IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received October 10, 2018, accepted October 23, 2018, date of publication November 2, 2018, date of current version December 3, 2018.

Digital Object Identifier 10.1109/ACCESS.2018.2879386

A Simplified DOA Estimation Method Based on Correntropy in the Presence of Impulsive Noise

QUAN TIAN[1], TIANSHUANG QIU[1], JITONG MA[1], JINGCHUN LI[2], AND RONG LI[2]
[1]Faculty of Electronic Information and Electrical Engineering, Dalian University of Technology, Dalian 116024, China
[2]State Radio Monitoring Center, Beijing 100037, China

Corresponding author: Tianshuang Qiu (qiutsh@dlut.edu.cn)

This work was supported by the National Natural Science Foundation of China under Grant 61671105, Grant 61139001, Grant 61172108, and Grant 81241059.

ABSTRACT Many approaches have been studied in the field of array signal processing when impulsive noise is modeled with an alpha-stable distribution. By introducing the correntropy, which exhibits a robust statistics property, this paper defines a correntropy-based (COB) operator that provides a powerful mechanism to eliminate the detrimental effect of outliers in alpha-stable distributed noise environments. To improve computational efficiency, we apply the unitary transformation to the COB upper triangular Toeplitz matrix and construct a novel real-valued approximate estimation matrix with a MUSIC-like algorithm. On the basis of guaranteeing the accuracy of the direction of arrival estimation, the proposed algorithm also significantly reduces the computational complexity of the calculations. Comprehensive Monte Carlo simulation results demonstrate that the proposed algorithm is more robust than the existing algorithms in terms of the probability of resolution and root-mean-square-error, especially in the presence of highly impulsive noise or low generalized signal-to-noise ratio.

INDEX TERMS Alpha-stable distribution, correntropy, Toeplitz matrix, unitary transformation, MUSIC, DOA estimation.

I. INTRODUCTION

The problem of DOA estimation has been the focus of research for decades [1]–[5], and numerous effective methods have been proposed, for example, subspace-based high-resolution parameter estimation methods such as ESPRIT [6]–[8] and MUSIC [9]–[11], maximum-likelihood [12], [13] and conjugate augmented MUSIC [14] techniques including wireless communications, radar [15], sonar and microphone array systems. In the past, DOA estimation has been extensively investigated in the presence of additive noise that is assumed to be modeled by the Gaussian distribution, and a variety of algorithms has been proposed. However, many signals and noises encountered in practice are decidedly non-Gaussian, for example, low-frequency atmospheric noise, underwater acoustic signals and many types of human-made noises. Therefore, it is inappropriate to model the noise under Gaussian conditions. Recent studies also show that due to the influence of sea waves and mountains, the array outputs exhibit sharp spikes [16]. In [17], these classes of noises often result in significant performance degradation for systems optimized under the Gaussian assumption; that is, if the noise statistics deviate from the Gaussian distribution, a severe deterioration in performance may occur. As a consequence of heavy tails of probability density function (pdf), the alpha-stable distribution provides a more useful theoretical tool than the Gaussian distribution to describe the signal and noise sources mentioned above. As a variable parameter α controls the heaviness of tails, the alpha-stable distribution is a very flexible modeling tool. Unfortunately, no closed-form expression exists for the pdf of the alpha-stable distribution except for the Gaussian and Cauchy distributions. The alpha-stable distribution is generally defined by its characteristic function as

$$\varphi(t) = \exp\left\{\mathrm{j}\mu t - \gamma|t|^{\alpha}\left[1 + \mathrm{j}\beta\,\mathrm{sgn}(t)\omega(t,\alpha)\right]\right\} \tag{1}$$

where

$$\omega(t,\alpha) = \begin{cases} \tan\left(\dfrac{\pi\alpha}{2}\right) & \alpha \neq 1 \\ \dfrac{2}{\pi}\log|t| & \alpha = 1 \end{cases} \tag{2}$$

and

$$\mathrm{sgn}(t) = \begin{cases} 1 & t > 0 \\ 0 & t = 0 \\ -1 & t < 0 \end{cases} \tag{3}$$

for some real $\alpha, \beta, \gamma, \mu$ with

$$0 < \alpha \leq 2, \quad -1 \leq \beta \leq 1, \quad \gamma > 0, \quad -\infty < \mu < +\infty \tag{4}$$

α is the characteristic exponent, which measures the thickness of the tails of the distribution and is uniquely determined. β is a symmetry parameter and determines the sign and degree of asymmetry. $\beta = 0$ implies a distribution symmetric about μ. In this case, the distribution is called symmetric alpha-stable ($S\alpha S$). γ is a scale parameter, also called the dispersion. It is similar to the variance of the Gaussian distribution and can measure the degree of discretization of the data samples. μ is a location parameter. Furthermore, for the $S\alpha S$ distribution, μ represents the mean when $1 < \alpha \leq 2$ or the median when $0 < \alpha < 1$. An $S\alpha S$ distribution is said to be standard if $\gamma = 1, \mu = 0$.

Note that $S\alpha S$ degenerates into the Gaussian distribution when $\alpha = 2$. Due to the infinite covariance of the $S\alpha S$ distribution, the conventional DOA estimation algorithms based on the second-order statistics degrade dramatically in their performance under impulsive noise environments [18]. For details about the $S\alpha S$ distribution, see [19] and the references therein.

One idea for improving the performance of the conventional MUSIC algorithms based on second-order statistics is to adopt the fractional lower-order statistics instead of the second-order covariance, e.g., FLOM-MUSIC [20]. Fractional lower-order statistics-based methods are robust against outliers, but large sample sizes are required to achieve a satisfactory performance [22]. The fractional lower-order cyclic statistics can effectively restrain both the impulsive noise and the co-channel interference [21].

In addition, there is a class of impulsive noise robust algorithms introduced within different applications based on the ℓ_p-norm as the fidelity criterion. For example, the ℓ_p-norm is used via ADMM for robust matrix completion [23]. To overcome the drawbacks of the two-step method of time delay estimation (TDE), such as error accumulation, estimation bias and nonrobust impulsive noise or outliers, [24] proposed the ℓ_p-norm minimization algorithm. Moreover, an ℓ_p-ℓ_1 minimization approach named Lp-ADM was proposed in [25], in which ℓ_p-norm was used for error penalization and ℓ_1-norm was employed for sparsity. In [26], a new channel estimation and an equalizer design criterion based on the minimization of the absolute error function with an ℓ_1-norm regularization term were proposed, which involved the ℓ_1-regularized least-absolutes (ℓ_1-LA) and the linear least-absolutes (LLA) algorithms. By using the ℓ_1-modulus of complex numbers instead of the standard modulus, the channel estimation and equalization problem is recast into linear programming that can be efficiently solved.

The zero-memory non-linearity (ZMNL) [27] based MUSIC methods were proposed for limiting the influence of impulsive noise by clipping the amplitude of the received signals. ZMNL may provide more accurate DOA estimation than the fractional lower-order based methods. However, its performance might degrade with an increase in the rank of the signal subspace.

The unitary transformation method [28] was presented to transform the complex covariance matrix of an equally spaced linear array into a real symmetric matrix. Since the computations of performing the unitary transformation are much less than those of performing the eigenvalue decomposition and spectral peak search, this method saves a significant number of calculations. The unitary transformation was also applied to spherical arrays in [29] to achieve a lower computational cost and better performance.

In [30], the correntropy was proposed as a new statistic that can quantify both the statistical distribution and the time structure of two random processes. The correntropy, which contains higher-order statistics, can extract more information than the correlation function and provide a robust method to suppress the influence of outliers. By virtue of correntropy, [37] defined a robust correlation criterion, the correntropy-based correlation (CRCO), to estimate DOA by MUSIC. It introduced correntropy as an adaptive factor into the covariance matrix to reduce the effect of impulsive noise.

In this paper, we focus on passive radio monitoring and signal source target location and address the DOA estimation problems of robust subspace under impulsive noise environments including the scenarios where the bursts or spikes generally exhibit in the array signals from the perspective of correntropy. The main contributions of this work are summarized as follows:

- An effective estimation operator is derived, and a novel correntropy-based MUSIC algorithm is proposed to estimate DOAs in the presence of impulsive noise. The proposed algorithm can significantly improve the estimation performance in terms of resolution and accuracy. We also prove the boundedness of this operator.
- To reduce the computational complexity of the MUSIC-like algorithm, we construct a Toeplitz matrix and define a new real-valued estimation matrix, which maintains a low level of computational complexity but provides reasonable DOA estimation performance.

Consider CRCO-MUSIC and ZMNL-MUSIC algorithms; although CRCO applies correntropy to suppress impulsive noise, we rederive a new estimation matrix that is more robust than the estimation matrix used by CRCO. ZMNL handles heavy-tailed noise by passing it through a zero-memory nonlinearity; however, although the algorithm is more immune to impulsive noise than the fractional lower-order statistics such as FLOM and phased fractional lower-order moment (PFLOM) [15], it has lower noise suppression ability than correntropy. The proposed algorithm not only achieves better DOA estimation performance but also has lower computational complexity.

The remainder of this paper is organized as follows. After a brief introduction of the signal model of the sensor array outputs, Section II describes the definition and some

essential properties of the correntropy of independent and identical distribution (i.i.d.) complex $S\alpha S$ random variables. To suppress the impulsive noise and simplify the solution, Section III defines the correntropy-based (COB) operator and derives several properties of the COB. Furthermore, we present the difference between MSE and COB and discuss the advantage of correntropy. On this basis, we construct the COB Toeplitz matrix. Considering the unitary transformation, we define an approximate estimation matrix that is employed to substitute the covariance matrix of the conventional MUSIC-like algorithm for DOA estimation in Section IV. Finally, we demonstrate simulation results in Section V, and Section VI summarizes the conclusions and future works.

The following notations are used throughout this paper. The superscripts $'*'$, $'T'$ and $'H'$ denote the conjugate, transpose and conjugate transpose, respectively. $\mathbf{E}[\cdot]$, $\Re(\cdot)$ and $\Im(\cdot)$ represent the mathematical expectation and real and imaginary part operators, respectively.

II. PROBLEM FORMULATION

A. SIGNAL MODEL

Consider L narrow-band independent, complex isotropic incoherent sources with a known center frequency ω and locations $\{\theta_1, \theta_2, \cdots, \theta_L\}$ impinging on a uniform linear array (ULA) with M sensors. In the following, we assume that the underlying noises are i.i.d. complex $S\alpha S$ random processes with the same characteristic exponent α and that the real and imaginary parts of these noises are statistically independent of one another [31] both along the array sensors and over time. In the subspace-based algorithms, the number of signal sources needs to be less than the number of sensors, that is, $L < M$. As the propagation delay across the array is much smaller than the reciprocal of the signal bandwidth, the complex envelope of the array output at the ith sensor can be employed to estimate the parameters in the following model:

$$x_i(t) = \sum_{k=1}^{L} a_i(\theta_k) s_k(t) + n_i(t), \quad i = 1, 2, \cdots, M \tag{5}$$

where $a_i(\theta_k) = \mathrm{e}^{-\mathrm{j}\frac{2\pi}{\lambda}(i-1)d\sin(\theta_k)}$ is the steering coefficient of the ith sensor toward direction θ_k, the distance between adjacent sensors is d, λ is the wavelength of the carrier, $s_k(t)$ is the kth signal received at the ith sensor and $n_i(t)$ is the additive noise at the ith sensor.

Equation (5) can be rewritten in a compact form as

$$\boldsymbol{x}(t) = \boldsymbol{A}(\boldsymbol{\theta})\boldsymbol{s}(t) + \boldsymbol{n}(t) \tag{6}$$

where we have the following:

- $\boldsymbol{x}(t) = [x_1(t), x_2(t), \cdots, x_M(t)]^T$ is the $M \times 1$ vector of the signals received by the sensor array.
- $\boldsymbol{s}(t) = [s_1(t), s_2(t), \cdots, s_L(t)]^T$ is the $L \times 1$ vector of the complex signal amplitudes for the sources.
- $\boldsymbol{A}(\boldsymbol{\theta}) = [\boldsymbol{a}(\theta_1), \boldsymbol{a}(\theta_2), \cdots, \boldsymbol{a}(\theta_L)]$ is the $M \times L$ matrix of the array steering vector, and each vector $\boldsymbol{a}(\theta_k)$ has the following special structure: $\boldsymbol{a}(\theta_k) = [1, \mathrm{e}^{-\mathrm{j}\frac{2\pi}{\lambda}d\sin(\theta_k)}, \cdots, \mathrm{e}^{-\mathrm{j}\frac{2\pi}{\lambda}(M-1)d\sin(\theta_k)}]^T$.
- $\boldsymbol{n}(t) = [n_1(t), n_2(t), \cdots, n_M(t)]^T$ is the $M \times 1$ vector of additive measurement noise.

$\boldsymbol{x}(t)$ and $\boldsymbol{n}(t)$ are multivariate independent and zero-mean; $\boldsymbol{n}(t)$ is assumed to be spatially uncorrelated. The problem addressed in this paper is estimating the DOAs $\{\theta_1, \theta_2, \cdots, \theta_L\}$ of the sources from the N snapshots of the array outputs. The number of sources L is assumed to be known.

B. CORRENTROPY AND ITS PROPERTIES

Inspired by information theoretic learning, based on the viewpoint of kernel methods, the correntropy has been proposed as a novel robust statistic. Ref. [30] shows that correntropy is directly related to the probability of how similar two random variables are in a neighborhood of the joint space controlled by the kernel size. Furthermore, the correntropy also provides a powerful mechanism to eliminate the detrimental effect of outliers.

By extending the fundamental definition of the correlation function for random processes with a generalized correlation function and containing higher-order moments of the pdf, the correntropy has been employed in problems [34]–[36], including the DOA estimation problems [37]–[40].

