
(19)
\begin{align*}
\Sigma &= \Lambda - \Lambda A^H C^{-1} A \Lambda,
\end{align*}
(20)

where \(C = (\beta^{-1}I_{2N_l} + AA^H)\) [32].

According to the expectation maximisation (EM) algorithm, the following likelihood function is maximized to obtain the estimation of the corresponding parameters, which is
\begin{align*}
\mathcal{L}(\lambda, \beta) &= E(\ln(p(Z|S, \lambda, \beta)))_{p(S|Z, \lambda, \beta)} \\
&= E(\ln(p(Z|S, \beta)p(S|\lambda)p(\beta)))_{p(S|Z, \lambda, \beta)},
\end{align*}
(21)

For the updating of parameter \(\lambda\), the term associated with \(\beta\) is ignored. Then, Equation (21) can be rewritten as
\begin{align*}
E(\ln(p(S|\lambda)p(\lambda)))_{p(S|Z, \lambda, \beta)} \\
&= E(\ln(p(S|\lambda)))_{p(S|Z, \lambda, \beta)} + \log(p(\lambda)) \\
&= -\sum_{l=1}^{L} \ln|\Lambda_l| - \sum_{l=1}^{L} E(s_l^H \Lambda_l^{-1} s_l) - q \sum_{k=1}^{K} \lambda_k + \text{const}
\end{align*}
(22)

where \(\Delta_l = \|\mu_l\|^2 + \Sigma\). Similarly, when estimating the parameter \(\beta\), the other irrelevant terms are ignored. Then, Equation (21) can be simplified to
\begin{align*}
E(\ln(p(Z|S, \beta)p(\beta)))_{p(S|Z, \lambda, \beta)} \\
&= E(\ln(p(Z|S, \beta)))_{p(S|Z, \lambda, \beta)} + \ln(p(\beta)) \\
&= 2LN_l \ln(\beta) + (i-1) \ln(\beta) - \omega(\beta) \\
&- \beta \sum_{l=1}^{L} E(\|Z_l - AS\|^2_p) + \text{const}
\end{align*}
(23)

where \(\Xi = \sum_{j=1}^{L} \|Z_j - AS\|^2_p / L + \text{tr}(AA^H)\). Then, by taking the derivatives of Equations (22) and (23) with respect to \(\lambda\) and \(\beta\), respectively, and making them zero, the updating rules for \(\lambda\) and \(\beta\) can be obtained as follows:
\begin{align*}
\lambda^{\text{new}}_n &= \frac{\sqrt{L^2 + 4q \sum_{l=1}^{L} \langle \Delta_l \rangle_{n,n}} - L}{2q},
\end{align*}
(24)
\begin{align*}
\beta^{\text{new}} &= \frac{2N_l + (i-1)}{\langle \Xi \rangle + \omega/L}.
\end{align*}
(25)

After all, the modified signal \(\hat{S}\) can be obtained as
\begin{align*}
\hat{S} \equiv [\mu_1, \mu_2, \ldots, \mu_n].
\end{align*}
(26)
4.3. Alternating Estimate Outlier Matrix $\hat{H}$

When $\hat{H}$ is regarded as a signal, the prior probability of receiving data matrix is rewritten as

$$p(\mathbf{Z}|\hat{H}, \hat{\beta}) = \prod_{l=1}^{L} CN(\mathbf{z}_l|Q_1, \hat{\beta}^{-1}I_{N_1}). \quad (27)$$

Since $\mathbf{S}$ in Equation (14) is replaced by the updated $\mathbf{S}$ in equation Equation (26), we find that the noise variance has changed, which is denoted as $\hat{\beta}$ here.

Similar to the procedure from Equation (21) to Equation (23), the posterior probability of $\hat{H}$ with respect to data $\mathbf{Z}$ is deduced as

$$p(\hat{H}|\mathbf{Z}, \Omega) = \prod_{l=1}^{L} CN(h_l|\bar{\mathbf{p}}_l, \Sigma), \quad (28)$$

where

$$\bar{\mathbf{p}}_l = \hat{\beta}\Sigma Q^H \mathbf{z}_l, \quad l = 1, 2, \ldots, L,$$

$$\Sigma = (\hat{\beta}I_{N_1} + \text{diag}(\omega_1)^{-1})^{-1}. \quad (30)$$

Then, it can be easily obtained that the updating rule for parameters is

$$\hat{\beta}_{\text{new}} = \frac{2N_1L + (t - 1) \langle \hat{\varepsilon} \rangle}{\langle \hat{\varepsilon} \rangle + \sum_{n=1}^{L} \sum_{l=1}^{L} \nabla + \omega},$$

