Signal Processing 104 (2014) 346–357

Contents lists available at ScienceDirect

Signal Processing

journal homepage: www.elsevier.com/locate/sigpro

A novel correntropy based DOA estimation algorithm in impulsive noise environments

Jinfeng Zhang [a,b], Tianshuang Qiu [a,*], Aimin Song [a], Hong Tang [a]

[a] Faculty of Electronic Information and Electrical Engineering, Dalian University of Technology, Dalian 116024, China
[b] Shenzhen Key Lab of Advanced Communications and Information Processing, Shenzhen University, Shenzhen 518060, China

ARTICLE INFO

Article history:
Received 6 November 2013
Received in revised form
11 March 2014
Accepted 28 April 2014
Available online 6 May 2014

Keywords:
Correntropy
Direction of arrival
Fractional lower order statistics
Impulsive noise
Symmetric alpha-stable process

ABSTRACT

The Direction of Arrival (DOA) estimation under impulsive noise environments remains open in the field of array signal processing. Inspired by the advantage of correntropy which exhibits a 'robust statistics' property, this paper proposes a new operator, namely, the correntropy based correlation (CRCO), for independent and identically distributed (i.i.d) symmetric alpha-stable ($S\alpha S$) random variables. We define the CRCO based matrix for the array sensor outputs and show that it can be applied with MUSIC to estimating DOAs in the presence of complex symmetric alpha-stable noise. The comprehensive Monte-Carlo simulation results demonstrate that CRCO–MUSIC outperforms the existing fractional lower order statistics (FLOS) based MUSIC algorithms especially in highly impulsive noise environments.

1. Introduction

The second order statistics has been the major methodology of considerable academic research for statistical array processing by assuming that the additive noise is Gaussian distributed [1]. For example, the well-known eigendecompostion-based high-resolution direction finding method, namely, the multiple signal classification (MUSIC) is to perform an eigendecompositon on the spatial covariance matrix of the array outputs to estimate the direction of arrivals (DOA) for the source signals [2]. For a fairly common case, it is inappropriate to model the noise as Gaussian. The high-order statistics has been applied to extract more information than the second-order methods for non-Gaussian situations. Mohler [3] performed eigenstructure analysis on the second-order convolution of the received data. Chiang and Nikias [4] developed a fourth-order ESPRIT algorithm based on generalized eigenstructure analysis. Porat and Friedlander [5] derived a MUSIC-like algorithm by performing eigendecompositon of the fourth-order cumulants of the array data. All these methods are based on finite second-order and high-order statistics assumption. However, a broader and increasingly important class of non-Gaussian phenomena encountered in practice can be characterized as impulsive [6]. In these scenarios, sudden bursts or sharp spikes are exhibited due to clutter sources such as the signals affected by mountains, forests and sea waves [7–9]. As a result, their density functions decay in the tails less rapidly than those of the Gaussian which even lead to infinite second or higher-order statistics.

Recent studies show that this impulsive environment can be well modeled by alpha stable distribution [6]. Alpha stable distribution is defined by its characteristic function $\varphi(\omega)=\exp(ja\omega-\gamma|\omega|^{\alpha})$, where α is the characteristic exponent taking values $0<\alpha\leq 2$, and γ is the dispersion which is similar to the variance of the Gaussian distribution. Specially, Gaussian processes are stable processes with $\alpha=2$. These processes satisfy the stability property which

* Corresponding author.
E-mail address: qiutsh@dlut.edu.cn (T. Qiu).

http://dx.doi.org/10.1016/j.sigpro.2014.04.033
0165-1684/

states that linear combinations of jointly stable variables are indeed stable. They arise as limiting processes of sums of i.i.d random variables via the generalized central limit theorem. These distributions have heavier tails than those of Gaussian distribution and possess finite pth-order moments only for $p<\alpha<2$, which induce that the methods based on the second order or the higher-order statistics have degraded dramatically in their performance. For details about $S\alpha S$ processes, see [6], and the references therein.

In order to suppress these impulsive outliers, fractional lower order statistics (FLOS) based subspace methods [10–12] have been developed to the DOA estimation problem. In [10], the robust covariation-based MUSIC (ROC-MUSIC) used covariations under the assumption that the signals and the additive noise are jointly $S\alpha S$. Liu and Mendel [11] relaxed the jointly $S\alpha S$ constraints for signals and noise and formulated fractional lower order moments (FLOM) of the received array data. In [12], the phased fractional lower order moment (PFLOM) based subspace algorithm was presented for retrieving the directions of arrival of signals. Despite the robustness of these FLOS based methods, they still have some limitations: (1) similar to the traditional MUSIC algorithm, these methods need a fairly high signal to noise (SNR) ratio to yield the robust estimation performance. However, in low SNR situations, a significant degradation in their performance may be encountered; (2) when the desired signals are not circularly symmetrical, these methods suffer from model mismatch and have poor performance especially in small α value or in low SNR situations.

By applying a translation-invariant kernel, the correntropy has been proposed as a new statistics that can quantify both the time structure and the statistical distribution of two stochastic random processes [13,14]. As correntropy starts directly from Parzen estimation of probability density function and involves higher-order statistics, it is a much stronger measure of independence for the random variables. Therefore, it can extract more information than the conventional correlation functions. In this paper, we address the issue of DOA estimation based on correntropy. We introduce a novel operator based on correntropy, namely the correntropy based correlation (CRCO), which can be used with MUSIC in the presence of a wide range of impulsive noise. The derived DOA estimation algorithm through combining CRCO with MUSIC, as we call the CRCO–MUSIC algorithm, has two appealing advantages which are listed as follows: (1) in low SNR conditions or in high impulsive environments, the CRCO–MUSIC exhibits an evident performance improvement over the FLOS based MUSIC algorithms; (2) the CRCO–MUSIC is robust not only for circular signals but also for noncircular signals, while the FLOS based MUSIC algorithms are limited to retrieving the DOA estimates from circular signals.

This paper is organized as follows: in Section 2, we define the problem of interest and review the existing FLOS based methods. In Section 3, we present a brief introduction on correntropy and propose the new operator for two i.i.d $S\alpha S$ random variables, namely the CRCO, which can be applied for the development of subspace techniques in the presence of α-stable distributed signals and noise. In Section 4, the CRCO based subspace method for DOA estimation is presented and the implementation of the algorithm is given. Finally, we demonstrate the simulation results in Section 5 and draw some conclusions in Section 6.

2. Problem definition and the FLOS based solutions

2.1. Problem definition

Assume a uniformly linear array of M sensors receives signals generated by P narrow-band independent, complex isotropic $S\alpha S$ $(1<\alpha\le 2)$ random processes with known center frequency ω and locations $\theta_1,\theta_2,\cdots,\theta_P$. The underlying noises are also complex isotropic $S\alpha S$ random processes with the same characteristic exponent α as the signals. Using the complex envelope representation, the array output can be expressed as

$$x_m(t)=\sum_{k=1}^{P} a(\theta_k)s_k(t)+n_m(t)\quad m=1,2,\cdots,M \tag{1}$$

where $a(\theta_k)=e^{-j\frac{2\pi}{\lambda}(m-1)d\ \sin(\theta_k)}$ is the steering coefficient of the mth sensor toward direction θ_k, λ is the wavelength of the carrier and d is the distance between two sensors, $s_k(t)$ is the kth signal received at the sensor array, and $n_m(t)$ is the underlying noise at the mth sensor.

Eq. (1) can be expressed as the following vector form:

$$\mathbf{x}(t)=\mathbf{A}(\boldsymbol{\theta})\mathbf{s}(t)+\mathbf{n}(t) \tag{2}$$

where $\mathbf{A}(\boldsymbol{\theta})$ is the $M\times P$ matrix of the array steering vectors $\mathbf{A}(\boldsymbol{\theta})=[\mathbf{a}(\theta_1),\mathbf{a}(\theta_2),\cdots,\mathbf{a}(\theta_P)]$, in which $\mathbf{a}(\theta_k)=[1,e^{-j\frac{2\pi}{\lambda}d\ \sin(\theta_k)},\cdots,e^{-j\frac{2\pi}{\lambda}d(M-1)\sin(\theta_k)}]^T$, $\mathbf{x}(t)$ is the $M\times 1$ vector of signals received by the array sensors $\mathbf{x}(t)=[x_1(t),x_2(t),\cdots,x_M(t)]^T$, $\mathbf{s}(t)$ is the $P\times 1$ vector of the signals $\mathbf{s}(t)=[s_1(t),s_2(t),\cdots,s_P(t)]^T$, and $\mathbf{n}(t)$ is the $M\times 1$ vector of the noise $\mathbf{n}(t)=[n_1(t),n_2(t),\cdots,n_M(t)]^T$.