Refs. [30] and [32] define the correntropy between two arbitrary random variables X and Y as follows

$$\boldsymbol{V}_\sigma(X, Y) = \mathbf{E}\left[\kappa_\sigma(X - Y)\right] = \int \kappa_\sigma(x - y) dF_v(x, y) \tag{7}$$

where $\kappa_\sigma(\cdot)$ is a shift-invariant kernel that satisfies Mercer's theorem and $F_v(\cdot, \cdot)$ denotes the joint distribution function of (X, Y) in [33].

However, for some cases of practical interest, since we do not know the joint pdf and only use a finite amount of data$\{(x_i, y_i)\}_{i=1}^N$, the sample estimator of correntropy [30] can be obtained through

$$\hat{\boldsymbol{V}}_\sigma(X, Y) = \frac{1}{N}\sum_{i=1}^{N} \kappa_\sigma(x_i - y_i) \tag{8}$$

The Gaussian kernel is generally chosen as the kernel function that can be expressed as

$$\kappa_\sigma(x_i - y_i) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x_i - y_i)^2}{2\sigma^2}\right), \quad i = 1, 2, \cdots, N \tag{9}$$

where $\sigma > 0$ denotes the kernel size, and N is the number of the data points. The kernel size can be determined by using Silverman's rule or maximum-likelihood concepts.

It is worth noting that by utilizing a Taylor series expansion with the Gaussian kernel, the correntropy can be formed as

$$\boldsymbol{V}_\sigma(X, Y) = \frac{1}{\sqrt{2\pi}\sigma}\sum_{n=0}^{\infty}\frac{(-1)^n}{2^n n!}\mathbf{E}\left[\frac{(X - Y)^{2n}}{\sigma^{2n}}\right] \tag{10}$$

which involves all the even-order moments of the random variable $X-Y$. Consequently, it is beneficial for nonlinear and non-Gaussian signal processing. Equation (10) also indicates that the correntropy contains the conventional covariance function, and kernel size controls the weights of higher-order moments. However, from (8), it is easy to see that the correntropy is much simpler to estimate directly from data samples than the conventional moments.

In 2D sample space, correntropy induces a metric named the correntropy induced metric (CIM). Considering two vectors $X = [x_1, x_2, \cdots, x_n]^T$ and $Y = [y_1, y_2, \cdots, y_n]^T$, the CIM function is defined by

$$\mathrm{CIM}(X, Y) = \sqrt{\kappa_\sigma(0) - \boldsymbol{V}_\sigma(X, Y)} \tag{11}$$

the contours of CIM that measure the distance from the origin to a point (x_i, y_i), $i = 1, 2, \cdots, n$, are shown in Fig. 1.

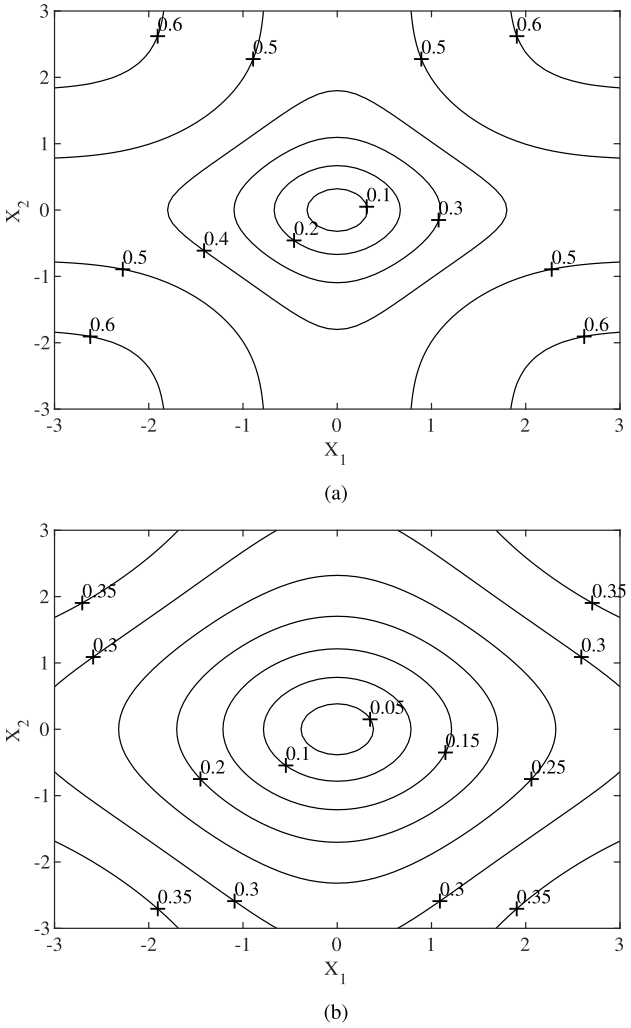

FIGURE 1. Contours of CIM in 2D sample space with different kernel size: (a) $\sigma = 1$; (b) $\sigma = 1.8$.

CIM is translation invariant and not homogeneous, so it does not induce a single norm. We can also observe some interesting facts from Fig. 1: when the point is close to the origin, CIM behaves similarly to an L2 norm called the Euclidean zone; when the point is apart from the origin, CIM behaves similarly to an L1 norm; and if the point is farther from the origin, CIM behaves similarly to an L0 norm. These findings demonstrate that the correntropy is a local similarity measure and more robust to outliers and the scale of CIM can be controlled by the kernel size.

Additionally, we list four main properties of CIM with complex joint $S\alpha S$ random variables X and Y:

- CIM is nonnegative: $\mathrm{CIM}(X, Y) \geq 0$.
- CIM is symmetric: $\mathrm{CIM}(X, Y) = \mathrm{CIM}(Y, X)$.
- CIM reaches its minimum if and only if $X = Y$, then $\mathrm{CIM}(X, Y) = 0$.
- CIM satisfies the triangle inequality: $\mathrm{CIM}(X, Z) \leq \mathrm{CIM}(X, Y) + \mathrm{CIM}(Y, Z)$ if the random variable Z has the same distribution as X or Y.

III. PROPOSED SOLUTION

We define a novel operator based on the correntropy to efficiently suppress outliers in the presence of impulsive noise.

A. CORRENTROPY-BASED OPERATOR

1.Definition

Considering two i.i.d. $S\alpha S$ random variables X and Y, the COB operator V_T is defined by

$$\boldsymbol{V}_T = \mathbf{E}\left[\kappa_\sigma(X - Y)(X - Y)^2\right] \tag{12}$$

2.Properties

Some important properties of the COB operator are presented next. The first three are obtained after straightforward derivations and will therefore not be proved here. The proof of the forth property is shown in APPENDIX A.

1) COB operator is symmetric: $V_T(X, Y) = V_T(Y, X)$.
2) COB operator is centrosymmetric: $V_T(X, Y) = V_T(-X, -Y)$.
3) COB operator is nonnegative for real signals: $V_T \geq 0$. It is easy to see that $\kappa_\sigma(X - Y) > 0$ and $(X - Y)^2 \geq 0$, so the COB operator reaches its minimum if and only if $X = Y$.
4) COB operator is bounded.

B. COMPARISON BETWEEN MEAN SQUARE ERROR (MSE) AND COB

Considering two random variables X and Y, the residual error is defined by $\varepsilon = X - Y$. The MSE can be expressed as

$$\begin{aligned}\mathrm{MSE}\,(X, Y) &= \mathbf{E}[(X - Y)^2] \\ &= \iint (x - y)^2 f_{XY}(x - y)dxdy \\ &= \int \varepsilon^2 f\,(\varepsilon)\,d\varepsilon\end{aligned} \tag{13}$$

and

$$\begin{aligned}\boldsymbol{V}_T\,(X, Y) &= \mathbf{E}[\kappa_\sigma(X - Y)(X - Y)^2] \\ &= \iint \kappa_\sigma(x - y)(x - y)^2 f_{XY}(x - y)dxdy \\ &= \int \kappa_\sigma(\varepsilon)\varepsilon^2 f(\varepsilon)d\varepsilon\end{aligned} \tag{14}$$

From (13), we can observe that the MSE is a second-order statistic in the joint space with a valley along the line $y = x$. Since the MSE quantifies the differences between x and y and is quadratically increased for values away from this line, it can be employed as a similarity measure in the joint space.

However, the MSE reflects the influence of global error and can provide optimality for the Gaussian noise. If the underlying noise is modeled with the alpha-stable distribution that has outliers, the impulsive residuals will be accumulated in the MSE. This is why other distributions will make the MSE nonoptimal, especially if the measurement noise contains impulsive components; then, the noise has a nonzero mean or is nonsymmetric.

In (14), we employ the exponential kernel function that can be expressed as

$$\kappa_\sigma(\varepsilon) = \exp\left(-\frac{|\varepsilon|}{2\sigma^2}\right) \tag{15}$$

Assume ε contains impulsive components; as ε increases, the $\kappa_\sigma(\varepsilon)$ decay quickly, and if ε tends to $+\infty$, the $\kappa_\sigma(\varepsilon)$ tends to zero. Consequently, the outliers will not tend to dominate, and the COB can obtain better performance than the MSE.

Now, we can summarize that the COB operator introduces an adaptive weight factor into the second-order statistics. By using the kernel size as a zoom adjuster of the data samples and decreasing the influence of impulsive noises exponentially, the COB operator improves the performance of the parameter estimation on the basis of the boundedness of the second-order statistics.

C. COB TOEPLITZ MATRIX

In the $m \times m$ square case, the Toeplitz matrix has the following structure:

$$\boldsymbol{T} = \begin{bmatrix} b_{1,1} & b_{1,2} & \cdots & b_{1,m} \\ b_{2,1} & b_{1,1} & \ddots & \vdots \\ \vdots & \ddots & \ddots & b_{1,2} \\ b_{m,1} & \cdots & b_{2,1} & b_{1,1} \end{bmatrix}_{m\times m} \tag{16}$$

Accordingly, the corresponding COB Toeplitz matrix $\boldsymbol{Q}$ of the sensor array outputs can be constructed, and the (i, j)th entry $Q_{i,j}$ can be expressed as

$$Q_{i,j} = \mathbf{E}\left[\kappa_\sigma(x_i(t) - x_j^*(t))(x_i(t) - x_j^*(t))^2\right] \tag{17}$$

where $x_i(t)$ and $x_j(t)$ are the ith and jth components, respectively, of the vector $\boldsymbol{x}(t)$.

Furthermore, consider the characteristics of the received signal. If the array outputs $\{x_i(t)\}_{i=1}^M$ are real-valued, the $Q_{i,j}$ will be zero for $i = j$; that is, the singular value decomposition (SVD) of $\boldsymbol{Q}$ does not exist. To solve this problem, we introduce a weight factor ξ and further generalize the definition of (17) as

$$Q'_{i,j} = \mathbf{E}\left[\kappa_\sigma(x_i(t) - x_j^*(t))(x_i(t) - \xi x_j^*(t))^2\right] \tag{18}$$

where ξ is a given positive constant. Here, we apply the parameter ξ to ensure that the estimation matrix $\boldsymbol{Q}'$ satisfies the semipositive definite conditions. It is easy to determine that when the array outputs are complex-valued, (18) is equivalent to (17) in the case of $\xi = 1$. When the array outputs are real-valued, the condition $\xi \neq 1$ can guarantee the estimation matrix to avoid a nonpositive definite case. Therefore, (18) is more robust than (17).

To simplify the derivation of the expression given in the following equations, we define $\kappa_\sigma^{(i,j)} := \kappa_\sigma(x_i, x_j^*)$; moreover, to reduce the high computational cost, the $M \times M$ estimation matrix $\boldsymbol{Q}'$ is triangulated, and the upper triangular Toeplitz matrix can be defined in (19), which appears at the top of the next page.

IV. MUSIC-LIKE SUBSPACE ALGORITHM

A. REAL-VALUED TRANSFORMATION METHOD

In this section, we propose a novel method based on the MUSIC-like algorithm according to the fact that it is formulated regarding real-valued computations within.

The $p \times p$ exchange matrix is defined as follows

$$\boldsymbol{J} = \begin{bmatrix} 0 & 0 & \cdots & 1 \\ 0 & 0 & 1 & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 1 & 0 & \cdots & 0 \end{bmatrix}_{p\times p} \tag{20}$$

We define the following sparse matrices for an even or odd number sensor array outputs.

If M is an even value, we define

$$\boldsymbol{U} = \frac{1}{\sqrt{2}}\begin{bmatrix} \boldsymbol{I} & \mathrm{j}\boldsymbol{I} \\ -\boldsymbol{J} & \mathrm{j}\boldsymbol{J} \end{bmatrix} \tag{21}$$

where $\boldsymbol{J}$ and $\boldsymbol{I}$ are the $M/2 \times M/2$ exchange and identity matrices, respectively. It can easily be shown that $\boldsymbol{U}$ satisfies $\boldsymbol{U}^H\boldsymbol{U} = \boldsymbol{U}\boldsymbol{U}^H = \boldsymbol{I}$. However, we can also set M to be odd; in this case, the transformation matrix can be defined by

$$\boldsymbol{U} = \frac{1}{\sqrt{2}}\begin{bmatrix} \boldsymbol{I} & \boldsymbol{0} & \mathrm{j}\boldsymbol{I} \\ \boldsymbol{0}^T & \sqrt{2} & \boldsymbol{0}^T \\ -\boldsymbol{J} & \boldsymbol{0} & \mathrm{j}\boldsymbol{J} \end{bmatrix} \tag{22}$$

where $\boldsymbol{0} = (0, 0, \cdots, 0)^T$ is the $(M-1)/2 \times 1$ vector, $\boldsymbol{J}$ and $\boldsymbol{I}$ are of size $(M-1)/2 \times (M-1)/2$. Moreover, without loss of generality, let us assume that M is even.