$$\omega_{n,l}^{\text{new}} = \frac{\sqrt{1 + 4\rho \nabla - 1}}{2\rho}, \quad (32)$$

where $\hat{\varepsilon} = \sum_{l=1}^{L} \| \mathbf{z}_l - Q\bar{\mathbf{p}}_l \|_2^2$, and $\nabla = (\hat{\beta}^{\text{new}} + (\omega_{n,l}^{\text{new}})^{-1})^{-1}$. Then, the updated $\hat{H}$ can be expressed as

$$\hat{H} \triangleq [\bar{\mathbf{p}}_1, \bar{\mathbf{p}}_2, \ldots, \bar{\mathbf{p}}_L]. \quad (33)$$

When the parameter exceeds the set threshold condition $\| \lambda^t - \lambda^{t-1} \|_2 / \| \lambda^{t-1} \|_2 < 10^{-4}$ or exceeds the set maximum number of iterations $T = 1500$, the update is stopped. Then, the rough estimation of DOA is obtained by searching for the first $K$ peaks of $\hat{S}$.

As can be seen from the spatial domain discretization model shown in Figure 1, the real signal source may not fall exactly on the divided grid points. The gap between true DOA and grid point is called the off-grid error. To achieve high-resolution DOA estimation, the influence of off-grid errors should be minimized. Thus, a fine one-dimensional search procedure around $K$ spectral peaks is used to obtain more accurate DOA estimation [33], which is specifically expressed as

$$\theta_{\text{new}} = \arg \max_{\theta \in \Theta_k} \left\{ \mathbf{H}^H(\theta) \mathbf{U}_k [h(\theta) \mathbf{h}^H(\theta) \mathbf{U}_k] \mathbf{z} \mathbf{z}^H \right. \left. - \mathbf{z}^H \mathbf{U}_k h(\theta) \mathbf{h}^H(\theta) \mathbf{U}_k \frac{\partial h(\theta)}{\partial \theta} \right\}, \quad (34)$$

where $h(\theta) = h(\theta)$. $[\cdot]_k$ represents a matrix without the column related to $\theta_k$ in matrix $[\cdot]$, and $\mathbf{U}_k = (\beta^{-1}I_{N_1} + \mathbf{A}_k \Lambda_{-k} \mathbf{A}_k^H)^{-1}$. $\Lambda_{-k}$ denotes the matrix after removing the row and column related to $\theta_k$ in the matrix $\Lambda$. Generally, $\Theta_k$ represents the peak search range of the $k$-th peak, and the search step is usually set to 0.01 or smaller.

Based on the above-mentioned derivations, the robust DOA estimation can be finally realized. The flow of the proposed algorithm is summarized as in Algorithm 1.
Algorithm 1 Robust DOA Estimation Scheme for Non-Circular Sources in Impulse Noise and Mutual Coupling.

1. Give the data receiving matrix $Z$ and the array flow matrix $A$ after eliminating the mutual coupling effect and expanding the array, and the orthogonal matrix $Q$ obtained by the singular value decomposition;
2. Initialize $\beta = \hat{\beta} = 0, \lambda = 0$, $\Omega = 0$, $S = 0$, $\bar{A} = QA, \varrho = 0.01, \omega = 0.001$ and step length of fine search $\delta = 0.01$;
3. while Convergence = 0 do
   4. The correction process of $\bar{S}$;
   5. Reconstruct $\bar{Z} = Q(Z - \bar{H})$ according to Equation (14);
   6. Estimate $\mu$ and $\Sigma$ by Equations (19) and (20);
   7. Update $\lambda^{\text{new}}$ and $\beta^{\text{new}}$ by Equations (24) and (25);
   8. Revise $\bar{S}$ according to Equation (26);
   9. The update process of $\bar{H}$;
   10. Reconstruct $Z = Q(Z - \bar{A}S)$ according to Equation (15);
   11. Estimate $\bar{\mu}$ and $\Sigma$ by Equation (29) to Equation (30);
   12. Update $\beta^{\text{new}}$ and $\Omega^{\text{new}}$ by Equations (31) and (32);
   13. Revise $\bar{H}$ according to Equation (33);
   14. end while
15. Obtain $\bar{S}$;
16. After the spatial spectrum search, a rough estimate of the $K$ peaks by Equation (34) to obtain a fine estimate of the signal source.