Our aim is to estimate the directions of arrival $\theta_1,\theta_2,\cdots,\theta_P$ of the sources from $\mathbf{x}(t)$.

2.2. FLOS based MUSIC

The covariance based MUSIC algorithm plays an essential role in high-resolution direction finding under Gaussian noise assumption. However, due to the lack of finite variance, the covariance does not exist on the space of $S\alpha S$ random processes. Instead, the FLOS based statistics has been applied with MUSIC to obtain the DOA estimates [10–12].

Assume X and Y are jointly $S\alpha S$ random variables with the characteristic exponent $1<\alpha\le 2$, the covariation of X and Y is defined by

$$[X,Y]_\alpha=\frac{E\{XY^{\langle p-1\rangle}\}}{E\{|Y|^p\}}\gamma_Y\quad 1\le p<\alpha \tag{3}$$

where γ_Y is the dispersion of Y. For complex number z, $z^{\langle\beta\rangle}=|z|^{\beta-1}z^*$.

In [11], Liu simplified the covariation and defined the FLOM for X and Y through

$$W_{\mathrm{FLOM}}=E\{XY^{\langle p-1\rangle}\}\quad 1\le p<\alpha \tag{4}$$

Both covariation and FLOM are proved to be bounded for jointly $S\alpha S$ random variables. The operator $(\cdot)^{\langle p\rangle}$ is mainly used to suppress the outliers in a $S\alpha S$ distribution. Based on the well defined covariation and FLOM, the (i,j)th entries of two $M\times M$ matrices for the array output $\mathbf{x}(t)$ are formulated by $[x_i(t),x_j(t)]_\alpha$ in ROC-MUSIC and $E\{x_i(t)|x_j(t)|^{p-2}x_j^*(t)\}$ in FLOM-MUSIC respectively. However, we can see from both Eqs. (3) and (4) that $(\cdot)^{\langle p\rangle}$ only acts on the random variable Y but not on X at all. So the performance of either covariation or FLOM based MUSIC algorithms will degenerate significantly when the impulsiveness of both $x_i(t)$ and $x_j(t)$ is remarkable [12].

Belkacemi [12] defined the conjugate of the FLOM operator $X^{-\langle p\rangle}=(X^*)^{\langle p\rangle}=(X^{\langle p\rangle})^*$ and proposed the PFLOM based operator for X and Y by

$$W_{\mathrm{PFLOM}}=E\{X^{\langle\frac{b}{2}\rangle}Y^{\langle-\frac{b}{2}\rangle}\}\quad 0<b<\alpha \tag{5}$$

Thereby, the corresponding matrix for the array output $\mathbf{x}(t)$ based on PFLOM can be derived with the (i,j)th entry given by $E\{x_i^{\langle\frac{b}{2}\rangle}(t)x_j^{\langle-\frac{b}{2}\rangle}(t)\}$. Since the impulsiveness in $x_i(t)$ and $x_j(t)$ has been restrained simultaneously by the operator $(\cdot)^{\langle\frac{b}{2}\rangle}$ and $(\cdot)^{\langle-\frac{b}{2}\rangle}$, it is easy to understand why the PFLOM based MUSIC algorithm performs superior to the covariation and FLOM based ones especially in highly impulsive noise environments. However, compared with the performance in less impulsive noise situations (e.g. $\alpha>1.8$), all the FLOS based MUSIC algorithms (including the PFLOM based MUSIC) exhibit a distinct degradation in the strong impulsive noise environments (e.g. $\alpha<1.5$). In addition, for the scenario that noncircular signals (e.g. BPSK, AM signals) are contaminated by $S\alpha S$ noise, these FLOS based MUSIC algorithms perform poorer due to the model mismatch than they do in the scenario of circular signals [11].

3. Correntropy based correlation

Here, we propose a new operator based on correntropy, namely the correntropy based correlation, and then apply it with MUSIC to gain a more robust DOA estimation results in highly impulsive noise envrionments.

3.1. Correntropy

Inspired by kernel-based methods [15] and information theoretic learning (ITL) methods [16], based on the information potential (IP), the correntropy for two arbitrary random variables X and Y is defined as follows [14]:

$$V_\sigma(X,Y)=E[\kappa_\sigma(X-Y)] \tag{6}$$

where $\kappa_\sigma(\cdot)$ is the kernel function that satisfies Mercer's theory [17] and $E[\cdot]$ denotes the mathematical expectation. A brief interpretation for Mercer' theory and kernel function can be found in Appendix A.

Using a Taylor series expansion for the widely used Gaussian kernel, the correntropy can be rewritten as

$$V_\sigma(X,Y)=\frac{1}{\sqrt{2\pi}\sigma}\sum_{n=0}^{\infty}\frac{(-1)^n}{2^n\sigma^{2n}n!}E[(X-Y)^{2n}] \tag{7}$$

which involves all the even-order moments of the random variable $(X-Y)$. It is notable that the term corresponding to $n=1$ in (7) is proportional to $E(X^2)+E(Y^2)-2E(XY)$, which indicates that the conventional covariance function (the autocorrelation for zero-mean processes) is also included within correntropy. Liu [13] has compared correntropy with the constrained covariance [18] and concluded that the former is a much simpler, while possibly weaker measure of independence than the latter but a much stronger measure than the traditional covariance.

Correntropy can also induce a metric which can be described as

$$\mathrm{CIM}(X,Y)=\mathrm{sqrt}\{E[\kappa(0)-\kappa_\sigma(X-Y)]\} \tag{8}$$

Apparently, for Gaussian kernel, $\kappa(0)=1/\sqrt{2\pi}\sigma$. The metric correntropy induced metric (CIM) possesses a "mix norm" property, that is, it behaves like an L2 norm when two points are close, and like an L1 norm when the two points are getting apart and eventually like a L0 norm when the two points get further apart. This property demonstrates the inherent robustness to outliers of correntropy. Liu [13] analyzed the influence to CIM of the kernel size and concluded that a small kernel size leads to a tight linear region and to a large L0 region, while a larger kernel size will enlarge the linear region and shrink the L0 region, that is, the kernel size controls the level of the outlier suppression.

In practice, the joint probability density function is often unknown, and only a finite number of data $\{(x_i,y_i)\}_{i=1}^N$ for X and Y are available, where the sample estimator of correntropy can be obtained through

$$\hat{V}_\sigma(X,Y)=\frac{1}{N}\sum_{i=1}^{N}\kappa_\sigma(x_i-y_i) \tag{9}$$

The merit of correntropy is that it can convey information not only about the correlation but also the statistical distribution for stochastic processes as the correntropy possesses a lot of properties quantifying the PDF of the data directly. Meanwhile, correntropy is robust against outliers due to the inner product in the feature space computed via the Gaussian kernel. All these features would inspire us to develop new array signal processing methods using conventional correlation matrices, such as, signal and noise subspace decompositions, projections, etc.

3.2. Correntropy based correlation

In this paper, we define a new operator based on correntropy which can be applied in $S\alpha S$ processes for a wide range of $1<\alpha\le 2$, and that would make it as an effective substitute for the conventional correlation functions.

Theorem 1. *Let X and Y be i.i.d $S\alpha S$ random variables with the characteristic exponent $1<\alpha\le 2$, the correntropy-based correlation (CRCO) by applying Gaussian kernel of X and Y given by*

$$R_{\mathrm{CRCO}}=E\left[\exp\left(-\frac{|X-Y|^2}{2\sigma^2}\right)XY\right] \tag{10}$$

is bounded. Where σ is the kernel size. See Appendix B for the proof of the boundedness of CRCO.

As we know, the diagonal entries of the covariance matrix of the array outputs corresponding to the traditional MUSIC algorithm are the autocorrelation functions of $x_i(t)$ $\{i=1,2,\cdots,M\}$. However, Theorem 1 only defines the correntropy-based cross-correlation for two i.i.d $S\alpha S$ random variables. It's obvious that the CRCO will be infinite for $X=Y$, that is, the correntropy-based autocorrelation dose not exist. To solve this problem, we can further generalize the definition of (10) to be

$$R=E\left[\exp\left(-\frac{(X-\mu Y)^2}{2\sigma^2}\right)XY\right]\quad \mu\neq 1 \tag{11}$$

where μ is a given positive constant. Here we introduce the parameter μ to make sure the corretropy-based autocorrelation to be finite. The proof for the boundedness of (11) omitted in this paper is similar to the proof of Eq. (10).