In this section, we will briefly review the unitary MUSIC algorithm: firstly, a unitary matrix $\boldsymbol{U}_0$ is defined, e.g., $\boldsymbol{U}_0 = (1/\sqrt{2})[\boldsymbol{I}, \boldsymbol{J}; \mathrm{j}\boldsymbol{J}, -\mathrm{j}\boldsymbol{I}]$ [28], and it satisfies ${\boldsymbol{U}_0}^{-1} = {\boldsymbol{U}_0}^H$ and ${\boldsymbol{U}_0}^*\boldsymbol{J} = \boldsymbol{U}_0$. Secondly, for any $M \times M$ Hermitian persymmetric matrix $\boldsymbol{G}$, $\boldsymbol{U}_0\boldsymbol{G}{\boldsymbol{U}_0}^H$ is real and symmetric, i.e.,

$$\begin{aligned}\left(\boldsymbol{U}_0\boldsymbol{G}{\boldsymbol{U}_0}^H\right)^* &= {\boldsymbol{U}_0}^*\boldsymbol{G}^*{\boldsymbol{U}_0}^{\mathrm{T}} \\ &= {\boldsymbol{U}_0}^*\boldsymbol{J}\boldsymbol{J}\boldsymbol{G}^*\boldsymbol{J}\boldsymbol{J}{\boldsymbol{U}_0}^{\mathrm{T}} \\ &= \boldsymbol{U}_0\boldsymbol{G}{\boldsymbol{U}_0}^H\end{aligned} \tag{23}$$

$$\boldsymbol{R}_T = \begin{bmatrix} \mathbf{E}\left[\kappa_\sigma^{(1,1)}\left(x_1 - \xi x_1^*\right)^2\right] & \mathbf{E}\left[\kappa_\sigma^{(1,2)}\left(x_1 - \xi x_2^*\right)^2\right] & \cdots & \mathbf{E}\left[\kappa_\sigma^{(1,M)}\left(x_1 - \xi x_M^*\right)^2\right] \\ 0 & \mathbf{E}\left[\kappa_\sigma^{(1,1)}\left(x_1 - \xi x_1^*\right)^2\right] & \ddots & \vdots \\ \vdots & \ddots & \ddots & \mathbf{E}\left[\kappa_\sigma^{(1,2)}\left(x_1 - \xi x_2^*\right)^2\right] \\ 0 & \cdots & 0 & \mathbf{E}\left[\kappa_\sigma^{(1,1)}\left(x_1 - \xi x_1^*\right)^2\right] \end{bmatrix} \tag{19}$$

Furthermore, $\boldsymbol{U}_0\boldsymbol{G}\boldsymbol{U}_0{}^H$ is Hermitian since $\boldsymbol{G}$ is Hermitian. With this theorem, we can form a real symmetric transformed covariance matrix $\boldsymbol{U}_0\boldsymbol{R}\boldsymbol{U}_0{}^H$. It is well known that the eigenvalues and eigenvectors of a real-valued matrix are real, so calculating the eigencomponents of $\boldsymbol{U}_0\boldsymbol{G}\boldsymbol{U}_0{}^H$ requires real compilations only. We can derivate that if we define the real-valued eigenvectors and eigenvalues as $\boldsymbol{\eta}_i'$ and λ_i' of $\boldsymbol{U}_0\boldsymbol{G}\boldsymbol{U}_0{}^H$, the following relations can be obtained as

$$\boldsymbol{\eta}_i' = \boldsymbol{U}_0\boldsymbol{\eta}_i \tag{24}$$

and

$$\lambda_i' = \lambda_i \tag{25}$$

where $\boldsymbol{\eta}_i$ and λ_i $i = 1, 2, \cdots, M$ are the eigenvectors and eigenvalues of $\boldsymbol{R}$, respectively. Therefore, the signal and noise subspaces can be constructed from both (24) and (25). Finally, after performing the SVD and searching for the spectral peak, the DOAs of the signal sources can be obtained.

To reduce the computational complexity, the solution we adopted is converting the complex-valued estimation matrix to a real-valued matrix. Due to the unitary matrix $\boldsymbol{U}$ defined by (21), there is $\boldsymbol{U}^*\boldsymbol{J} \neq \boldsymbol{U}$; it can be seen that $\boldsymbol{U}$ does not satisfy (23). That is, $\boldsymbol{U}\boldsymbol{R}_T\boldsymbol{U}^H$ is a complex-valued matrix. Therefore, we define the real part of $\boldsymbol{U}^H\boldsymbol{R}_T\boldsymbol{U}$ as a new approximate estimation matrix

$$\boldsymbol{C}_T := \Re\left(\boldsymbol{U}^H\boldsymbol{R}_T\boldsymbol{U}\right) \tag{26}$$

which can be obtained via the forward COB Toeplitz matrix $\boldsymbol{R}_T$. This new estimation matrix reduces the computational cost by at least a factor of four without sacrificing accuracy, and our previous works proved that the matrix $\boldsymbol{U}$ performs better than the unitary matrix involved in [28].

As $\boldsymbol{R}_T$ is an upper triangular Toeplitz matrix, we can divide it into four $M/2 \times M/2$ submatrices: $\boldsymbol{A}$, $\boldsymbol{B}$, $\boldsymbol{C}$ and $\boldsymbol{0}$, which are shown as follows:

$$\boldsymbol{R}_T = \begin{bmatrix} \boldsymbol{A} & \boldsymbol{C} \\ \boldsymbol{0} & \boldsymbol{B} \end{bmatrix} \tag{27}$$

in which $\boldsymbol{A}$ and $\boldsymbol{B}$ are also upper triangular matrices.

By replacing the expression of $\boldsymbol{U}$ with (21), it follows that

$$\begin{aligned} \boldsymbol{U}^H\boldsymbol{R}_T\boldsymbol{U} &= \frac{1}{2}\begin{bmatrix} \boldsymbol{I} & \mathrm{j}\boldsymbol{I} \\ -\boldsymbol{J} & \mathrm{j}\boldsymbol{J} \end{bmatrix}^H \begin{bmatrix} \boldsymbol{A} & \boldsymbol{C} \\ \boldsymbol{0} & \boldsymbol{B} \end{bmatrix} \begin{bmatrix} \boldsymbol{I} & \mathrm{j}\boldsymbol{I} \\ -\boldsymbol{J} & \mathrm{j}\boldsymbol{J} \end{bmatrix} \\ &= \frac{1}{2}\begin{bmatrix} \boldsymbol{A} & \boldsymbol{C} - \boldsymbol{J}\boldsymbol{B} \\ -\mathrm{j}\boldsymbol{A} & -\mathrm{j}\boldsymbol{C} - \mathrm{j}\boldsymbol{J}\boldsymbol{B} \end{bmatrix} \begin{bmatrix} \boldsymbol{I} & \mathrm{j}\boldsymbol{I} \\ -\boldsymbol{J} & \mathrm{j}\boldsymbol{J} \end{bmatrix} \\ &= \frac{1}{2}\begin{bmatrix} \boldsymbol{A} - \boldsymbol{C}\boldsymbol{J} + \boldsymbol{J}\boldsymbol{B}\boldsymbol{J} & \mathrm{j}\boldsymbol{A} + \mathrm{j}\boldsymbol{C}\boldsymbol{J} - \mathrm{j}\boldsymbol{J}\boldsymbol{B}\boldsymbol{J} \\ -\mathrm{j}\boldsymbol{A} + \mathrm{j}\boldsymbol{C}\boldsymbol{J} + \mathrm{j}\boldsymbol{J}\boldsymbol{B}\boldsymbol{J} & \boldsymbol{A} + \boldsymbol{C}\boldsymbol{J} + \boldsymbol{J}\boldsymbol{B}\boldsymbol{J} \end{bmatrix} \end{aligned} \tag{28}$$

The matrix $\boldsymbol{R}_T$ has an upper triangular Toeplitz structure; thus, it is easy to determine that the submatrices satisfy the condition: $\boldsymbol{A} = \boldsymbol{B}$. Therefore, we can simplify $\boldsymbol{R}_T$ by replacing $\boldsymbol{B}$ with $\boldsymbol{A}$; then, (28) can be rewritten as follows:

$$\begin{aligned} &\boldsymbol{U}^H\boldsymbol{R}_T\boldsymbol{U} \\ &\quad = \frac{1}{2}\begin{bmatrix} \boldsymbol{A} - \boldsymbol{C}\boldsymbol{J} + \boldsymbol{J}\boldsymbol{A}\boldsymbol{J} & \mathrm{j}\boldsymbol{A} + \mathrm{j}\boldsymbol{C}\boldsymbol{J} - \mathrm{j}\boldsymbol{J}\boldsymbol{A}\boldsymbol{J} \\ -\mathrm{j}\boldsymbol{A} + \mathrm{j}\boldsymbol{C}\boldsymbol{J} + \mathrm{j}\boldsymbol{J}\boldsymbol{A}\boldsymbol{J} & \boldsymbol{A} + \boldsymbol{C}\boldsymbol{J} + \boldsymbol{J}\boldsymbol{A}\boldsymbol{J} \end{bmatrix} \end{aligned} \tag{29}$$

$$\begin{aligned} \boldsymbol{C}_T &= \frac{1}{2}\begin{bmatrix} \Re(\boldsymbol{A} + \boldsymbol{A}^T - \boldsymbol{C}\boldsymbol{J}) & \Re(\mathrm{j}\boldsymbol{A} + \mathrm{j}\boldsymbol{C}\boldsymbol{J} - \mathrm{j}\boldsymbol{A}^T) \\ \Re(-\mathrm{j}\boldsymbol{A} + \mathrm{j}\boldsymbol{C}\boldsymbol{J} + \mathrm{j}\boldsymbol{A}^T) & \Re(\boldsymbol{A} + \boldsymbol{A}^T + \boldsymbol{C}\boldsymbol{J}) \end{bmatrix} \\ &= \frac{1}{2}\begin{bmatrix} \Re(\boldsymbol{A}) + \Re(\boldsymbol{A}^T) - \Re(\boldsymbol{C}\boldsymbol{J}) & -\Im(\boldsymbol{A}) - \Im(\boldsymbol{C}\boldsymbol{J}) + \Im(\boldsymbol{A}^T) \\ \Im(\boldsymbol{A}) - \Im(\boldsymbol{C}\boldsymbol{J}) - \Im(\boldsymbol{A}^T) & \Re(\boldsymbol{A}) + \Re(\boldsymbol{A}^T) + \Re(\boldsymbol{C}\boldsymbol{J}) \end{bmatrix} \end{aligned} \tag{30}$$

whose real parts can be expressed by (30).

To build the estimation matrix under the assumption of the signal model in (6), we have to compute an $M \times M$ matrix. However, if we want to achieve the proposed estimation matrix $\boldsymbol{C}_T$ in (30), we only need to calculate two $M/2 \times M/2$ matrices $\boldsymbol{A}$ and $\boldsymbol{C}$, which involve merely computing $M/2$ and $M - 1$ entries, respectively.

Comparing the computations of performing the SVD and searching for the spectral peak, as multiplication occupies most of the computational load and a complex-valued multiplication requires four real multiplications and three real additions, whatever the construction of covariance matrix, SVD or spectral peak search, the proposed algorithm will save a considerable number of computations.

B. THE IMPLEMENTATION OF THE TOEPLITZ REAL-VALUED MUSIC

As analyzed in Sections III and IV, the COB-based unitary MUSIC algorithm for finding a source's direction, referred to as the COBU-MUSIC algorithm, can be implemented by the following six steps.

1) Compute the $M \times M$ upper triangular Toeplitz estimation matrix $\boldsymbol{Q}'$, and its nonzero (i, j)th entry can be written as

$$Q_{i,j}' = \frac{1}{N}\sum_{t=1}^{N}\left[\exp\left(-\frac{|x_i(t) - x_j^*(t)|}{2\sigma^2}\right)\left(x_i(t) - \xi x_j^*(t)\right)^2\right] \tag{31}$$

2) Compute the $M/2 \times M/2$ matrices $\Re(\boldsymbol{A})$, $\Im(\boldsymbol{A})$, $\Re(\boldsymbol{A}^T)$, $\Im(\boldsymbol{A}^T)$, $\Re(\boldsymbol{CJ})$ and $\Im(\boldsymbol{CJ})$, which constitute the estimation matrix $\boldsymbol{C}_T$;
3) Execute SVD on $\boldsymbol{C}_T$ and obtain the $M \times (M-L)$ singular matrix $\boldsymbol{D}$, which associates with the smallest $M-L$ singular values of the estimation matrix $\boldsymbol{C}_T$.
4) Compute the spatial spectrum of COBU-MUSIC

$$\boldsymbol{P}_{\mathrm{COBU}}(\theta) = \frac{1}{\boldsymbol{a}^H(\theta)\boldsymbol{D}\left(\boldsymbol{D}^H\boldsymbol{D}\right)^{-1}\boldsymbol{D}^H\boldsymbol{a}(\theta)} \tag{32}$$

where $\boldsymbol{a}(\theta)$ is the ULA steering vector;
5) Find L local peaks of $\boldsymbol{P}_{\mathrm{COBU}}(\theta)$;
6) Estimate the DOAs $\{\theta_1, \theta_2, \ldots, \theta_L\}$ of the multiple sources from the L peaks.

C. COMPLEXITY ANALYSIS

In practical applications, not only the DOA estimation performance of the methods but also their computational complexity is important. Therefore, we compare the computational complexity of the MUSIC, FLOM-MUSIC, PFLOM-MUSIC, ZMNL-MUSIC and CRCO-MUSIC algorithms with that of the proposed algorithm in this section. It is well known that the MUSIC-like DOA estimation algorithms involve constructing a covariance matrix, eigenvalue decomposition and spectral peak search. In general, the computational complexity of the MUSIC algorithm is concentrated in the eigenvalue decomposition and spectral peak search stages.