5. Algorithm Analysis

5.1. Analysis of the Maximum Identified Target Number

Under the signal model shown in Section 3 of this paper, the analysis of the maximum identified number of sources for the algorithm is as follows:

First, the left and right sides of Equation (6) are vectorized and expressed as

$$\text{vec}(Z) = \text{vec}(I \otimes B) \cdot \text{vec}(\tilde{S}) + \text{vec}(\bar{N}), \quad (35)$$

where $Z$ contains $2N_1L$ measurements and the dimension of $S$ after the null domain discretization is $2K^{*}L$ (where $2KL$ positions are large values compared to other positions). Since $\tilde{S}$ in the above equation is reconstructed from $S$ and $\bar{H}$, the outlier matrix $\bar{H}$ provides $E$ outliers to $\tilde{S}$—in summary, when satisfying

$$2(E + KL) \leq 2N_1L \quad (36)$$

Equation (36) has a consistent solution. Therefore, the maximum number of identified sources is $K \leq N_1 - (E/L)$ [24]. In theory, the array expansion shown in Equation (6) in Section 3 constructs a virtual array that doubles the array aperture. However, since the matrix $\Xi$ generated by the decoupling process shown in Equation (5) reconstructs the signal matrix $\tilde{S}$, the proposed scheme can only identify $K \leq N_1 - (E/L)$ sources.

5.2. Convergence Analysis

The technique used in this paper is a DOA estimation method based on alternating updates of the SBL. The core of the method is that an alternating iterative strategy is used to update the signal power $\bar{\mu}$ and the noise power $\bar{\mu}$. As with the conventional EM-based SBL estimation algorithm, there is good convergence. The specific analysis is presented in [34,35]. The simulation results below verify this property, where the error is set to

$$\text{Error} = ||\lambda^I - \lambda^{I-1}||_2. \quad (37)$$
The physical meaning of this error representation is the difference in signal power estimated from two adjacent iterations. Based on the convergence analysis diagram of the algorithm shown in Figure 2, it can be seen that the error curve tends to decrease asymptotically in one direction as the number of iterations increases. In addition, the trend continues to hold when the signal-to-noise ratio changes. What is clear is that all three curves shown in the graph flatten out and approach zero when the number of iterations is greater than 1000. This phenomenon indicates that the difference in power estimates at this point is small and can be considered constant. The iterative convergence of the algorithm is also demonstrated from a simulation perspective.

![Image of Figure 2](image-url)

**Figure 2.** Convergence of the iterative estimation process of the spatial power spectrum.

### 6. Simulation Results

To evaluate the performance of DOA estimation, we directly verify the effectiveness and robustness of the proposed DOA estimation method. Numerous simulation experiments are performed in this section. First of all, the physical meaning of some parameters and environmental settings is explained.

The impulse noise considered in this paper is described by the $\alpha$-stable distribution with probability density function (PDF), which is

$$\varphi(n) = e^{(-\gamma^\alpha |n|^\alpha + j\zeta(\tan(\pi/2\text{sign}(n))) + \varsigma n)}, \quad (38)$$

where $\alpha$ is the characteristic index. The smaller the value of $\alpha$, the more obvious the characteristic of impulse noise. When $\alpha = 2$, it satisfies the Gaussian property. $\gamma$ is the dispersion parameter, which is against the variance of Gaussian noise and the location parameter $\varsigma$. In symmetric $\alpha$-stable distribution, the symmetric parameters are both zero; then, the PDF is rewritten as

$$\varphi(n) = e^{(-\gamma^\alpha |n|^\alpha + j\varsigma n)}, \quad (39)$$

Since SNR is an invalid description for $\alpha < 2$, the generalized signal-to-noise ratio (GSNR) is used instead of SNR, which is defined by

$$\text{GSNR} = 10\log_{10} \frac{\delta_s^2}{\gamma^\alpha}, \quad (40)$$

where $\delta_s^2$ denotes signal power, and $\gamma^\alpha$ represents noise power.

In order to illustrate the superiority of the proposed algorithm, the RIM-SBLfast algorithm in [30], and the IM-SBLfast algorithm proposed in [24] are introduced to compare with the proposed algorithm. During simulations, unless otherwise specified, the uniform linear array (ULA) is composed of $N = 10$ array elements, and assume that there are $K = 2$ uncorrelated signal sources, whose direction is $\theta = [-10.56^\circ, 9.7^\circ]$. The mutual coupling coefficient is set to $c = [1, 0.9174 + j \times 0.3577]$, the number of snapshots is set to $L = 20$, GSNR = 10dB, $\alpha = 1.2$, and the grid interval is set to $r = 1^\circ$. In addition, the root mean
square error (RMSE) is used to evaluate the performance of different algorithms, which is denoted by

\[
\text{RMSE} = \sqrt{\frac{1}{M} \sum_{m=1}^{M} \sum_{k=1}^{K} (\theta_k - \theta_{m,k})^2}
\]

(41)

where \( \theta_k \) is the true angle of the \( k \)-th signal, and \( \theta_{m,k} \) is the corresponding estimated angle in the \( m \)-th Monte-Carlo experiment. \( M \) represents the total number of Monte Carlo experiments, which is \( M = 200 \) in this paper. Simulation results demonstrate the performance of the proposed algorithm from two aspects. Figure 3 illustrates the effectiveness of the proposed algorithm and Figures 4–7 illustrate the robustness of the algorithm.