4. Subspace techniques under CRCO framework

4.1. The CRCO based matrix of the array sensor outputs

In Proposition 1, we propose a new class of matrices of the array outputs based on the CRCO which can be used with MUSIC for DOA estimation.

Proposition 1. *Let* $\mathbf{R}$ *be an* $M\times M$ *matrix, whose* (i,j)th *entry* R_{ij} *is given by*

$$R_{ij}=E\left\{\exp\left(-\frac{|x_i(t)-\mu x_j^*(t)|^2}{2\sigma^2}\right)x_i(t)x_j^*(t)\right\}\quad \mu\neq 1\quad 1<\alpha\leq 2 \tag{12}$$

where x_i *and* x_j *are the ith and jth components of the vector* $\mathbf{x}(t)$ *respectively, and* α *is the characteristic exponent of the underlying noise in* $\mathbf{x}(t)$. μ *is a given positive constant.*

The key for applying $\mathbf{R}$ with MUSIC is that $\mathbf{R}$ can be separated into two subspaces, namely, the signal subspace and the noise subspace, therefore, we can conduct eigen-decomposition on $\mathbf{R}$ and formulate the corresponding 'MUSIC spectrum' to derive the DOA estimates. In fact, separating $\mathbf{R}$ is very complicated for correntropy does not present an homogenous moment property, however, it involves all the higher order moments of X and Y, and exhibits a 'robust statistics' property like the M-estimators.

In the following content, we will describe that R_{ij} would behave as a correlation function for x_i and x_j in Gaussian noise situations and transition to a covariation function in the noise environments with significant outliers. Therefore, extending the traditional covariance based MUSIC to the CRCO based one is straightforward due to the inherent robustness of $\mathbf{R}$ to impulsive noise environments.

Consider the cost function by applying the metric CIM in (8)

$$J(w)=E[\kappa(0)-\kappa_\sigma(x_i(t)-wx_j(t))] \tag{13}$$

where w is a weight parameter. Letting $e=x_i(t)-wx_j(t)$ and defining $\rho(e)=\kappa(0)-\kappa_\sigma(e)$, it is easy to obtain that ρ satisfies all the needed conditions that a robust M-estimate cost function should meet [13]. Therefore, minimizing $J(w)$ can be rewritten as minimizing

$$J(w)=E[\rho(e)] \tag{14}$$

The optimal weight parameter w can be determined by setting the first-order partial derivatives of $J(w)$ with respect to w to zero. This yields $E[\psi(e(t))x_j(t)]=0$, where $\psi(e)\triangleq\partial\rho(e)/\partial e$ is called the score function. And then, the following M-estimate normal equation is obtained:

$$C_{jj}(t)w=C_{ij}(t) \tag{15}$$

where $C_{jj}(t)=E[q(e(t))x_j(t)x_j^*(t)]$ and $C_{ij}(t)=E[q(e(t))x_i(t)x_j^*(t)]$ are called the M-estimate autocorrelation of $x_j(t)$ and the M-estimate cross-correlation of $x_i(t)$ and $x_j(t)$ by setting $q(e)\triangleq\psi(e)/e$ [19]. For $\rho(\cdot)=\kappa(0)-\kappa_\sigma(\cdot)$ with the Gaussian kernel function $\kappa_\sigma(\cdot)=\exp-\left(\frac{(\cdot)^2}{2\sigma^2}\right)$, we can obtain $q(e)=\exp-\left(\frac{e^2}{2\sigma^2}\right)$, thereby we can get $C_{ij}(t)$ expressed as follows:

$$C_{ij}(t)=E\left[\exp-\left(\frac{|x_i(t)-wx_j^*(t)|^2}{2\sigma^2}\right)x_i(t)x_j^*(t)\right] \tag{16}$$

Apparently, substituting w with μ, we can have R_{ij} in (12) described by C_{ij}, which means that the proposed CRCO is a special case of the M-estimate cross-correlation between $x_i(t)$ and $x_j(t)$ by employing the CIM square function as the M-estimate cost function $J(w)$ in (13). Unfortunately, CIM dose not induce a norm on the sample space. Fig. 1 shows the contours of the distance from X to the origin in a 2-D space.

From Fig. 1 we can observe that, when two points are close, CIM behaves like an L2 norm, outside of this L2 norm zone CIM behaves like an L1 norm, eventually approaches L0 norm as two points are further apart.

That is, when $e=x_i(t)-wx_j(t)$ is small which corresponds to less impulsive noise environments (eg. Gaussian noise), the cost function $J(w)$ can be equivalent to

$$J_1(w)=\|x_i(t)-wx_j(t)\|_2^2 \tag{17}$$

in this case, the corresponding M-estimate cross-correlation between $x_i(t)$ and $x_j(t)$ can be obtained as

$$C_{ij}^{J_1}(t)=E[x_i(t)x_j^*(t)] \tag{18}$$

which is the (i,j)th entry of the covariance matrix $E[\mathbf{x}(t)\mathbf{x}^{\mathrm{H}}(t)]$ exactly.

Then, when e gets larger, the CIM behaves as an L1 norm, that is, the cost function $J(w)$ is equivalent to

$$J_2(w)=\|x_i(t)-wx_j(t)\|_1^2 \tag{19}$$

Based on the orthogonality principle of linear regression for stable processes with $1\leq\alpha<2$ [6], we have

$$[x_i(t)-wx_j(t),x_j(t)]_1=0 \tag{20}$$

Employing the linear property of covariation on (20), we have

$$[x_i(t),x_j(t)]_1=w[x_j(t),x_j(t)]_1 \tag{21}$$

where $[x_i(t),x_j(t)]_1$ is exactly the (i,j)th entry of the covariation based matrix of the array outputs employed by ROC-MUSIC $p=1$ [10]. Furthermore, inspecting the relationship between the covariation in (3) and the FLOM in (4), and rewriting (21) with FLOM, we can get

$$E\{x_i(t)|x_j(t)|^{-1}x_j^*(t)\}=wE\{x_j(t)|x_j(t)|^{-1}x_j^*(t)\} \tag{22}$$

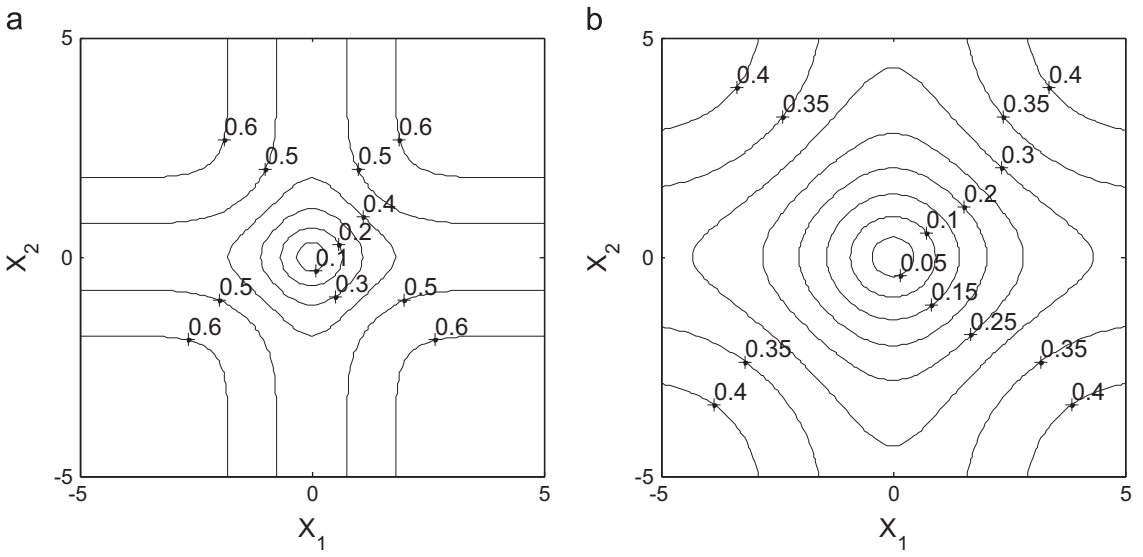

Fig. 1. Contours of CIM $(X,0)$ in 2-D sample space. (a) Kernel size $\sigma=1$ and (b) Kernel size $\sigma=2$.

interestingly, $C_{ij}^{J_2}(t)\triangleq E[x_i(t)|x_j(t)|^{-1}x_j^*(t)]$ in (22) is just the (i,j)th entry of the FLOM based matrix of the array outputs employed by FLOM-MUSIC [11] by setting $p=1$. See Appendix C for the properties of covariation.