For the spectral peak search, $(\pi/\varpi)+1$ steps are required, in which ϖ is the step length. The calculation for each step is $2M(M-L)+M$ real multiplications. Therefore, the total number of operation in the spectral peak search is $\mathcal{O}([2M(M-L)+M](\pi/\varpi+1))$.

Considering the eigenvalue decomposition of the $M \times M$ complex matrix, $\mathcal{O}(20M^3)$ real multiplications are required [41]; therefore, the complexity of MUSIC, FLOM-MUSIC, PFLOM-MUSIC, ZMNL-MUSIC and CRCO-MUSIC is $\mathcal{O}(20M^3)$ real multiplications, and the complexity of COBU-MUSIC is $\mathcal{O}(17/3M^3 + 2NM^2)$ real multiplications [42].

For constructing the covariance matrix, MUSIC needs $\mathcal{O}(4NM^2)$ real multiplications, and the complexity of FLOM-MUSIC, PFLOM-MUSIC and ZMNL-MUSIC is $\mathcal{O}(6NM^2+11NM)$, $\mathcal{O}(4NM^2+14NM)$ and $\mathcal{O}(4NM^2+15NM)$, respectively. In terms of real multiplications, the total computational complexity required by the CRCO-MUSIC can be approximately expressed as $\mathcal{O}(13NM^2)$, and the complexity of the proposed algorithm is $\mathcal{O}(11NM)$. It can readily be seen that the computational complexity of the proposed algorithm is dominated by the requirements of estimating the DOAs and among all the algorithms, the proposed algorithm has the best execution efficiency.

D. CRAMÉR-RAO BOUND

Consider that the complex-valued noise components $\omega(t) = \bar{\omega}(t) + \mathrm{j}\tilde{\omega}(t)$ are i.i.d. in time and space, and assume that the pdf is symmetric; therefore, we can represent the pdf as

$$\begin{aligned} f_{\mathrm{pd}}(\bar{\omega}, \tilde{\omega}) &= f_{\mathrm{pd}}(\pm\bar{\omega}, \pm\tilde{\omega}) \\ &= f(\sqrt{\bar{\omega}^2+\tilde{\omega}^2}) = f(\rho) \end{aligned} \tag{33}$$

where $\rho = |\omega| = \sqrt{\bar{\omega}^2+\tilde{\omega}^2}$, and $f(\cdot)$ is defined on $[0, +\infty)$. Assume that $f_{\mathrm{pd}}(\bar{\omega}, \tilde{\omega})$ satisfies the regularity conditions (some details of the regularity conditions can be found in APPENDIX B), which are necessary for the existence of Cramér-Rao bound (CRB), the CRB on the accuracy of estimating the DOA parameter vector $\boldsymbol{\theta} = [\theta_1, \theta_2, \ldots, \theta_L]^T$ for non-Gaussian noise has been derived in [12] and [43]. In [12], the variance of any unbiased estimation is given to obtain bounded below by the following

$$\begin{aligned} \mathrm{CRB}^{-1}(\boldsymbol{\theta}) = I_c \sum_{t=1}^{T} \Re\Big[&\boldsymbol{S}^H(t)\boldsymbol{Z}^H(\boldsymbol{\theta})(\boldsymbol{I} - \boldsymbol{A}(\boldsymbol{\theta}) \\ &\times(\boldsymbol{A}^H(\boldsymbol{\theta})\boldsymbol{A}(\boldsymbol{\theta}))^{-1}\boldsymbol{A}^H(\boldsymbol{\theta}))\boldsymbol{Z}(\boldsymbol{\theta})\boldsymbol{S}(t)\Big] \end{aligned} \tag{34}$$

where $\boldsymbol{S}(t) = \mathrm{diag}\{s_1(t), s_2(t), \cdots, s_L(t)\}$ is a diagonal matrix, $\boldsymbol{Z}(\boldsymbol{\theta}) = [\partial\boldsymbol{a}(\theta_1)/\partial\theta_1, \partial\boldsymbol{a}(\theta_2)/\partial\theta_2, \cdots, \partial\boldsymbol{a}(\theta_L)/\partial\theta_L]$ and I_c can be written as

$$I_c = \pi \int_0^{\infty} \frac{\left[f'(\rho)\right]^2}{f(\rho)} \rho d\rho \tag{35}$$

where $f(\cdot)$ is the noise pdf. Equation (34) illustrates that CRB is affected only by the scalar factor $1/I_c$.

Unfortunately, no closed-form expressions exist for the alpha-stable distributions, except for the Cauchy ($\alpha = 1$) and Gaussian ($\alpha = 2$) distributions. However, power series expansions of stable density functions are available and shown in APPENDIX C.

Let us consider the complex Gaussian distributed noise with variance 2γ

$$f(\rho) = \frac{1}{2\pi\gamma}\exp\left(-\frac{\rho^2}{2\gamma}\right) \tag{36}$$

and the complex Cauchy distributed noise

$$f(\rho) = \frac{v}{2\pi\left(v^2+\rho^2\right)^{3/2}} \tag{37}$$

where v is the dispersion. Evaluation of (36) and (37) yields $I_c(\gamma) = 1/\gamma$ and $I_c(v) = 3/5v^2$ for the Gaussian and Cauchy distributions, respectively. As suggested by [44], if $1 \leq \alpha \leq 2$, the CRB scalar factor I_c can be well-described as a quadratic monotonically increasing function with the characteristic exponent α.

V. SIMULATION RESULTS

A. SIMULATION SETUP

In these simulations, we assume that two statistically independent narrow-band quadrature phase-shift keying (QPSK) signals with unit power are received by a ULA with $M = 8$ sensors and $d = \lambda/2$, where d is the distance between adjacent sensors and λ is the wavelength of both signals.

The underlying noises are additive complex isotropic $S\alpha S$ distribution with the characteristic exponent $0 < \alpha \leq 2$ and the location parameter $\mu = 0$. However, due to the infinite variance of the alpha-stable family for $0 < \alpha < 2$, we define an effective alternative signal-to-noise ratio (SNR), namely, the generalized signal-to-noise ratio (GSNR), which is utilized to evaluate the rate of the signal power over noise dispersion by

$$\mathrm{GSNR} = 10 \log \frac{1}{\gamma} \mathbf{E}\left[|s(t)|^2\right] \tag{38}$$

where γ is the dispersion parameter of the SαS distribution.

We evaluate the estimation capability of the proposed algorithm through two statistical measures: the probability of resolution and the root mean square error (RMSE). The two signals are considered to be resolvable if the following resolution criterion can be satisfied

$$\delta(\theta_{\mathrm{I}}, \theta_{\mathrm{II}}) = g(\theta_m) - \frac{1}{2}\left[g(\theta_{\mathrm{I}}) + g(\theta_{\mathrm{II}})\right] > 0 \tag{39}$$

where θ_{I} and θ_{II} are two signal arrival angles, θ_m is the mean of θ_{I} and θ_{II}, and $g(\theta)$ is the reciprocal of the spatial spectrum $P(\theta)$. The probability of resolution is the ratio of the successful runs to the total Monte Carlo runs, and the RMSE of those successful runs is defined by

$$\mathrm{RMSE} = \frac{1}{2}\left[\sqrt{\frac{1}{K}\sum_{i=1}^{K}\left(\hat{x}_1(i) - x_1\right)^2} + \sqrt{\frac{1}{K}\sum_{i=1}^{K}\left(\hat{x}_2(i) - x_2\right)^2}\right] \tag{40}$$

where K is the number of the successful runs and $\hat{x}_1$ and $\hat{x}_2$ are the successful estimates of x_1 and x_2, respectively.

In each simulation, we perform three hundred Monte Carlo runs and compute the probability of resolution and RMSE of the DOA estimation. The performance of several algorithms is reviewed as the functions of seven parameters, namely, the weight factor, kernel size, GSNR, characteristic exponent, number of snapshots, angular separation and SNR.

We compare the FLOM-MUSIC, PFLOM-MUSIC, ZMNL-MUSIC, CRCO-MUSIC and conventional MUSIC algorithms.

The directions of the two impinging sources are $\theta_1 = 5°$ and $\theta_2 = 15°$ except in Section V-G.

B. THE SELECTION OF THE WEIGHT FACTOR

In (18), we introduce a weight factor ξ to ensure that the estimation matrix satisfies the semipositive definite conditions. When the weight factor equals 1, the data samples used to build the estimation matrix are original array outputs without any adjustment. Therefore, DOA estimation can achieve high precision. However, when $\xi \neq 1$, the amplitude of the data samples is changed by the multiple of the weight factor. The direct effects are that the original information will be lost and that the accuracy of the DOA estimation will decrease. Therefore, if the array outputs are real-valued, we choose $\xi \neq 1$, the solution for the direction finding is at the expense of loss accuracy.

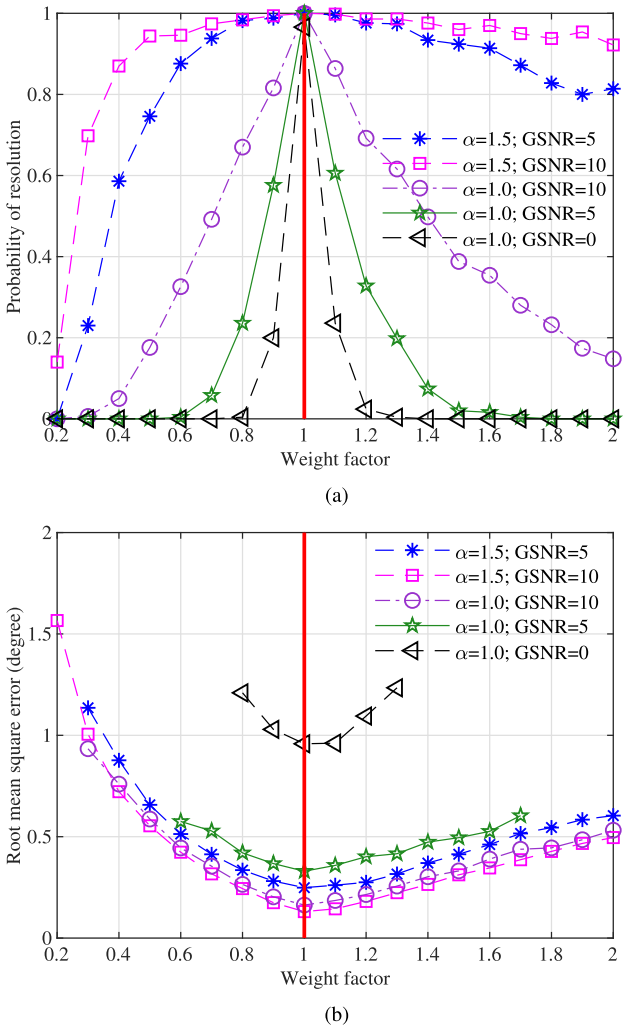

FIGURE 2. Performance of different weight factors: (a) Probability of resolution; (b) Root mean square error.

In Fig. 2, we see that when the weight factor $\xi = 1$, both the RMSE and the probability of resolution are optimal. The performance degrades with the increase or decrease in the weight factor. Furthermore, increasing the impulsive components in the noise or decreasing the GSNR will lead to a reduction in the performance of DOA estimation. Therefore, to obtain the optimal performance of the proposed algorithm, setting the weight factor to 1 is the most reasonable under the complex-valued array output conditions. Fig. 2 also indicates that, in practice, the effective interval of the weight factor can be limited to (0, 1].

C. THE SELECTION OF THE KERNEL SIZE

Fig. 3 illustrates the performance of the proposed algorithm versus kernel size σ varied from 0.2 to 4.0 and the characteristic exponent α varied from 1.0 to 1.6. As recommended in [30], the correntropy is directly related to the probability of how similar two random variables are in a neighborhood of the joint space controlled by the kernel size. Furthermore, the kernel size can also affect the performance

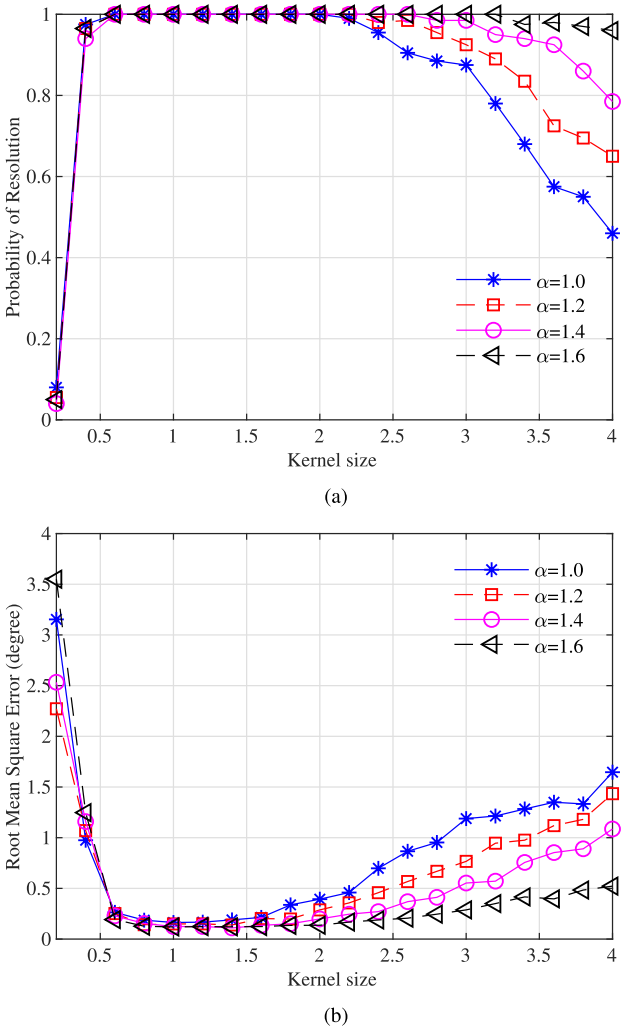

FIGURE 3. Performance of different kernel sizes: (a) Probability of resolution; (b) Root mean square error.