![Figure 3](image)

**Figure 3.** The effectiveness of the proposed algorithm in DOA estimation under different conditions. (a) spectrum at GSNR = [0, 5, 10] dB; (b) spectrum at alpha = [1, 1.2, 2]; (c) spectrum at grid = [1, 2, 3].

As shown in Figure 3, this results shows that the proposed algorithm is effective to eliminate the mutual coupling and off-grid errors, and the performance of the proposed algorithm is stable under different conditions. Figure 3a shows that the proposed algorithm has relatively accurate DOA estimation when the GSNR changes from 0 dB to 10 dB. Moreover, in the process of increasing the GSNR, the main lobe of the signal spectrum is sharper and the side lobe is lower because the influence of noise is weaker and weaker.

![Figure 4](image)

**Figure 4.** Comparison of RMSE of different methods in terms of GSNR.

Figure 3b shows that, under the different alpha conditions, our proposed algorithm is effective for the identification and processing of impulse noise. Figure 3c is an experiment conducted under different grid intervals. These results show that the proposed algorithm can reduce the off grid error to a certain extent. However, when the grid interval is set too large, the efficiency of the proposed algorithm will be improved, but some estimation accuracy will be lost.

As shown in Figure 4, the RMSE of different algorithms in terms of GSNR are compared. In order to obtain the validity of the results, the value range of GSNR is set as from \(-10 \) dB to \( 20 \) dB with a step of \( 5 \) dB. The results show that the influence of noise gradually decreases with the increase of GSNR, hence the three curves show an obvious downward...
trend. In addition, the RMSE of our algorithm has always been smaller than RIM-SBLfast and IM-SBLfast. The advantage of our proposed algorithm is more obvious when the GSNR is higher. When GSNR > 15 dB because the signal is dominant and our algorithm can effectively handle the mutual coupling problem, the performance is also better than the comparison algorithms.

Figure 5 shows the performance comparison of different algorithms in terms of characteristic index \( \alpha \). In this simulation, the characteristic index \( \alpha \in [1.1, 2] \) and GSNR = 10 dB. The results show that, when \( \alpha \) is smaller, the pulse characteristics are more obvious, and our proposed algorithm has better performance with the increasing of \( \alpha \). However, on the whole, the performance of our algorithm is better than the comparison algorithms in the whole range of \( \alpha \).

![Figure 5. Comparison of RMSE of different methods in terms of Alpha.](image)

Last but not least, in sparse recovery-based DOA estimation algorithms, the high resolution of the DOA estimate often relies on how well the algorithm can attenuate the effects of off-grid errors. Figures 6 and 7 illustrate the performance of the proposed algorithm from different perspectives.

![Figure 6. Comparison of RMSE of different methods in terms of grid interval.](image)

Figure 6 shows the RMSE comparison of the proposed algorithm in terms of grid interval when the GSNR is 0 dB and 10 dB, respectively. Obviously, the proposed algorithm has a smaller RMSE when GSNR = 10 dB because the influence of impulse noise is significantly reduced in this case. However, with the increasing grid interval, DOA estimation performance of the proposed algorithm is obviously reduced. Hence, it is reasonable to set grid interval to 1° or 2°.

On the other hand, Figure 7 is the performance comparison of different algorithms in terms of grid interval. According to Figure 7, we can find that the estimation accuracy
of these three algorithms decreases with the increase of grid interval (the RMSE of the RIM-SBLfast algorithm decreases when grid = 3° compared to grid = 2°, which may be because of the fact that the grid falls just near the true signal). At the same time, it can be seen that our algorithm has the best performance among the other algorithms. This result in Figure 7 effectively illustrates the robustness of the proposed algorithm against off-grid error.

![Figure 7. Comparison of RMSE of different methods in terms of grid interval.](image)

7. Conclusions

In this paper, a robust DOA estimation scheme based on sparse Bayesian learning has been proposed for non-circular sources in impulse noise and mutual coupling. The proposed algorithm can effectively reduce the influence of array mutual coupling, and can extend the array aperture by using the non-circular property. On the other hand, the proposed algorithm can achieve accurate DOA estimation under impulsive noise while balancing the accuracy and efficiency of DOA estimation. Finally, the robustness and superiority of the proposed algorithm have been demonstrated through numerous simulations.