On the basis of our analysis, we can conclude that the CRCO based array output matrix $\mathbf{R}$ should perform equivalent to the covariance matrix in MUSIC for Gaussian noise environments and to the covariation matrix in ROC-MUSIC for highly impulsive noise environments. Therefore, we can conduct eigen-decomposition on $\mathbf{R}$ and formulate the corresponding 'MUSIC spectrum' to gain the DOA estimates.

Moreover, as the scale of the CIM norm is fully controlled by the kernel size σ, that is, a small kernel size will lead to a tight linear region (L2 norm) and to a large L1 and L0 region while a large kernel size will enlarge the linear region and shrink the L1 and L0 region on the other hand, the transition process of $R_{ij}(t)$ from $C_{ij}^{J_1}(t)$ to $C_{ij}^{J_2}(t)$ is asymptotical which means we can extract a robust DOA estimation under a well tuned kernel size σ according to the level of the impulsiveness of the noise.

4.2. The implementation of CRCO–MUSIC

Here, we extend noise-subspace-based MUSIC for Gaussian noise to CRCO–MUSIC for $S\alpha S$ noise environments. The implementation procedures are as follows:

Step 1. Compute the $M\times M$ matrix $\hat{\mathbf{R}}$, whose (i,j)th entry is

$$\hat{R}_{ij}=\frac{1}{N}\sum_{t=1}^{N}\left[\exp\left(\frac{|x_i(t)-\mu x_j^*(t)|^2}{2\sigma^2}\right)x_i(t)x_j^*(t)\right]\quad \mu\neq 1 \tag{23}$$

Section 5 gives the recommendations for the selection of the kernel size σ and the parameter μ.

Step 2. Execute SVD on $\hat{\mathbf{R}}$ and construct the $M\times(M-P)$ matrix $\mathbf{E}_n\triangleq[\hat{e}_{P+1},\hat{e}_{P+2},\cdots,\hat{e}_M]$, where $\hat{e}_{P+1},\hat{e}_{P+2},\cdots,\hat{e}_M$ are the left singular vectors associated with the smallest $M-P$ singular values of $\hat{\mathbf{R}}$.

Step 3. Compute the corresponding CRCO–MUSIC spectrum

$$P_{\text{CRCO–MUSIC}}(\phi)=\frac{1}{\mathbf{a}^H(\phi)\mathbf{E}_n\mathbf{E}_n^H\mathbf{a}(\phi)}\quad -90^\circ\le\phi\le 90^\circ \tag{24}$$

where the $M\times 1$ is linear array steering vector $\mathbf{a}(\phi)$ is $\mathbf{a}(\phi)=[1\ \ e^{-j2\pi(d/\lambda)\sin\ \phi}\cdots e^{-j2\pi((M-1)d/\lambda)\sin\ \phi}]^T$.

Step 4. Select P local peaks of $P_{\text{CRCO–MUSIC}}(\phi)$ as the estimates of DOAs, $\{\hat{\phi}_1,\hat{\phi}_2,\cdots,\hat{\phi}_P\}$.

4.3. Kernel substitutes

Correntropy is directly related to the information potential (IP) of Renyi's quadratic entropy [14] where the probability of density function is estimated by Parzen kernels. From the viewpoint of kernel methods, the kernel function induces a nonlinear mapping Φ which transforms data from the input space to another high dimensional Hilbert space and then we can perform inner product of data pairs $\kappa(x_i,x_j)=\langle\Phi(x_i),\Phi(x_j)\rangle$ on this new feature space like the covariance function on the feature space induced by the input data. In [13], the translation invariant Gaussian kernel is applied to construct the correntropy. In fact, like in the Parzen estimation for PDF, other symmetric kernels can be used. As an example, we can reformulate the operator in (10) by applying Laplace kernel as

$$R_{Laplace}=E\left[\exp\left(-\frac{|X-Y|}{\sigma}\right)XY\right] \tag{25}$$

Accordingly, we can construct the corresponding matrix $\mathbf{Q}$ of the array output $\mathbf{x}(t)$ whose (i,j)th entry Q_{ij} is given by

$$Q_{ij}=E\left\{\exp\left(-\frac{|x_i(t)-\mu x_j^*(t)|}{\sigma}\right)x_i(t)x_j^*(t)\right\}\quad \mu\neq 1 \tag{26}$$

and then apply **Q** with MUSIC to gain the DOA estimates by a similar strategy to CRCO–MUSIC.

Apparently, the methodology for the construction of $R_{Laplace}$ is similar to the CRCO. Furthermore, we can also seek other symmetric kernels as the substitutes for the Gaussian kernel which is beyond the scope of this paper.

5. Simulations

In the simulations, we assume two independent signals of the same power are received by a linear array of 5 sensors which space a half-wavelength apart. The underlying noise is assumed to be complex isotropic $S\alpha S$ alpha-stable distributed. For $\alpha=2$, it is simplified to be Gaussian distributed. For $\alpha<2$, due to the lack of finite variance, the generalized signal-to-noise ratio (GSNR) is defined to evaluate the ratio of the signal power over noise dispersion γ by $\mathrm{GSNR}=10\log(E|s(t)|^2/\gamma)$.

The two signals are considered to be resolvable if the following resolution criterion holds [20]

$$\Lambda(\theta_1,\theta_2)\triangleq S(\theta_m)-\frac{1}{2}\{S(\theta_1)+S(\theta_2)\}>0 \tag{27}$$

where θ_1 and θ_2 are the two arrival angles for the two signals, $\theta_m=(\theta_1+\theta_2)/2$ is the mid-range between them, and the *null spectrum* $S(\theta)\triangleq 1/P(\theta)$ is defined as the reciprocal of the MUSIC spectrum. The performance of the algorithm is evaluated by two quantities: the probability of resolution and the mean square error (MSE) of the DOA estimates. The probability of resolution is defined as the ratio of the number of successful runs to the total number of the Monte-Carlo (MC) runs. MSE is the sample mean-squared error averaged for the two sources and can be defined by

$$\mathrm{MSE}\triangleq\frac{1}{2L}\sum_{l=1}^{L}(\hat{\theta}_1(l)-\theta_1)^2+\frac{1}{2L}\sum_{l=1}^{L}(\hat{\theta}_2(l)-\theta_2)^2 \tag{28}$$

where $\hat{\theta}_1(l)$ and $\hat{\theta}_2(l)$ are the respective estimates of θ_1 and θ_2 for the lth MC run, and L is the total number of the MC runs.

Considering the stability property of $S\alpha S$ processes which states that linear combinations of jointly stable variables are indeed stable, for analytic convenience, we set the assumption that the signals are $S\alpha S$ distributed and are with the same characteristic exponent α with the noise. However, our proposed algorithm is not limited to the scenarios stated by the above assumption. Considering $S\alpha S$ signals are not commonly encountered in practical applications, in Sections 5.1and 5.2, we verify the robustness of the proposed CRCO–MUSIC algorithm and compare its performance with three existing FLOS based MUSIC algorithms: ROC-MUSIC, FLOM-MUSIC, and PFLOM-MUSIC, for those frequently encountered communication signals, such as quaternary amplitude modulation (QAM), binary phase-shift keying (BPSK), quaternary phase-shift keying (QPSK), and Gaussian sources. Each simulation experiment is executed over 200 independent MC runs. In Section 5.1, we assume two independent QAM communication signals representing circular signals of the same power are received by a linear array of 5 sensors. The directions of arrival are $\theta_1=5^\circ$ and $\theta_2=15^\circ$ respectively. The performance of these algorithms is reviewed as a function of four parameters, namely the GSNR, the noise characteristic exponent α, the number of snapshots, and the angular separation of the two sources. In Section 5.2, three different types of circular signals, namely, the QAM signals, the QPSK signals, and the circularly symmetrical Gaussian signals, added with the BPSK signals representing non-circular signals embedded in $S\alpha S$ noise respectively, are employed to assess the performance of these algorithms. In Section 5.3, to consist with the $S\alpha S$ distribution assumption for both the signals and the noise in Section 3, and to give some references for other possible applications, we further examine the robustness of CRCO–MUSIC under the assumption that the signals and noise are both $S\alpha S$ distributed.