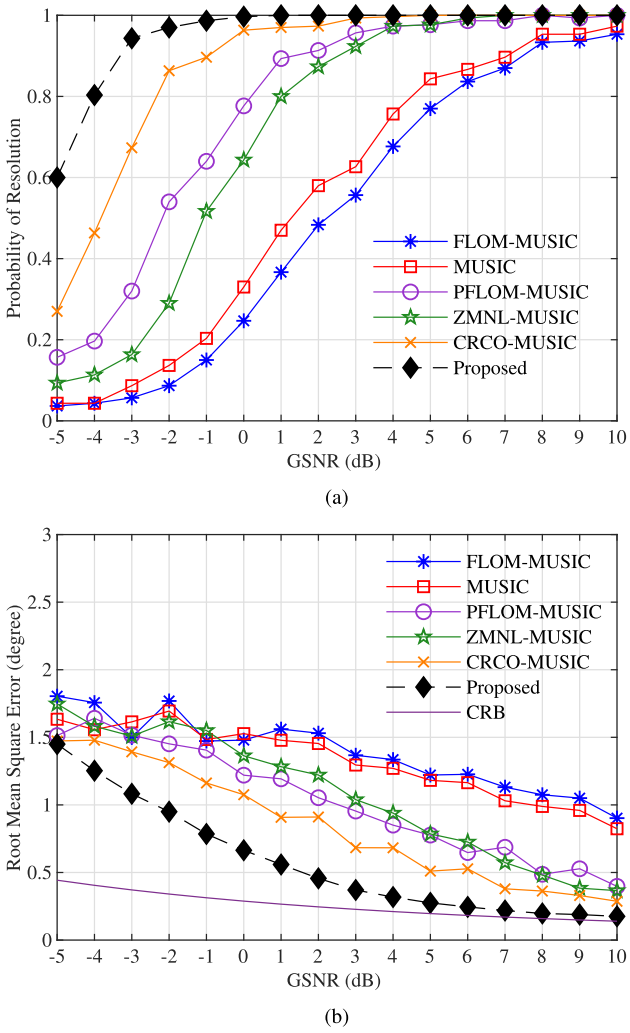

FIGURE 4. Performance of different methods versus GSNR at $\alpha = 1.5$: (a) Probability of resolution; (b) Root mean square error.

of the COBU-MUSIC. Fig. 3(a) shows that as σ increases from 0.2 to 0.4, the probability of resolution monotonically increases from 0.05 to 1 and has almost no difference with different α values. When σ is in $[0.5, 2.2]$, the probability of resolution equals 1; that is to say, the resolution of signal sources is 100%.

However, if $\sigma > 2.2$, the probability of resolution presents a downward trend. The RMSEs are minimum in $\sigma \in [0.7, 1.5]$ as shown in Fig. 3(b). Fig. 3 also illustrates that if the kernel size is fixed, the performance of the COBU-MUSIC will decrease with decreasing the characteristic exponent α. From what has been discussed above, the kernel size is chosen as $\sigma = 1.0$ in our simulations.

D. THE EFFECT OF THE GSNR

Fig. 4 and Fig. 5 display the performance of the COBU-MUSIC compared with the other algorithms versus GSNRs over a wide range from -5 dB to 10 dB. The number of snapshots available for these algorithms is set to 300. In Fig. 4, the underlying noise is set to be modulated impulsive noise with the characteristic exponent $\alpha = 1.5$; the performance of all algorithms improved with the increasing of GSNR is demonstrated. However, the conventional MUSIC, FLOM-MUSIC and PFLOM-MUSIC algorithms are inferior to the other algorithms. For low GSNR, the CRCO-MUSIC, ZMNL-MUSIC and COBU-MUSIC have a higher success probability than the other algorithms, and the RMSE of COBU-MUSIC is lower than the RMSEs of the other algorithms.

A fairly strong impulsive noise environment with $\alpha = 1.0$ (Cauchy noise) is employed, and the DOA estimations are presented in Fig. 5. As expected, due to the increase of the impulsive components in the noise, the performance of all algorithms degrades; however, the performance of the COBU-MUSIC is still the best and very close to the CRB.

E. THE EFFECT OF THE CHARACTERISTIC EXPONENT

Fig. 6 shows the probability of resolution and RMSE curves as functions of the characteristic exponent α, which varies

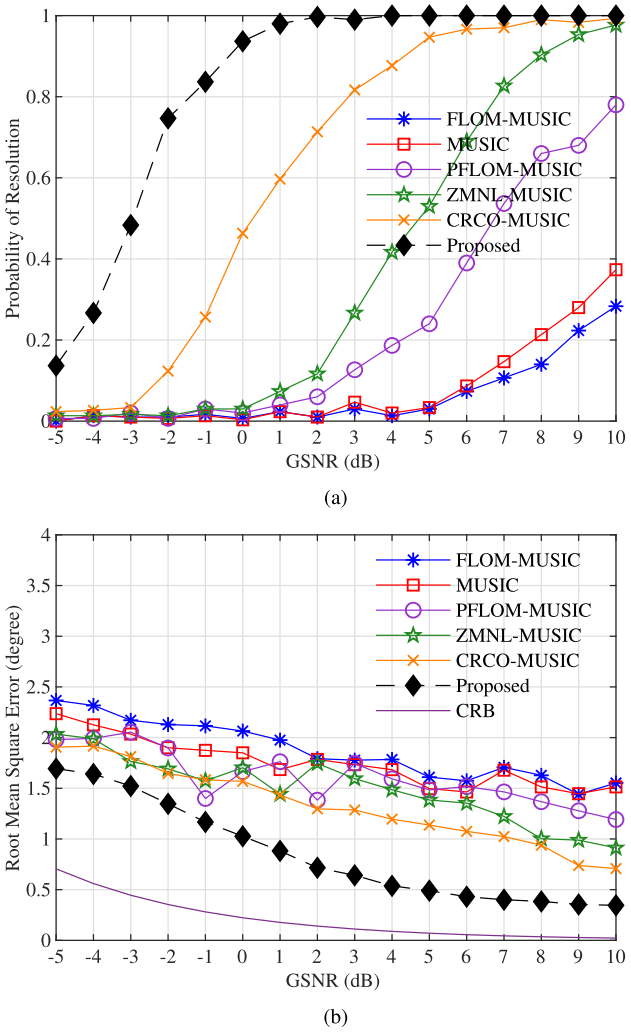

FIGURE 5. Performance of different methods versus GSNR at $\alpha = 1.0$: (a) Probability of resolution; (b) Root mean square error.

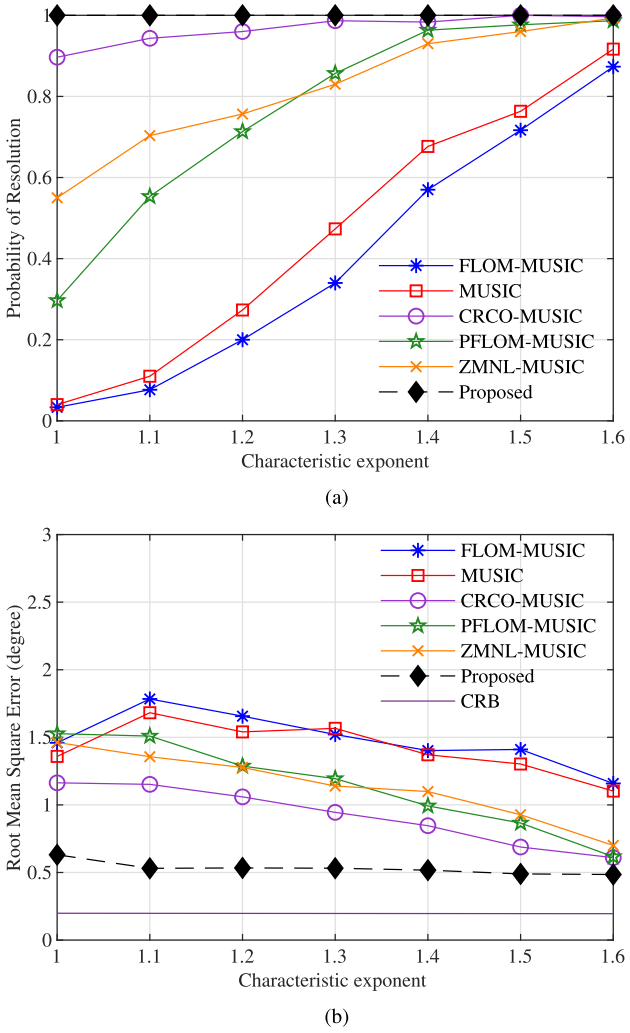

FIGURE 6. Performance of different methods versus the characteristic exponent α: (a) Probability of resolution; (b) Root mean square error.

from highly impulsive conditions with $\alpha = 1.0$ to moderately impulsive conditions with $\alpha = 1.6$. To allow a direct comparison of the estimation performance of various algorithms, the GSNR is assumed as 0 dB in these simulations, and the number of snapshots available to the algorithms is $N = 300$.

In general, all of the algorithms perform very well for $\alpha = 1.6$ and poorly for $\alpha = 1.0$. In Fig. 6(a), it is evident that the change in α has a slight impact on the probability of resolution of the proposed algorithm. In $\alpha \in [1.0, 1.6]$, the probability of resolution is close to or equal to 1. In Fig. 6(b), the COBU-MUSIC based on the $S\alpha S$ noise assumption has the lowest RMSE and is very close to the CRB; in other words, the performance of the COBU-MUSIC is closer to optimal than the other algorithms. The results indicate that it is beneficial to apply the COBU-MUSIC instead of the other algorithms if quite impulsive noise is involved.

F. THE EFFECT OF THE NUMBER OF SNAPSHOTS

In this experiment, we study the influence of the number of snapshots N on the performance of these algorithms. The GSNR is kept almost constant at 5 dB. We choose the characteristic exponent $\alpha = 1.2$ corresponding to a highly impulsive noise condition. From Fig. 7(a), we can observe that although the CRCO-MUSIC, PFLOM-MUSIC and ZMNL-MUSIC are superior to the FLOM-MUSIC and conventional MUSIC algorithms, they are inferior to the COBU-MUSIC regarding the probability of resolution.

Fig. 7(b) shows the resulting RMSE of the estimated DOA as a function of the number of snapshots N. The CRB is also plotted. As expected, when N increases from 100 to 1000, the RMSE decreases. The reason is that when the number of snapshots is small, the difference between the estimated value and the true value is large. It is noted that the proposed algorithm can obtain the best accurate DOA estimation.

G. THE EFFECT OF THE ANGULAR SEPARATION

The importance of this experiment rests in its study of the performance of the proposed algorithm and other contrast algorithms concerning the angular separation of the two incoming signals. In this experiment, a fairly strong impulsive

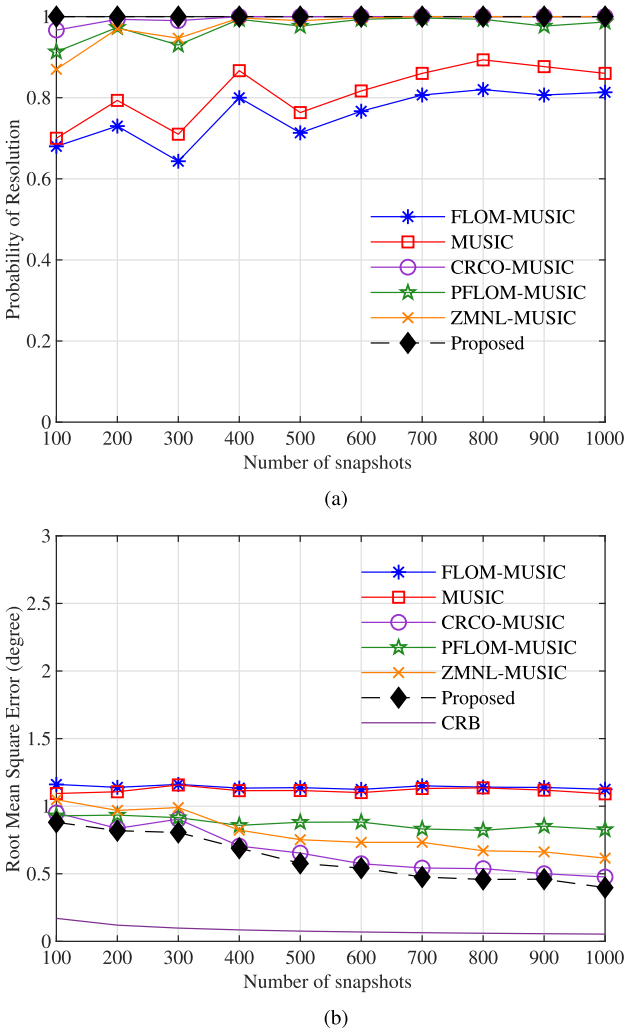

FIGURE 7. Performance of different methods versus the number of snapshots N: (a) Probability of resolution; (b) Root mean square error.