Author Contributions: Conceptualization, formal analysis, writing—original draft preparation, L.J.; Methodology, supervision, review and editing, X.W.; review and editing, X.L. and J.S. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the Important Science and Technology of Hainan Province (ZDKJ202010), the Key Research and Development Program of Hainan Province (No. ZDYF2020019), the National Natural Science Foundation of China (No. 61861015 and 61961013), the National Key Research and Development Program of China (No. 2019CXTD400, No. SQ2020YF0405680), the Young Elite Scientists Sponsorship Program by CAST (No. 2018QNRCC001), the National Natural Science Foundation of China (No. 62101165), and the Innovative Research Project of Postgraduates in Hainan Province (No. Qhys2021-47).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

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

References
1. Krim, H.; Viberg, M. Two decades of array signal processing research: The parametric approach. *IEEE Signal Process. Mag.* 1996, 13, 67–94. [CrossRef]
2. Schmidt, R. Multiple emitter location and signal parameter estimation. *IEEE Trans. Antennas Propag.* 1986, 34, 276–280. [CrossRef]
3. Wang, X.; Yang, L.T.; Meng, D.; Dong, M.; Ota, K.; Wang, H. Multi-UAV cooperative localization for marine targets based on weighted subspace fitting in SAGIN environment. *IEEE Internet Things J.* 2021. [CrossRef]
4. Cong, J.; Wang, X.; Yan, C.; Yang, L.; Dong, M.; Ota, K. CRB Weighted Source Localization Method Based on Deep Neural Networks in Multi-UAV Network. *IEEE Internet Things J.* 2022, in press. [CrossRef]
5. Wu, X.; Yan, J. A Second-Order Statistics-Based Mixed Sources Localization Method With Symmetric Sparse Arrays. *IEEE Commun. Lett.*, 2020, 24, 1695–1699. [CrossRef]

6. Porat, B.; Friedlander, B. Analysis of the asymptotic relative efficiency of the MUSIC algorithm. *IEEE Trans. Acoust. Speech Signal Process.* 1988, 36, 532–544. [CrossRef]

7. Kundu, D. Modified MUSIC algorithm for estimating DOA of signals—ScienceDirect. *Signal Process.* 1996, 48, 85–90. [CrossRef]

8. Roy, R.; Kailath, T. ESPRIT-estimation of signal parameters via rotational invariance techniques. *IEEE Trans. Acoust. Speech Signal Process.* 1989, 37, 984–995. [CrossRef]

9. Gao, F.; Gershman, A.B. A generalized ESPRIT approach to direction-of-arrival estimation. *IEEE Signal Process. Lett.* 2005, 12, 254–257. [CrossRef]

10. Dai, J.; Ye, Z. Sensor array processing based on subspace fitting. *IEEE Trans. Signal Process.* 1991, 39, 1110–1121. [CrossRef]

11. Stoica, P.; Sharman, K. Maximum likelihood methods for direction-of-arrival estimation. *IEEE Trans. Acoust. Speech Signal Process.* 1990, 38, 1132–1143. [CrossRef]

12. Stoica, P.; Gershman, A. Maximum-likelihood DOA estimation by data-supported grid search. *IEEE Signal Process. Lett.* 1999, 6, 273–275. [CrossRef]

13. Donoho, D.L. Compressed sensing. *IEEE Trans. Inf. Theory* 2006, 52, 1289–1306. [CrossRef]

14. Gerstoft, P.; Mecklenbruker, C.F.; Xenaki, A. Multi Snapshot Sparse Bayesian Learning for DOA Estimation. *IEEE Signal Process. Lett.* 2016, 23, 1469–1473. [CrossRef]

15. Dai, J.; Bao, X.; Xu, X.; Chang, C. Root Sparse Bayesian Learning for Off-Grid DOA Estimation. *IEEE Signal Process. Lett.* 2016, 24, 46–50. [CrossRef]

16. Dai, J.; Ye, Z.; Zhao, D. Robust Weighted Subspace Fitting for DOA Estimation via Block Sparse Recovery. *IEEE Commun. Lett.* 2020, 24, 563–567. [CrossRef]

17. Wang, X.; Wan, L.; Huang, M.; Shen, C.; Zhu, T. Low-Complexity Channel Estimation for Circular and Noncircular Signals in Virtual MIMO Vehicle Communication Systems. *IEEE Trans. Veh. Technol.* 2020, 69, 3916–3928. [CrossRef]