For the FLOS based MUSIC algorithms, we often have to estimate the α value from the noise contaminated signals because the selection of the parameter p or b are directly related to the value of α, so that we can ensure the boundedness of the (i,j)th entries $[x_i(t),x_j(t)]_\alpha$, $E\{x_i(t)|x_j(t)|^{p-2}x_j^*(t)\}$ and $E\{x_i^{\langle b\rangle}(t)x_j^{\langle -b\rangle}(t)\}$, which are utilized to formulate the corresponding FLOS based sensor array output matrices. However, it is often not easy to evaluate the α value in priori especially in the case that the value of α dynamically changes. For ROC-MUSIC and FLOM-MUSIC, one straightforward solution is to choose $p=1$ since the characteristic exponent α satisfies $1<\alpha<2$ and p should meet $1\le p<\alpha$. It is understandable due to Eqs. (3)–(5) that the smaller the value of p is, the deeper the outliers are suppressed. Studies also show that the performance of ROC-MUSIC and FLOM-MUSIC is better by $p=1$ than by other p values [10,11]. Therefore, $p=1$ is chosen for ROC-MUSIC and FLOM-MUSIC in the following simulations. For PFLOM-MUSIC, referring to the simulation results in [12], the parameter b is chosen as $b=0.2$ for performance comparison.

Finally, in Section 5.4, we discuss the selection of the parameters of CRCO–MUSIC algorithm. The value of the kernel size σ and the parameter μ in Sections 5.1, 5.2, and 5.3 is chosen under the guidance of the results in Section 5.4.

5.1. DOA estimation in circular signals

5.1.1. Experiment 1: effect of GSNR

Fig. 2 illustrates the performance of CRCO–MUSIC and the FLOS based MUSIC algorithms under a wide range of GSNRs from 1 dB to 10 dB. The number of snapshots available to the algorithms is 1000. A moderate impulsive noise condition with $\alpha=1.6$ is employed in this experiment. From Fig. 2 we can see that CRCO–MUSIC outperforms the other algorithms both in resolution probability and in MSE. Another observation is that the corresponding curves of ROC-MUSIC and FLOM-MUSIC nearly overlap which indicate the performance of ROC-MUSIC and FLOM-MUSIC are approximately the same. The same conclusion has been drawn in [11]. It is easy to understand this behavior as the covariation and the FLOM are formulated similarly and only differ with each other by the constant $\gamma_Y/E\{|Y|^p\}$ when the GSNR is fixed. For this reason, in the following experiments, we will not review the performance of ROC-MUSIC algorithm and we will

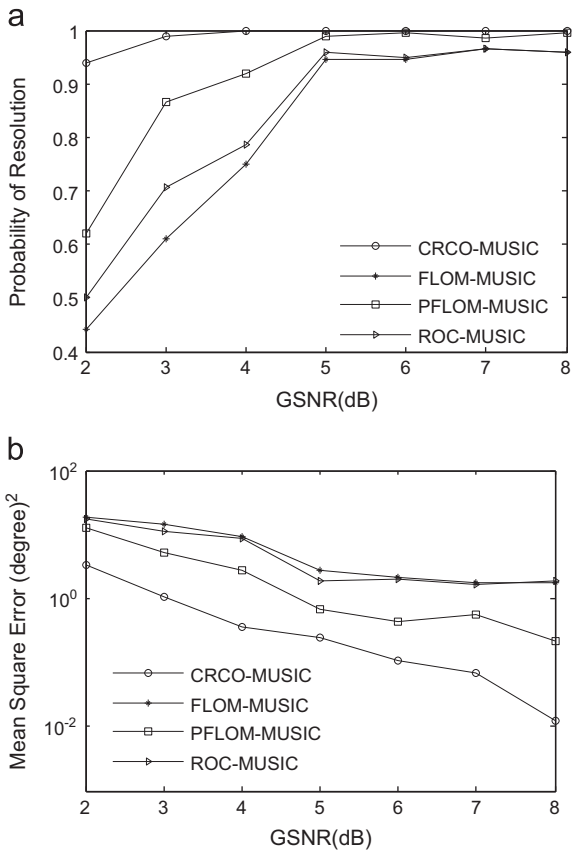

Fig. 2. Performance as a function of the GSNR ($\alpha = 1.6$). (a) Probability of resolution. (b) Mean square error.

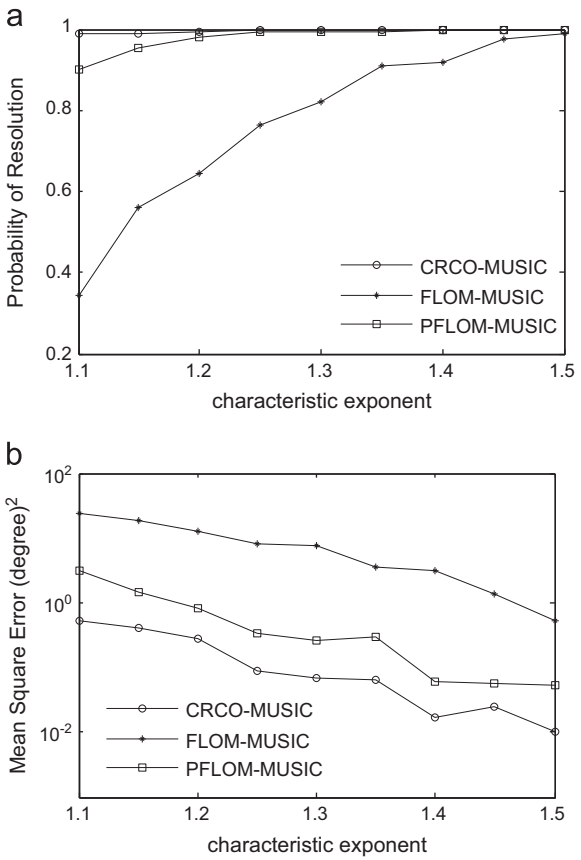

Fig. 3. Performance as a function of the characteristic exponent α. (a) Probability of resolution. (b) Mean square error.

only perform the performance comparison between CRCO–MUSIC and the other two fractional lower order statistics based MUSIC algorithms.

5.1.2. Experiment 2: effect of the characteristic exponent α

In this experiment, two directions of arrival are $\theta_1 = 5^\circ$ and $\theta_2 = 15^\circ$. The robustness of CRCO–MUSIC algorithm to strong impulsive noise environments is tested in the presence of $S\alpha S$ noise from extremely impulsive with $\alpha = 1.1$ to moderately impulsive with $\alpha = 1.5$. The GSNR is set as a constant at 10 dB. The number of snapshots available to the algorithms is 1000.

Fig. 3 depicts the overall performance improvement of CRCO–MUSIC over the FLOS based MUSIC algorithms both in resolution probability and in MSE. The results would suggest that it is beneficial to employ the CRCO-based MUSIC instead of the FLOS-based MUSIC methods if the quite impulsive noise environments are involved.

5.1.3. Experiment 3: effect of the number of snapshots M

The performance in terms of resolution probability and MSE as a function of the number of snapshots are investigated in this experiment and the results are exhibited in Fig. 4(a) and (b). The directions of arrival of the two incoming signals are $\theta_1 = 5^\circ$ and $\theta_2 = 15^\circ$. For this experiment, we choose $\alpha = 1.6$ corresponding to a moderate impulsive noise environment. The GSNR is kept to be 10 dB. From Fig. 4 we can observe that CRCO–MUSIC gains a more evident decrease in MSE than the FLOS based MUSIC algorithms as the number of the snapshots increases.

5.1.4. Experiment 4: angular separation

The capability of angle separation for the two incoming signals of the CRCO–MUSIC versus the other algorithms is reviewed in this experiment. A moderate impulsive noise environment with $\alpha = 1.6$ is employed for this experiment and the GSNR is kept at 10 dB. The number of snapshots available to the algorithms is 1000. Fig. 5 displays the simulation results. By contrast with the performance of the two FLOS based MUSIC algorithms, the proposed CRCO–MUSIC exhibits the robustness to the angle separation. That is, for a given probability of resolution, CRCO–MUSIC always gains a much lower angle separation threshold than the FLOS based MUSIC algorithms; and for a given angular separation, CRCO–MUSIC always gains a much lower MSE than the FLOS based ones.