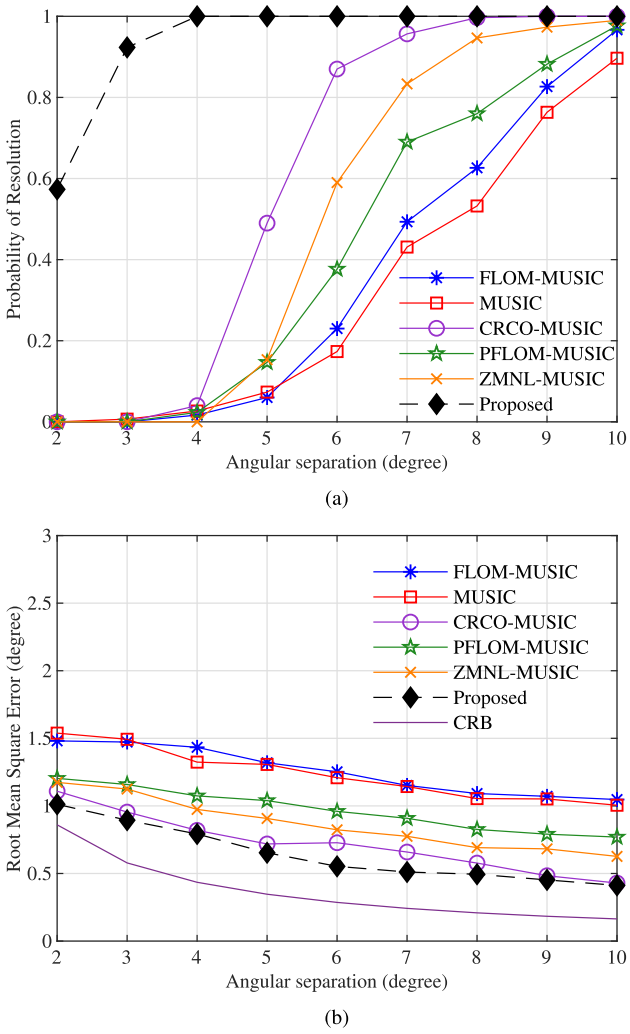

FIGURE 8. Performance of different methods versus the angular separation: (a) Probability of resolution; (b) Root mean square error.

noise with $\alpha = 1.2$ is applied, and the GSNR is kept at 5 dB. The number of snapshots N available to these algorithms is 300.

Fig. 8 contains the simulation results. As we can see, by comparing with the other algorithms, the COBU-MUSIC exhibits robustness to the angular separation; that is, for a given angular separation in Fig. 8(b), the COBU-MUSIC always obtains a much lower angular separation threshold than the other contrast algorithms. For a given probability of resolution in Fig. 8(a), the COBU-MUSIC still obtains a much smaller angular separation threshold than the others. We can also clearly observe that for the angular separation in [2°, 10°], the probability of resolution of the COBU-MUSIC is higher than 0.58; however, when the angular separation equals 2° or 3°, the other algorithms are not able to separate the two incident sources.

H. OTHER IMPULSIVE NOISE MODEL

In this section, we consider other two typical impulsive noise models, i.e., the generalized Gaussian distribution (GGD) and the Gaussian mixture model (GMM). In addition, we utilize the SNR to evaluate the rate of the signal power over noise power. The SNR can be expressed as

$$\mathrm{SNR} = 10\log\frac{\sigma_s^2}{\sigma_n^2} \tag{41}$$

where σ_s^2 and σ_n^2 are the variances of the signal and noise, respectively.

1) GGD

The pdf of a zero-mean GGD is given by

$$p_g(\upsilon) = \frac{\varsigma\Gamma(4/\varsigma)}{2\pi\sigma_\upsilon^2\Gamma^2(2/\varsigma)}\exp\left(-\frac{|\upsilon|^\varsigma}{\eta\sigma_\upsilon^\varsigma}\right) \tag{42}$$

where σ_υ^2 is the variance, $\varsigma > 0$ is the shape parameter, $\Gamma(\cdot)$ is the usual Gamma function defined by $\Gamma(x) = \int_0^\infty t^{x-1}\mathrm{e}^{-t}dt$ and $\eta = (\Gamma(2/\varsigma)/\Gamma(4/\varsigma))^{\varsigma/2}$. The case of $\varsigma < 2$ models a heavy-tailed distribution. Specifically, $\varsigma = 1$ corresponds to

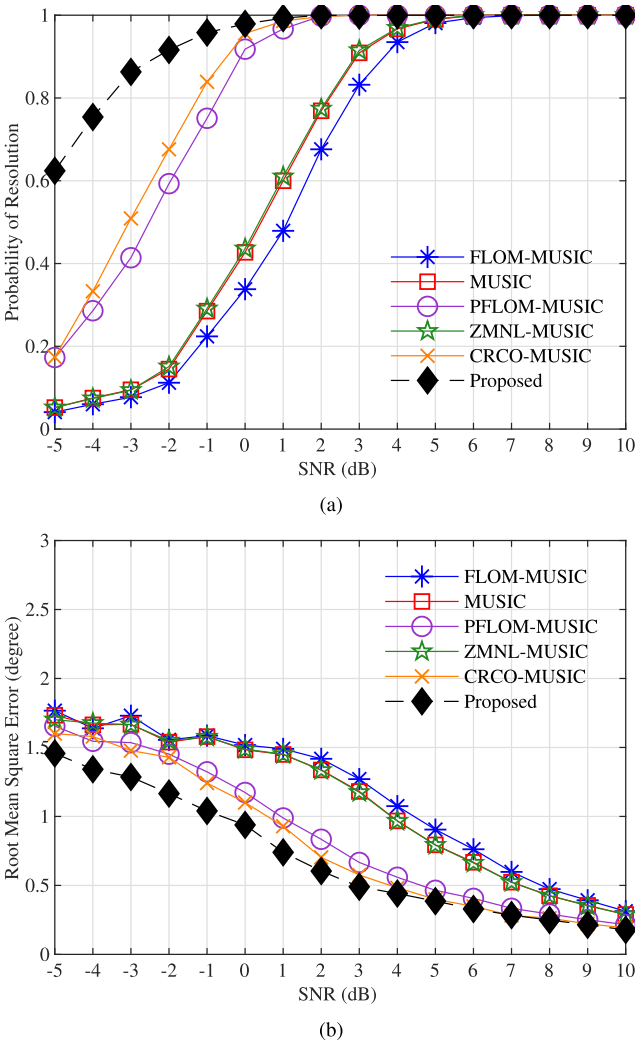

FIGURE 9. Performance as a function of the SNR in GGD: (a) Probability of resolution; (b) Root mean square error.

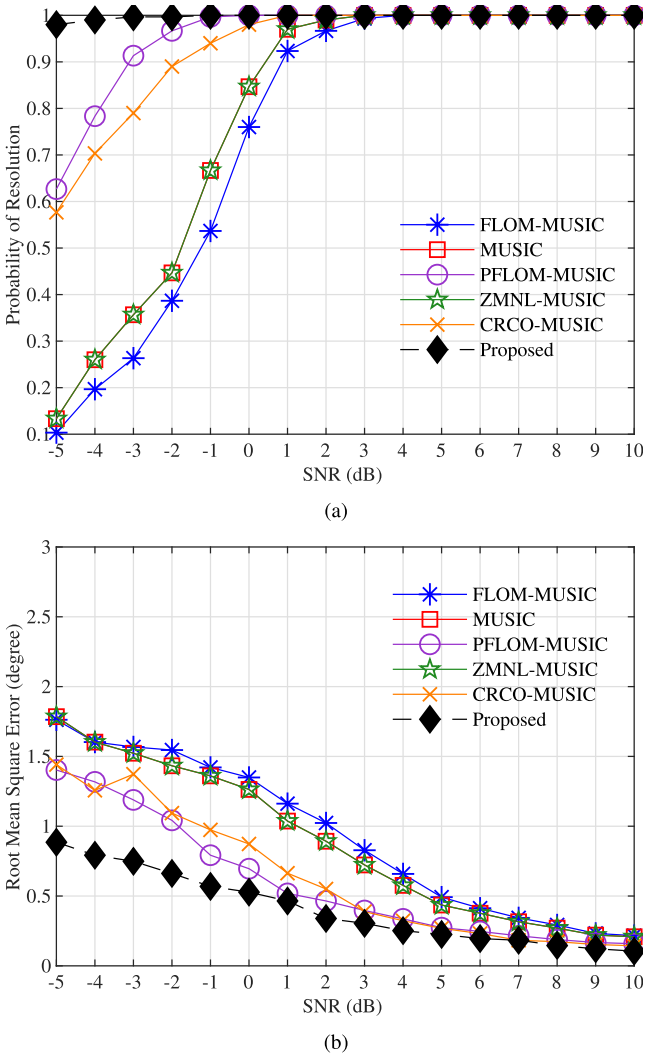

FIGURE 10. Performance as a function of the SNR in GMM: (a) Probability of resolution; (b) Root mean square error.

the Laplacian distribution, and when $\varsigma = 2$, the GGD reduces to the Gaussian distribution [47]–[50].

Note that the smaller the value of ς is, the more impulsive the noise. We take $\varsigma = 0.3$ to simulate the impulsive noise.

2) GMM

The pdf of the two-term GMM [51], [52] noise υ can be expressed as

$$p_m(\upsilon) = \sum_{i=1}^{2} \frac{\varrho_i}{\sqrt{2\pi}\sigma_i} \exp\left(-\frac{\upsilon^2}{2\sigma_i^2}\right) \tag{43}$$

where σ_i^2 is the variance of the ith term and $0 \leq \varrho_i \leq 1$ is the probability of the ith term with $\varrho_1 + \varrho_2 = 1$. The total variance of the GMM is $\sigma^2 = \sum_{i=1}^{2} \varrho_i \sigma_i^2$.

If we choose $\varrho_1 < \varrho_2$ and $\sigma_1^2 \gg \sigma_2^2$, we can view that large noise samples of variance σ_1^2 with a small probability ϱ_1 as sparse impulsive noises embed in Gaussian background noises with small noise samples of variance σ_2^2 and a large probability ϱ_2. Thus, the GMM can well model the phenomenon in both impulsive noise and Gaussian noise environments.

In this simulation, we set $\varrho_1 = 0.1$, $\varrho_2 = 0.9$ and $\sigma_1^2 = 100\sigma_2^2$, which satisfies that 10% of the impulsive noises are contained in the noises and 20 dB stronger than the Gaussian background noises [51], [52].

In Fig. 9 and Fig. 10, we consider GGD and GMM as impulsive noises and compare the performance of COBU-MUSIC with the different algorithms versus SNRs over a wide range from -5 dB to 10 dB. In Fig. 9(a), when SNR $= -5$ dB, the probability of resolution of the proposed algorithm is 0.62, whereas the possibility of resolution of the other algorithms does not exceed 0.2. When SNR > 6 dB, all algorithms can completely separate the incident signals. Fig. 9(b) illustrates that as the SNR increases from -5 dB to 10 dB, the RMSEs of all algorithms show a tendency to decay; however, among the algorithms assessed, the proposed algorithm has the lowest RMSE.

Fig. 10 shows the performance analysis of the comparison algorithms in the presence of GMM impulsive noise. It is worth nothing that when the SNR increases gradually, the proposed and several related algorithms have an approximate trend with GGD impulsive noise. Moreover, the COBU-MUSIC still has the best performance. However, when fixing the SNR, the suppression of GMM impulsive noise of the proposed algorithm is better than that of GGD impulsive noise.

Overall, these results indicate that the proposed algorithm can achieve better performance and higher robustness than others under the three typical impulsive noise ($S\alpha S$, GGD and GMM) conditions.

VI. CONCLUSIONS

We present a novel approach to the DOA estimation problem for array signal processing in the impulsive noise that is i.i.d. in time and space and modeled with complex isotropic $S\alpha S$, GGD and GMM, respectively. An estimation matrix is constructed based on the correntropy and Toeplitz transformation. To improve computational efficiency, we employ the unitary transformation and define a new approximate estimation matrix, which is used to build the COBU-MUSIC algorithm.

The comprehensive simulations are carried out to evaluate the performance of different MUSIC-like algorithms and analyze the effect of parameter changes on the estimation results. The results indicate that the COBU-MUSIC outperforms other algorithms and its RMSE is closer to the CRB than the RMSEs of other algorithms over wide ranges of operating conditions and the impulsive noise environments. Nevertheless, the COBU-MUSIC is also particularly powerful for direction finding in the presence of highly impulsive noise when reducing computational cost is a crucial design requirement. Furthermore, future research includes that we will extend our present work to the fields of time delay estimation and detection of the number of signal sources.

APPENDIX

A. THE PROOF OF THE BOUNDEDNESS OF COB

The translation invariance Gaussian kernel is commonly used to construct the correntropy. However, similar to the Parzen estimation of the pdf, other symmetric kernels can be applied. According to theoretical analysis and numerical experiments of our previous research, we have found that the performance of the exponential kernel is better than that of the Gaussian kernel. Therefore, the proof is based on the exponential kernel.

Proof: Based on the definitions as mentioned above of both X and Y, the COB operator can be rewritten as:

$$\boldsymbol{V}_T = \mathbf{E}\left[(X-Y)^2 \exp\left(-\frac{|X-Y|}{2\sigma^2}\right)\right] \tag{44}$$

According to the discussion in Section III-A, we construct a new random variable $z = X - Y$, which obeys the $S\alpha S$ distribution, then $\boldsymbol{V}_T$ can be written in the form of integral

$$\begin{aligned}\boldsymbol{V}_T &= \mathbf{E}\left[z^2 \exp\left(-\frac{|z|}{2\sigma^2}\right)\right] \\ &= \int_{-\infty}^{+\infty} \mathrm{e}^{-\frac{|z|}{2\sigma^2}} z^2 f(z)\, dz\end{aligned} \tag{45}$$

where $f(z)$ is the pdf of the random variable z.

The characteristic function of $S\alpha S$ can also be expressed as

$$\varphi(\omega) = \exp\left(\mathrm{j}\mu\omega - \gamma|\omega|^\alpha\right) \tag{46}$$

Without loss of generality, we assume that all $S\alpha S$ distributions are centered at the origin so that $\mu = 0$. In this case, an $S\alpha S$ distribution is only determined by the parameters $0 < \alpha \leq 2$ and $\gamma > 0$, through its characteristic function

$$\varphi(\omega) = \exp\left(-\gamma|\omega|^\alpha\right) \tag{47}$$

Substituting (47) into (45), we acquire

$$\begin{aligned}\boldsymbol{V}_T &= \frac{1}{2\pi}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} \mathrm{e}^{-\frac{|z|}{2\sigma^2}} z^2 \mathrm{e}^{-\gamma|\omega|^\alpha - \mathrm{j}\omega z} d\omega dz \\ &\leq \frac{1}{2\pi}\int_{-\infty}^{+\infty}\left(\int_{-\infty}^{+\infty} \mathrm{e}^{-\frac{|z|}{2\sigma^2}} z^2 dz\right)\mathrm{e}^{-\gamma|\omega|^\alpha} d\omega\end{aligned} \tag{48}$$

Denoting $g_1 = \int_0^{+\infty} \mathrm{e}^{-z/2\sigma^2} z^2 dz = \frac{1}{2}\int_{-\infty}^{+\infty} \mathrm{e}^{-|z|/2\sigma^2} z^2 dz$ and $u = \frac{z}{2\sigma^2}$ for short, we find that

$$\begin{aligned}g_1 &= 8\sigma^6 \int_0^{\infty} \mathrm{e}^{-u} u^2 du \\ &= 8\sigma^6\left[-\mathrm{e}^{-u}\left(u^2+2u+2\right)\right]\Big|_0^{\infty} \\ &= 8\sigma^6\left(2 - \lim_{u\to\infty}\mathrm{e}^{-u}u^2 - 2\lim_{u\to\infty}\mathrm{e}^{-u}u\right) \\ &= 16\sigma^6\end{aligned} \tag{49}$$

then, we only need to consider the integral as follows

$$\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} \mathrm{e}^{-\frac{|z|}{2\sigma^2}} z^2 dz \mathrm{e}^{-\gamma|\omega|^\alpha} d\omega = 32\sigma^6 \int_{-\infty}^{+\infty} \mathrm{e}^{-\gamma|\omega|^\alpha} d\omega \tag{50}$$

By defining $g_2 = \int_0^{+\infty} \mathrm{e}^{-\gamma\omega^\alpha} d\omega = \frac{1}{2}\int_{-\infty}^{+\infty} \mathrm{e}^{-\gamma|\omega|^\alpha} d\omega$, through some derivation, we can obtain a simplified form as (51) that appears at the top of the next page, where WM is the Whittaker M function that can be expressed as

$$\mathrm{WM}(\mu,\nu,z) = \mathrm{e}^{-\frac{z}{2}} z^{\frac{1}{2}+\nu} \sum_{n=0}^{\infty} \frac{\left(\frac{1}{2}+\nu-\mu\right)^{(n)} z^n}{(1+2\nu)^{(n)} n!} \tag{52}$$

and keep in mind that

$$\begin{aligned}a^{(0)} &= 1 \\ a^{(n)} &= a(a+1)\cdots(a+n-2)(a+n-1)\end{aligned} \tag{53}$$

To simplify the derivation of the expression given in (51), we define two partial derivative functions $d_1 = \frac{\partial}{\partial\alpha}\left(\frac{\gamma^{-\frac{1}{\alpha}}\Gamma\left(\frac{1}{\alpha}\right)}{\alpha}\right)$ and $d_2 = \frac{\partial}{\partial\gamma}\left(\frac{\gamma^{-\frac{1}{\alpha}}\Gamma\left(\frac{1}{\alpha}\right)}{\alpha}\right)$.

$$\int_0^{\infty} \mathrm{e}^{-\gamma\omega^{\alpha}} d\omega = \left[\frac{\alpha\omega^{-\alpha+1}\gamma^{-1}\left(\alpha\gamma\omega^{\alpha}+\alpha+1\right)\left(\gamma\omega^{\alpha}\right)^{-\frac{\alpha+1}{2\alpha}}\mathrm{e}^{-\frac{\gamma\omega^{\alpha}}{2}}\mathrm{WM}\left(\frac{1}{\alpha}-\frac{\alpha+1}{2\alpha},\frac{\alpha+1}{2\alpha}+\frac{1}{2},\gamma\omega^{\alpha}\right)}{(\alpha+1)(2\alpha+1)} + \frac{\omega^{-\alpha+1}\gamma^{-1}(\alpha+1)\left(\gamma\omega^{\alpha}\right)^{-\frac{\alpha+1}{2\alpha}}\mathrm{e}^{-\frac{\gamma\omega^{\alpha}}{2}}\mathrm{WM}\left(\frac{1}{\alpha}-\frac{\alpha+1}{2\alpha}+1,\frac{\alpha+1}{2\alpha}+\frac{1}{2},\gamma\omega^{\alpha}\right)}{2\alpha+1} \right]\Bigg|_0^{\infty} = \frac{\gamma^{-\frac{1}{\alpha}}\Gamma\left(\frac{1}{\alpha}\right)}{\alpha} \tag{51}$$

Therefore

$$\begin{aligned} d_1 &= \frac{\partial}{\partial\alpha}\left(\frac{\gamma^{-\frac{1}{\alpha}}\Gamma\left(\frac{1}{\alpha}\right)}{\alpha}\right) \\ &= \frac{\frac{\partial}{\partial\alpha}\left(\gamma^{-\frac{1}{\alpha}}\Gamma\left(\frac{1}{\alpha}\right)\right)\alpha-\gamma^{-\frac{1}{\alpha}}\Gamma\left(\frac{1}{\alpha}\right)}{\alpha^2} \\ &= \frac{\left[\frac{\partial}{\partial\alpha}\left(\gamma^{-\frac{1}{\alpha}}\right)\Gamma\left(\frac{1}{\alpha}\right)-\frac{\gamma^{-\frac{1}{\alpha}}\Psi\left(\frac{1}{\alpha}\right)\Gamma\left(\frac{1}{\alpha}\right)}{\alpha^2}\right]\alpha-\gamma^{-\frac{1}{\alpha}}\Gamma\left(\frac{1}{\alpha}\right)}{\alpha^2} \\ &= \frac{\left[\frac{\mathrm{e}^{-\frac{\ln(\gamma)}{\alpha}}\ln(\gamma)\Gamma\left(\frac{1}{\alpha}\right)}{\alpha^2}-\frac{\gamma^{-\frac{1}{\alpha}}\Psi\left(\frac{1}{\alpha}\right)\Gamma\left(\frac{1}{\alpha}\right)}{\alpha^2}\right]\alpha-\gamma^{-\frac{1}{\alpha}}\Gamma\left(\frac{1}{\alpha}\right)}{\alpha^2} \\ &= -\frac{\gamma^{-\frac{1}{\alpha}}\left(\alpha+\Psi\left(\frac{1}{\alpha}\right)-\ln(\gamma)\right)\Gamma\left(\frac{1}{\alpha}\right)}{\alpha^3} \end{aligned} \tag{54}$$

and

$$d_2 = \frac{\partial}{\partial\gamma}\left(\frac{\gamma^{-\frac{1}{\alpha}}\Gamma\left(\frac{1}{\alpha}\right)}{\alpha}\right) = -\frac{\gamma^{-\frac{1}{\alpha}}\Gamma\left(\frac{1}{\alpha}\right)}{\gamma\alpha^2} \tag{55}$$

where $\Psi(\cdot)$ is the digamma function defined by $\Psi(x) = \left(\frac{\partial}{\partial x}\Gamma(x)\right)/\Gamma(x)$. Under the definition of the $S\alpha S$ distribution, we know $\gamma > 0$ and $0 < \alpha \leq 2$. Then, $d_1 < 0$ and $d_2 < 0$. We can also obtain the following expression

$$\lim_{\substack{\alpha\to 0\\ \gamma\to 0}} \int_0^{\infty} \mathrm{e}^{-\gamma\omega^{\alpha}} d\omega = +\infty \tag{56}$$

from this, it is straightforward to see that

$$g_2 = \int_0^{\infty} \mathrm{e}^{-\gamma\omega^{\alpha}} d\omega < +\infty \tag{57}$$

which means that the COB operator $\boldsymbol{V}_T$ is bounded. This is the end of the proof. ■

B. THE REGULAR CONDITIONS

The following are the regularity conditions that are generally needed in statistics and developed to support the Fisher maximum-likelihood algorithms.

Consider a parametric model in which the joint distribution of $X = \{x_1, x_2, \cdots, x_n\}$ has a density function $f_X(X;\theta)$, where $\theta \in \Theta$. If the observations on X are i.i.d.with density function $f(x_i;\theta)$ for each observation.

- The parameter space Θ is compact.
- The unknown parameter value θ_c is identified by

$$\theta_c = \arg\max_{\theta\in\Theta} \mathbf{E}[\log f(X_i;\theta)] \tag{58}$$

- The likelihood function can be expressed as

$$\mathcal{L}(\theta) = \sum_{i=1}^{n} \log f(x_i;\theta) \tag{59}$$

  wihch is continuous in θ. In (58), the expression $\mathbf{E}[\log f(X;\theta)]$ exists.
- $\mathcal{L}(\theta)$ is twice continuously differentiable in a neighborhood of θ_c. The information matrix $\mathcal{I}(\theta_c) = \mathbf{E}[-\partial^2 \log f(X;\theta_c)/\partial\theta\partial\theta']$ exists and is nonsingular.

C. THE PDF OF $s\alpha s$ DISTRIBUTION

Consider the complex-valued $S\alpha S$ distributed noise components $\omega(t) = \bar{\omega}(t) + \mathrm{j}\tilde{\omega}(t)$ are i.i.d. in time and space with pdf defined in (33); by using the polar coordinate $\rho = |\omega| = \sqrt{\bar{\omega}^2+\tilde{\omega}^2}$, we can further express this pdf as $f_{\mathrm{pd}}(\bar{\omega},\tilde{\omega}) = \chi(\rho)$. When $\alpha \neq 1$ or $\alpha \neq 2$, no closed-form expressions exist for the pdf; fortunately, we can express $\chi(\rho)$ as a power series expression form in (60), which appears at the top of the next page, where

$$\vartheta(k) = \Gamma\left(\frac{\alpha k}{2}+1\right) \tag{61}$$

and

$$\psi(k) = \Gamma\left(\frac{2k+2}{\alpha}\right) \tag{62}$$

Furthermore, the $S\alpha S$ distribution has two important properties that play a crucial role in the modeling of uncertainty.

- **Stability property**: Assuming that the random variables X_1, X_2 are i.i.d. with the same distribution as X, for arbitrary constants c_1, c_2, there are constants a, b such that

$$c_1X_1 + c_2X_2 := aX + b \tag{63}$$

  means that $c_1X_1 + c_2X_2$ and $aX + b$ obey the same distribution as X_1 or X_2, then the random variable X has a stable distribution. A general statement can be deduced by using the characteristic function in (1): if

$$\chi(\rho)=\begin{cases}\dfrac{1}{\pi^2\gamma^{2/\alpha}}\displaystyle\sum_{k=1}^{\infty}\frac{(-1)^{k-1}}{k!}\vartheta^2(k)\sin\left(\frac{\pi\alpha k}{2}\right)2^{\alpha k}\left(\frac{\rho}{\gamma^{1/\alpha}}\right)^{-\alpha k-2} & 0<\alpha<1\\ \dfrac{\gamma}{2\pi\left(\rho^2+\gamma^2\right)^{3/2}} & \alpha=1\\ \dfrac{1}{\alpha\pi\gamma^{2/\alpha}}\displaystyle\sum_{k=0}^{\infty}\frac{(-1)^k}{2^{2k+1}(k!)^2}\psi(k)\left(\frac{\rho}{\gamma^{1/\alpha}}\right)^{2k} & 1<\alpha<2\\ \dfrac{1}{4\pi\gamma}\mathrm{e}^{-\rho^2/4\gamma} & \alpha=2\end{cases}\tag{60}$$

$X_1, X_2, \cdots, X_n$ are i.i.d. with the same characteristic exponent α and symmetry parameter β, all linear combinations of the form $\sum_{i=1}^{n} c_i X_i$ are stable, which have the same parameters α and β.

- **Generalized central limit theorem**: Assume the random variables $X_1, X_2, \cdots, X_n$ are i.i.d. under the condition $n \to \infty$, X is the limit in the distribution of normalized sums of the form

$$S_\alpha = \frac{1}{a_n}(X_1 + X_2 + \cdots + X_n) - b_n \tag{64}$$

if and only if the distribution of X satisfies stable distribution. In particular, the limit distribution is Gaussian when $\{X_i\}_{i=1}^{n}$ are i.i.d. and have finite variance.