5.2. DOA estimation in noncircular signals

DOA estimation for noncircular signals has brought considerable interest in array signal processing for recent years. The well-known NC-MUSIC algorithm is to formulate the extended spatial covariance matrix for the array

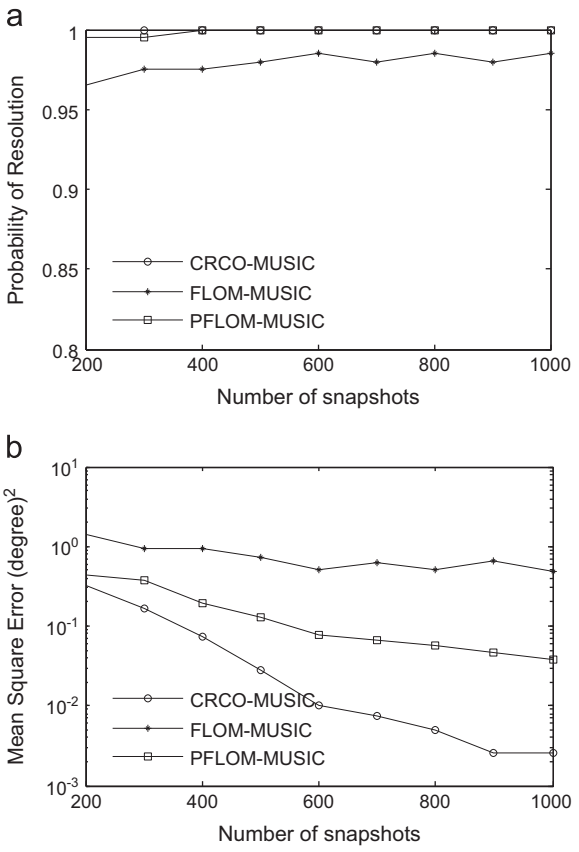

Fig. 4. Performance as a function of the number of snapshots. (a) Probability of resolution. (b) Mean square error.

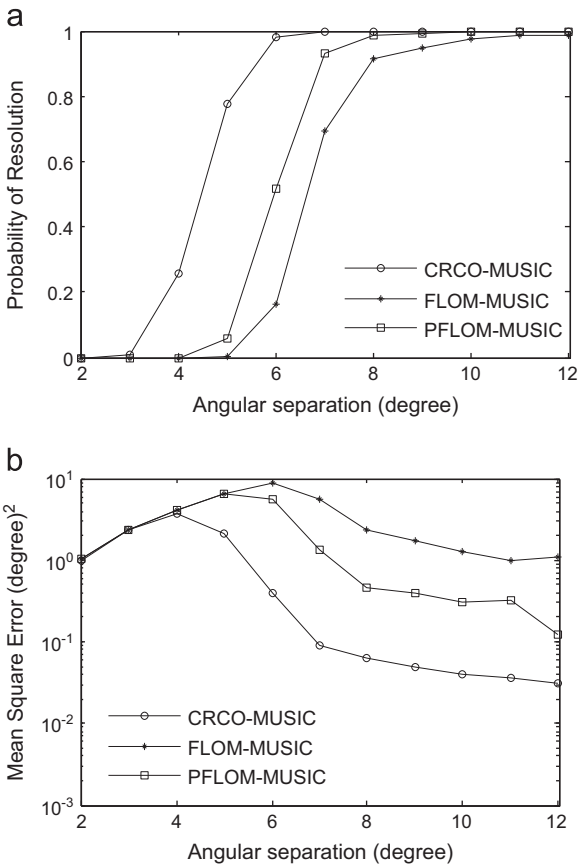

Fig. 5. Performance as a function of the source angular separation. (a) Probability of resolution. (b) Mean square error.

outputs by $\mathbf{R}_{\tilde{\mathbf{x}}} = {}^{\text{def}} E[\tilde{\mathbf{x}}_t \tilde{\mathbf{x}}_t^H]$ where $\tilde{\mathbf{x}}_t = {}^{\text{def}} \begin{pmatrix} \mathbf{x}_t \\ \mathbf{x}_t^* \end{pmatrix}$, and to conduct a MUSIC like strategy on it to gain DOA estimates for noncircular signals [21]. Besides the utilization of second order statistics in NC-MUSIC algorithm, Ref. [22] employed higher order statistics to gain better estimation performance. The advantages of these methods mainly lie in two aspects: (1) they have the ability of aperture expansion, that is, for the number of the array sensors M, these methods can handle more than $M-1$ incoming signals. (2) the estimation performance has been improved significantly especially in low SNRs compared with the traditional MUSIC algorithms. However, these methods still have drawbacks. For example, they achieve their robust performance under the assumption of finite second order of higher order statistics of signals and noise, which does not hold for alpha-stable processes.

In this experiment, we employ MC simulations to examine the robustness of CRCO–MUSIC for not only circular signals but also noncircular signals under four scenarios. Scenario 1, 2 and 3 are for two QAM, two QPSK and two circularly symmetrical Gaussian signals which stand for circular signals embedded in the additive $S\alpha S$ noise. Scenario 4 is for two BPSK signals representing noncircular signals involved in $S\alpha S$ noise. Table 1 depicts the performance in MSE of CRCO–MUSIC versus the FLOS based MUSIC algorithms under different noise conditions.

A distinct observation from Table 1 is that the MSE of FLOM-MUSIC algorithm and PFLOM–MUSIC algorithm corresponding to circular signals are relatively close. However, both algorithms exhibit poor performance for noncircular signals which indicates that these two FLOS-based MUSIC algorithms are limited to circular signals. By contrast, CRCO–MUSIC still yields quite good performance for noncircular signals. Moreover, inspired by NC-MUSIC, we can further improve the estimation accuracy and the capability of aperture expansion by formulating the extended spatial CRCO based matrix which is beyond the scope of this paper and omitted here.

5.3. DOA estimation for $S\alpha S$ signals

In this simulation, we assume two independent $S\alpha S$ signals impinge on the array sensors. The characteristic exponent α of the signals is set at $\alpha = 1.6$. The directions of arrival are $\theta_1 = 5^\circ$ and $\theta_2 = 15^\circ$. The performance of the algorithms is compared for the noise environments changing from extremely impulsive environments ($\alpha = 1.1$) to moderate impulsive ones ($\alpha = 1.6$). The GSNR is kept to be 10 dB. Fig. 6 shows that CRCO–MUSIC still yields better

Table 1
Performance of CRCO–MUSIC under four signal types measured by MSE.

α		1.3	1.4	1.5	1.6	1.7	1.8	1.6					
GSNR (Db)		8						3	4	5	6	7	8
Method	Signal type	Mean square error (degree^2)						Mean square error (degree^2)					
CRCO–MUSIC	QAM	0.265	0.168	0.100	0.070	0.033	0.033	8.220	2.353	0.483	0.258	0.125	0.070
	QPSK	0.275	0.033	0.028	0.012	0.003	0.003	0.993	0.345	0.218	0.120	0.050	0.012
	Gaussian	0.520	0.215	0.203	0.155	0.133	0.123	13.435	3.203	0.643	0.388	0.238	0.155
	BPSK	0.410	0.348	0.233	0.198	0.120	0.068	2.853	0.705	0.373	0.293	0.238	0.198
FLOM-MUSIC	QAM	22.270	12.773	5.755	2.768	1.638	0.233	24.633	17.333	11.760	5.935	4.190	2.768
	QPSK	16.125	6.743	2.475	1.578	0.633	0.078	15.380	7.685	5.043	3.440	3.070	1.578
	Gaussian	20.948	13.690	7.018	3.108	1.620	0.590	26.195	18.920	14.858	7.813	4.345	3.108
	BPSK	25.813	25.045	22.598	18.438	12.198	6.168	22.015	23.548	23.938	24.148	23.510	18.438
PFLOM-MUSIC	QAM	9.553	4.250	1.488	0.458	0.080	0.078	19.188	8.460	4.070	1.540	1.378	0.458
	QPSK	5.753	2.405	0.613	0.168	0.195	0.023	6.295	3.458	1.658	0.778	0.153	0.168
	Gaussian	13.505	5.758	1.570	0.873	0.338	0.095	18.963	12.780	5.680	2.963	1.593	0.873
	BPSK	20.948	23.128	23.280	24.128	25.778	25.645	14.885	14.713	16.240	21.075	24.468	24.128

performance over the other FLOS based MUSIC algorithms for $S\alpha S$ signals.

5.4. Parameter selection

In this section, we assume two independent QAM communication signals impinge on the array sensors. The directions of arrival are $\theta_1 = 5^\circ$ and $\theta_2 = 15^\circ$

5.4.1. Experiment 1: the selection of the kernel size

As shown in Section 3, correntropy behaves as a 'robust statistics' like M-estimators. It induces a new metric, namely CIM which is equivalent to the 2-norm, 1-norm and zero-norm, corresponding to the Euclidean zone, Transition zone and Rectification zone, when the points are close, getting further apart and far apart respectively. A small kernel size leads to a tight Euclidean region and to a large L0 region, while a larger kernel size will enlarge the linear region, that is, the scale of the metric is fully controlled by the kernel size σ. The essence of the robustness of correntropy based methods in impulsive environments is that correntropy removes the outliers by keeping them in the Rectification zone and approaches the data in the L2 norm sense simultaneously. Accordingly in this paper, we investigate the kernel size σ such that the noise-free signal dynamics lie in the Euclidean zone while the impulsive outliers are kept in Rectification zone as follows:

$$\sigma = l\hat{\sigma}_s \tag{29}$$

where l is a scale factor and $\hat{\sigma}_s^2$ is the estimated variance of the noise-free signal $s(t)$ which can be estimated through

$$\hat{\sigma}_s^2 = \frac{1}{N}\sum_{n=1}^{N}\hat{\sigma}_s^2(n) \tag{30}$$

where N is the number of snapshots and $\hat{\sigma}_s^2(n)$ is the instantaneous signal power of $s(t)$. Considering that the array output signal $x(t)$ received by the sensors which is contaminated by noise is the only signal we can handle with, $\hat{\sigma}_s^2(n)$ can be estimated through the following formula:

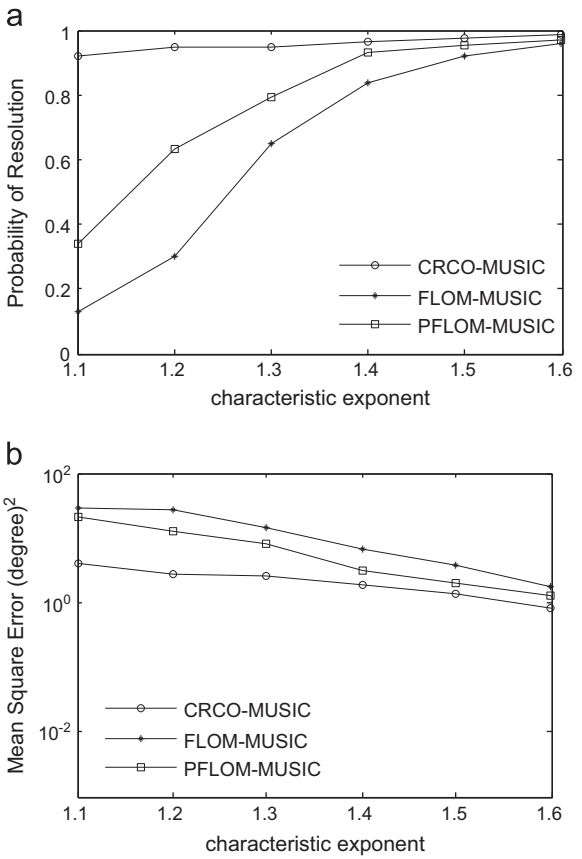

Fig. 6. Performance of CRCO–MUSIC for $S\alpha S$ signals. (a) Probability of resolution; and (b) mean square error.

$$\hat{\sigma}_s^2(n) = \rho\hat{\sigma}_s^2(n-1) + (1-\rho)med(A_x(n)) \tag{31}$$

where $A_x(n) = \{x^2(n), \cdots, x^2(n-N_w+1)\}$, $med(\cdot)$ denotes the sample median operation which is applied to remove the

additive impulsive noise in $x(t)$, N_w is the length of the estimation window and ρ is the forgetting factor.

In the following simulations, we set $\rho = 0.99$, $N_w = 5$ and $\sigma_s^2(1) = x^2(1)$ respectively and examine the estimation performance of CRCO–MUSIC as the scale factor l increases from 0.5 to 2.5, the results are depicted in Fig. 7.

From Fig. 7 we can observe that $l \in [1.3–1.5]$ would be the optimal domain for CRCO–MUSIC to achieve its best performance. Actually, compared with the improvements over the other FLOS-based methods, the variation in terms of MSE of CRCO–MUSIC is relatively small over a wide range of $l \in [1–2]$.

5.4.2. Experiment 2: the selection of the suppression parameter μ

As analyzed in [11], the reason for that the FLOM-based matrix of the array outputs can be processed by MUSIC to retrieve DOAs is its (i,j)th entry $E\{x_i(t)|x_j(t)|^{p-2}x_j^*(t)\}$ is the cross-correlation between $x_i(t)$ (which is a linear combination of the signals) and an FLOM-based function of $x_j(t)$ (which is used to reduce the effects of impulsive noise). From another perspective, we can consider that the FLOM operator implements suppression for different levels of outlier on $x_i(t)$ and $x_j(t)$, that is, only the outliers in $x_j(t)$ has been suppressed by the operator $x_j(t)^{\langle p-1\rangle}$ $(1 \le p < \alpha)$ while the amplitude in $x_i(t)$ remains unchanged.

Inspired by the FLOM operator, we set the parameter μ in Eq. (12) to control the outlier suppression level on $x_j(t)$. Apparently, $\mu = 0$ means the outlier suppression only affects $x_i(t)$. $\mu = 1$ indicates the same outlier suppression level has been implemented on both $x_i(t)$ and $x_j(t)$. In the following experiment, we perform MC simulations to examine the influence of the parameter μ on CRCO–MUSIC inside the range $0 \le \mu < 1$. The kernel size is fixed at $\sigma = 1.5 \cdot \hat{\sigma}_s$ as recommended in Section 5.4.1. $\mu > 1$ will not be considered since that will result in an exorbitant suppression level so that the (i,j)th entry in the matrix in Eq. (12) is too small to give reasonable results. The simulation results are demonstrated in Fig. 8. As we can see from Fig. 8, CRCO–MUSIC yields its best performance for the μ within the interval [0.2–0.5].

Based on the above analysis, the simulations in Sections 5.1, 5.2 and 5.3 are executed by setting $\sigma = 1.4\hat{\sigma}_s$ and $\mu = 0.5$.

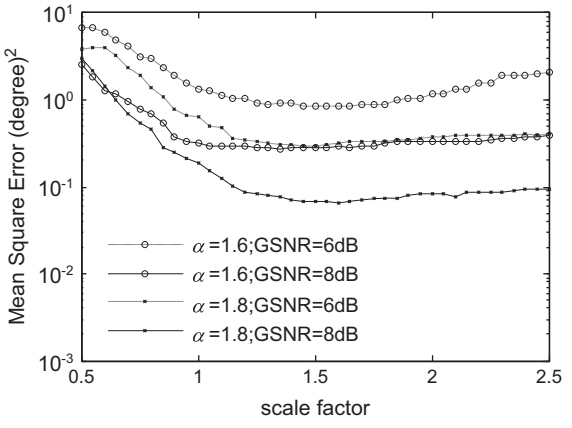

Fig. 7. The performance of CRCO–MUSIC as the factor of l.

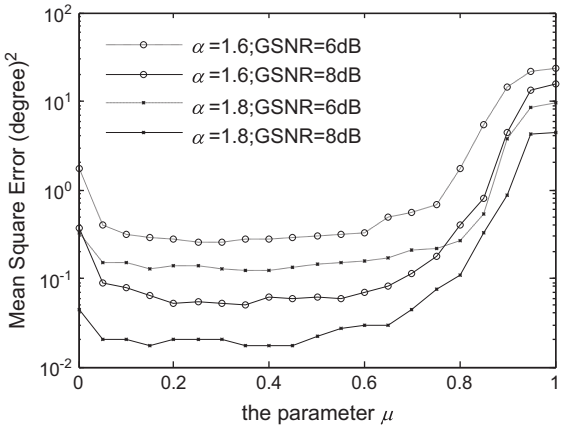

Fig. 8. Performance of CRCO–MUSIC as a function of μ.

6. Conclusion

This paper proposes a new operator referred to as the correntropy based correlation (CRCO) based on correntropy. It also proves the CRCO is bounded for i.i.d $S\alpha S$ random processes. We formulate the CRCO-based matrix of the array outputs and apply it with MUSIC to obtain DOA estimates. We show that the CRCO behaves like correlation in Gaussian noise and covariation in impulsive noise environments. The transition process is asymptotical and controlled by the kernel size. Experimental results illustrate that the CRCO–MUSIC can outperform the FLOS based MUSIC algorithms especially in highly impulsive noise environments.