Even more remarkably, for a complex $S\alpha S$ random variable Y, if $0 < \alpha < 2$, there exist

$$\mathbf{E}|Y|^p = \infty, \quad p \geq \alpha \tag{65}$$

and

$$\mathbf{E}|Y|^p < \infty, \quad 0 \leq p < \alpha \tag{66}$$

if $\alpha = 2$, then

$$\mathbf{E}|Y|^p < \infty, \quad p \geq 0 \tag{67}$$

REFERENCES

[1] H. Krim and M. Viberg, "Two decades of array signal processing research: The parametric approach," *IEEE Signal Process. Mag.*, vol. 13, no. 4, pp. 67–94, Jul. 1996.
[2] Q. Shen, W. Liu, W. Cui, and S. Wu, "Underdetermined DOA estimation under the compressive sensing framework: A review," *IEEE Access*, vol. 4, pp. 8865–8878, Nov. 2016.
[3] H. Zhai, X. Zhang, W. Zheng, and P. Gong, "DOA estimation of noncircular signals for unfolded coprime linear array: Identifiability, DOF and algorithm (May 2018)," *IEEE Access*, vol. 6, pp. 29382–29390, 2018.
[4] J. F. Li and X. F. Zhang, "Direction of arrival estimation of quasi-stationary signals using unfolded coprime array," *IEEE Access*, vol. 5, pp. 6538–6545, 2017.
[5] F. Shu *et al.*, "Low-complexity and high-resolution DOA estimation for hybrid analog and digital massive MIMO receive array," *IEEE Trans. Commun.*, vol. 66, no. 6, pp. 2487–2501, Jun. 2018.
[6] R. Roy and T. Kailath, "Esprit-estimation of signal parameters via rotational invariance techniques," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 37, no. 7, pp. 984–995, Jul. 1989.
[7] T. Wang, B. Ai, R. He, and Z. Zhong, "Two-dimension direction-of-arrival estimation for massive MIMO systems," *IEEE Access*, vol. 3, pp. 2122–2128, 2015.
[8] M. Boizard, G. Ginolhac, F. Pascal, S. Miron, and P. Forster, "Numerical performance of a tensor MUSIC algorithm based on HOSVD for a mixture of polarized sources," in *Proc. 21st Eur. Signal Process. Conf. (EUSIPCO)*, Marrakesh, Morocco, Sep. 2013, pp. 1–5.
[9] R. O. Schmidt, "Multiple emitter location and signal parameter estimation," *IEEE Trans. Antennas Propag.*, vol. 34, no. 3, pp. 276–280, Mar. 1986.
[10] J. Li, D. Li, D. Jiang, and X. Zhang, "Extended-aperture unitary root MUSIC-based DOA estimation for coprime array," *IEEE Commun. Lett.*, vol. 22, no. 4, pp. 752–755, Apr. 2018.
[11] M. Haardt, F. Roemer, and G. D. Galdo, "Higher-order SVD-based subspace estimation to improve the parameter estimation accuracy in multi-dimensional harmonic retrieval problems," *IEEE Trans. Signal Process.*, vol. 56, no. 7, pp. 3198–3213, Jul. 2008.
[12] R. J. Kozick and B. M. Sadler, "Maximum-likelihood array processing in non-Gaussian noise with Gaussian mixtures," *IEEE Trans. Signal Process.*, vol. 48, no. 12, pp. 3520–3535, Dec. 2000.
[13] B. Liao, S.-C. Chan, L. Huang, and C. Guo, "Iterative methods for subspace and DOA estimation in nonuniform noise," *IEEE Trans. Signal Process.*, vol. 64, no. 12, pp. 3008–3020, Jun. 2016.
[14] Z. Shan and T. S. P. Yum, "A conjugate augmented approach to direction-of-arrival estimation," *IEEE Trans. Signal Process.*, vol. 53, no. 11, pp. 4104–4109, Nov. 2005.
[15] H. Belkacemi and S. Marcos, "Robust subspace-based algorithms for joint angle/Doppler estimation in non-Gaussian clutter," *Signal Process.*, vol. 87, no. 7, pp. 1547–1558, Jul. 2007.
[16] P. Tsakalides and C. L. Nikias, "Robust space-time adaptive processing (STAP) in non-Gaussian clutter environments," *IEE Proc.-Radar, Sonar Navigat.*, vol. 146, no. 2, pp. 84–93, Apr. 1999.
[17] C. L. Nikias and M. Shao, *Signal Processing With Alpha-Stable Distributions and Applications*. New York, NY, USA: Wiley, 1995.
[18] W.-J. Zeng, H. C. So, and L. Huang, "ℓ_p-MUSIC: Robust direction-of-arrival estimator for impulsive noise environments," *IEEE Trans. Signal Process.*, vol. 61, no. 17, pp. 4296–4308, Sep. 2013.
[19] E. J. G. Pitman and J. Pitman, "A direct approach to the stable distributions," *Adv. Appl. Probab.*, vol. 48, pp. 261–282, Jul. 2016.
[20] T.-H. Liu and J. M. Mendel, "A subspace-based direction finding algorithm using fractional lower order statistics," *IEEE Trans. Signal Process.*, vol. 49, no. 8, pp. 1605–1613, Aug. 2001.
[21] G.-H. You, T.-S. Qiu, and A.-M. Song, "Novel direction findings for cyclostationary signals in impulsive noise environments," *Circuits Syst. Signal Process.*, vol. 32, no. 6, pp. 2939–2956, Dec. 2013.
[22] P. Tsakalides and C. L. Nikias, "The robust covariation-based MUSIC (ROC-MUSIC) algorithm for bearing estimation in impulsive noise environments," *IEEE Trans. Signal Process.*, vol. 44, no. 7, pp. 1623–1633, Jul. 1996.
[23] W. J. Zeng and H. C. So, "Outlier-robust matrix completion via ℓ_p-minimization," *IEEE Trans. Signal Process.*, vol. 66, no. 5, pp. 1125–1140, Mar. 2018.
[24] W.-J. Zeng, H.-C. So, and A. M. Zoubir, "An ℓ_p-norm minimization approach to time delay estimation in impulsive noise," *Digit. Signal Process.*, vol. 23, no. 4, pp. 1247–1254, 2013.
[25] F. Wen, P. Liu, Y. Liu, R. C. Qiu, and W. Yu, "Robust sparse recovery in impulsive noise via ℓ_1-ℓ_1 optimization," *IEEE Trans. Signal Process.*, vol. 65, no. 1, pp. 105–118, Jan. 2017.
[26] X. Jiang, T. Kirubarajan, and W.-J. Zeng, "Robust sparse channel estimation and equalization in impulsive noise using linear programming," *Signal Process.*, vol. 93, no. 5, pp. 1095–1105, May 2013.
[27] J. He, Z. Liu, and K. T. Wong, "Snapshot-instantaneous $\|\cdot\|_\infty$ normalization against heavy-tail noise," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 44, no. 3, pp. 1221–1227, Jul. 2008.

[28] K.-C. Huarng and C.-C. Yeh, "A unitary transformation method for angle-of-arrival estimation," *IEEE Trans. Signal Process.*, vol. 39, no. 4, pp. 975–977, Apr. 1991.
[29] Q. Huang, G. Zhang, L. Xiang, and Y. Fang, "Unitary transformations for spherical harmonics MUSIC," *Signal Process.*, vol. 131, pp. 441–446, Feb. 2017.
[30] W. Liu, P. P. Pokharel, and J. C. Principe, "Correntropy: Properties and applications in non-Gaussian signal processing," *IEEE Trans. Signal Process.*, vol. 55, no. 11, pp. 5286–5298, Nov. 2007.
[31] P. Chargé, Y. Wang, and J. Saillard, "A non-circular sources direction finding method using polynomial rooting," *Signal Process.*, vol. 81, no. 8, pp. 1765–1770, 2001.
[32] I. Santamaria, P. P. Pokharel, and J. C. Principe, "Generalized correlation function: Definition, properties, and application to blind equalization," *IEEE Trans. Signal Process.*, vol. 54, no. 6, pp. 2187–2197, Jun. 2006.
[33] B. Chen and J. C. Principe, "Maximum correntropy estimation is a smoothed MAP estimation," *IEEE Signal Process. Lett.*, vol. 19, no. 8, pp. 491–494, Aug. 2012.
[34] B. Chen, L. Xing, J. Liang, N. Zheng, and J. C. Principe, "Steady-state mean-square error analysis for adaptive filtering under the maximum correntropy criterion," *IEEE Signal Process. Lett.*, vol. 21, no. 7, pp. 880–884, Jul. 2014.
[35] L. Wang and C. Pan, "Robust level set image segmentation via a local correntropy-based K-means clustering," *Pattern Recognit.*, vol. 47, no. 5, pp. 1917–1925, May 2014.
[36] U. Melia *et al.*, "Correntropy measures to detect daytime sleepiness from EEG signals," *Physiol. Meas.*, vol. 35, no. 10, pp. 2067–2083, Oct. 2014.
[37] J. Zhang, T. Qiu, A. Song, and H. Tang, "A novel correntropy based DOA estimation algorithm in impulsive noise environments," *Signal Process.*, vol. 104, pp. 346–357, Nov. 2014.
[38] P. Wang, T.-S. Qiu, F.-Q. Ren, and A.-M. Song, "A robust DOA estimator based on the correntropy in alpha-stable noise environments," *Digit. Signal Process.*, vol. 60, pp. 242–251, Jan. 2017.
[39] J.-F. Zhang, T.-S. Qiu, P. Wang, and S.-Y. Luan, "A novel cauchy score function based DOA estimation method under alpha-stable noise environments," *Signal Process.*, vol. 138, pp. 98–105, Sep. 2017.
[40] B. Wang, Y. D. Zhang, and W. Wang, "Robust DOA estimation in the presence of miscalibrated sensors," *IEEE Signal Process. Lett.*, vol. 24, no. 7, pp. 1073–1077, Jul. 2017.
[41] Y. Zhou, Z. Fei, S. Yang, J. Kuang, S. Chen, and L. Hanzo, "Joint angle estimation and signal reconstruction for coherently distributed sources in massive MIMO systems based on 2-D unitary ESPRIT," *IEEE Access*, vol. 5, pp. 9632–9646, 2017.
[42] G. H. Golub and C. F. Van Loan, *Matrix Computations*. Baltimore, MD, USA: The Johns Hopkins Univ. Press, 1996.
[43] P. Tsakalides and C. L. Nikias, "Maximum likelihood localization of sources in noise modeled as a stable process," *IEEE Trans. Signal Process.*, vol. 43, no. 11, pp. 2700–2713, Nov. 1995.
[44] A. Swami and B. M. Sadler, "On some detection and estimation problems in heavy-tailed noise," *Signal Process.*, vol. 82, no. 12, pp. 1829–1846, Dec. 2002.
[45] X. Jiang, W.-J. Zeng, A. Yasotharan, H. C. So, and T. Kirubarajan, "Minimum dispersion beamforming for non-Gaussian signals," *IEEE Trans. Signal Process.*, vol. 62, no. 7, pp. 1879–1893, Apr. 2014.
[46] H. Li, X.-L. Li, M. Anderson, and T. Adali, "A class of adaptive algorithms based on entropy estimation achieving CRLB for linear non-Gaussian filtering," *IEEE Trans. Signal Process.*, vol. 60, no. 4, pp. 2049–2055, Apr. 2012.
[47] M. N. Do and M. Vetterli, "Wavelet-based texture retrieval using generalized Gaussian density and Kullback-Leibler distance," *IEEE Trans. Image Process.*, vol. 11, no. 2, pp. 146–158, Feb. 2002.
[48] M. Novey, T. Adali, and A. Roy, "A complex generalized Gaussian distribution—Characterization, generation, and estimation," *IEEE Trans. Signal Process.*, vol. 58, no. 3, pp. 1427–1433, Mar. 2010.
[49] Y. Ma, J. Zhu, and D. Baron, "Approximate message passing algorithm with universal denoising and Gaussian mixture learning," *IEEE Trans. Signal Process.*, vol. 64, no. 21, pp. 5611–5622, Nov. 2016.
[50] R. J. Kozick and B. M. Sadler, "Robust subspace estimation in non-Gaussian noise," in *Proc. IEEE Conf. Acoust., Speech, Signal Process. (ICASSP)*, Istanbul, Turkey, 2000, pp. 3818–3821.
[51] X. Jiang, W.-J. Zeng, H. C. So, T. Kirubarajan, and A. M. Zoubir, "Beamforming via nonconvex linear regression," *IEEE Trans. Signal Process.*, vol. 64, no. 7, pp. 1714–1728, Apr. 2016.
[52] F. Yin, C. Fritsche, D. Jin, F. Gustafsson, and A. M. Zoubir, "Cooperative localization in WSNs using Gaussian mixture modeling: Distributed ECM algorithms," *IEEE Trans. Signal Process.*, vol. 63, no. 6, pp. 1448–1463, Mar. 2015.

QUAN TIAN received the B.S. and M.S. degrees in measurement and control technology and instrument from the Dalian University of Technology, Dalian, China, in 2005 and 2008, respectively, where he is currently pursuing the Ph.D. degree with the Faculty of Electronic Information and Electrical Engineering. His research interests include non-Gaussian signal processing, image processing, and auto piloting.

TIANSHUANG QIU received the B.S. degree from Tianjin University, Tianjin, China, in 1983, and the M.S. degree from the Dalian University of Technology, Dalian, Liaoning, China, in 1993, and the Ph.D. degree from Southeastern University, Nanjing, China, in 1996, all in electrical engineering. He was a Post-Doctoral Fellow with the Department of Electrical Engineering, Northern Illinois University, USA, from 1996 to 2000. He is currently a Professor with the Faculty of Electronic Information and Electrical Engineering, Dalian University of Technology. His research interests include wireless signal processing, biomedical signal processing, and non-Gaussian signal processing.

JITONG MA received the B.S. degree in electronic information engineering from Hohai University, Nanjing, China, in 2013, and the M.S. degree from the Faculty of Electronic Information and Electrical Engineering, Dalian University of Technology, Dalian, China, in 2015, where he is currently pursuing the Ph.D. degree with the Faculty of Electronic Information and Electrical Engineering.

JINGCHUN LI received the B.S. degree from the Department of Electrical Engineering, Xi'an Jiaotong University, Xi'an, China, in 1988, and the M.S. degree from the Department of Management, Hebei University of Technology, Tianjin, China, in 1991.

RONG LI received the B.S. degree from the Department of Communication and Information Engineering, Xi'an University of Posts and Telecommunications, Xi'an, China, in 2007, and the Ph.D. degree from the Department of Information and Communication Engineering, Beijing University of Posts and Telecommunications, Beijing, China, in 2012.