Acknowledgments

The authors acknowledge the support form National Natural Science Foundation of China under Grants 60872122, 61172108, 61139001 and 81241059, and would like to thank the anonymous reviewers for their useful comments and suggestions that significantly improved the paper.

Appendix A. Mercer's theory and kernel function

Kernel methods, including support vector machines, kernel principal component analysis, and kernel canonical correlation analysis, have been widely used in some problems, such as classifications, filtering, predications, projections, etc. The basic idea of kernel algorithms is to transform the data $\mathbf{x}_i$ from the input space to a high dimensional feature space of vectors $\Phi(\mathbf{x}_i)$, where the inner products can be computed using a positive definite kernel function satisfying Mercer's conditions: $\kappa(\mathbf{x}_i, \mathbf{x}_j) = \langle \Phi(\mathbf{x}_i), \Phi(\mathbf{x}_j) \rangle$. A real-valued function $\kappa(u,v)$ is said to fulfill Mercer's condition if for all square integrable functions $g(u)$, one has $\iint \kappa(u,v)g(u)g(v)dudv \ge 0$.

Appendix B. The proof of Theorem 1

Let $R^1 = -R = E\left[\exp\left(-\frac{|X-Y|^2}{2\sigma^2}\right)(-XY)\right]$, we have

$$R^1 \le \frac{1}{2}E\left[\exp\left(-\frac{|X-Y|^2}{\sigma}\right)(X-Y)^2\right] \tag{32}$$

For the two i.i.d random variables X, Y with stable distribution and two arbitrary constants ξ_1 and ξ_2, there exists U which has the same distribution as X and constants ξ and c such that [6]

$$\xi_1 X + \xi_2 Y \overset{d}{=} \xi U + c \tag{33}$$

where the notation $X_1 \overset{d}{=} X_2$ means that X_1 and X_2 have the same distribution. The formula (33) shows that $X-Y$ obeys the same distribution as X or Y. Let $X-Y \overset{d}{=} Z$, where Z has the same distribution with X or Y, we have

$$2R^1 \le E\left[\exp\left(-\frac{|X-Y|^2}{\sigma}\right)(X-Y)^2\right] \Leftrightarrow E\left[\exp\left(-\frac{|Z|^2}{2\sigma^2}\right)Z^2\right] = \int_{-\infty}^{+\infty}\exp\left(-\frac{z^2}{2\sigma^2}\right)z^2 f(z)dz \tag{34}$$

where $f(z)$ is the probability density function (p.d.f) of z and can be given by

$$f(z) = \frac{1}{2\pi}\int_{-\infty}^{+\infty}\varphi(\omega)e^{-j\omega z}d\omega \tag{35}$$

For symmetric alpha stable variable z with the location parameter $a=0$, the characteristic function is denoted as $\varphi(\omega) = e^{-\gamma|\omega|^\alpha}$. Substituting $\varphi(\omega)$ into (35), we have Eq. (34) expressed as follows:

$$E\left[\exp\left(-\frac{Z^2}{2\sigma^2}\right)Z^2\right] = \frac{1}{2\pi}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\exp\left(-\frac{z^2}{2\sigma^2}\right)z^2 e^{-j\omega z}\exp(-\gamma|\omega|^\alpha)dzd\omega \tag{36}$$

By using Euler's formula on $e^{-j\omega z}$ in (36), and let $R_1 \triangleq \frac{1}{2\pi}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\exp(-\frac{z^2}{2\sigma^2})z^2\exp(-\gamma|\omega|^\alpha)(\cos\ \omega z)dzd\omega$ and $R_2 \triangleq -\frac{j}{2\pi}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\exp(-\frac{z^2}{2\sigma^2})z^2\exp(-\gamma|\omega|^\alpha)(\sin\ \omega z)dzd\omega$, Then, Eq. (36) becomes

$$E\left[\exp\left(-\frac{Z^2}{2\sigma^2}\right)Z^2\right] = R_1 + R_2 \tag{37}$$

Let us rewrite R_2 as $R_2 = -\frac{j}{2\pi}\int_{-\infty}^{+\infty}g_2(\omega)d\omega$, it is obvious that $g_2(\omega)$ is an odd function on ω from $-\infty$ to $+\infty$. Thus, we have $R_2 = \frac{1}{2\pi}\int_{-\infty}^{+\infty}g_2(\omega)d\omega = 0$.

Similarly, let $R_1 \triangleq \frac{1}{2\pi}\int_{-\infty}^{+\infty}g_1(\omega)d\omega$, where $g_1(\omega) = \int_{-\infty}^{+\infty}\exp(-\frac{z^2}{2\sigma^2})z^2\exp(-\gamma|\omega|^\alpha)(\cos\ \omega z)dz$ is an even function on ω from $-\infty$ to $+\infty$, Therefore,

$$R_1 = \frac{1}{\pi}\int_0^{+\infty}g_1(\omega)d\omega \le \frac{1}{\pi}\int_0^{+\infty}\int_{-\infty}^{+\infty}\exp\left(-\frac{z^2}{2\sigma^2}\right)z^2 dz\exp(-\gamma|\omega|^\alpha)d\omega \tag{38}$$

where $\int_{-\infty}^{+\infty}\exp\left(-\frac{z^2}{2\sigma^2}\right)z^2dz = \sqrt{2\pi}\sigma^3$, then we have

$$\frac{\sqrt{2\pi}}{\pi}\sigma^3\int_0^{+\infty}\exp(-\gamma|\omega|^\alpha)d\omega \le \sqrt{\frac{2}{\pi}}\sigma^3\int_0^{+\infty}\exp(-\gamma\omega)d\omega = \sqrt{\frac{2}{\pi}}\frac{\sigma^3}{\gamma} \quad 1<\alpha\le 2 \tag{39}$$

Eq. (39) shows that $2R^1 \le \sqrt{\frac{2}{\pi}}\frac{\sigma^3}{\gamma}$, which means the lower bound of R is $R \ge -\sqrt{\frac{2}{\pi}}\frac{\sigma^3}{\gamma}$.

Next, by investigating $R = E\left[\exp\left(-\frac{|X-Y|^2}{2\sigma^2}\right)XY\right]$, we can also get

$$R \le E\left[\left|\exp\left(-\frac{|X-Y|^2}{2\sigma^2}\right)XY\right|\right] \le E[|XY|] \tag{40}$$

Under the independence assumption between X and Y, we have

$$R \le E[|XY|] = E[|X|]E[|Y|] \tag{41}$$

For the random variable X with the characteristic exponent $1<\alpha\le 2$, there exists [6]

$$E[|X|^p] = 2^{p+1}\frac{\Gamma\left(\frac{p+1}{2}\right)\Gamma\left(-\frac{p}{\alpha}\right)}{\alpha\sqrt{\pi}\Gamma(-p/2)}\gamma^{p/\alpha}\quad 0<p<\alpha \tag{42}$$

where $\Gamma(\cdot)$ is the usual gamma function defined by $\Gamma(\cdot) = \int_0^\infty t^{x-1}e^{-t}dt$.

Substituting p with 1 in Eq. (42), the upper bound of Rcan be given by

$$R \le \left[4\frac{\Gamma(1)\Gamma(-1/\alpha)}{\alpha\sqrt{\pi}\Gamma(-1/2)}\gamma^{1/\alpha}\right]^2 \tag{43}$$

Combining (39) with (43), we obtain

$$-\sqrt{\frac{2}{\pi}}\frac{\sigma^3}{\gamma} \le R \le \left[4\frac{\Gamma(1)\Gamma(-1/\alpha)}{\alpha\sqrt{\pi}\Gamma(-1/2)}\gamma^{1/\alpha}\right]^2 \tag{44}$$

which means that R_{CRCO} is bounded. This is the end of the proof. □

Appendix C. The properties of covariation

1. If X_1, X_2 and Y are jointly $S\alpha S$, then
$$[a_1X_1 + a_2X_2, Y]_\alpha = a_1[X_1,Y]_\alpha + a_2[X_2,Y]_\alpha \tag{45}$$
for any complex constants a_1 and a_2.
2. If Y_1 and Y_2 are independent and Y_1, Y_2, X are jointly $S\alpha S$, then
$$[X, b_1Y_1 + b_2Y_2]_\alpha = b_1^{\langle\alpha-1\rangle}[X,Y_1]_\alpha + b_2^{\langle\alpha-1\rangle}[X,Y_2]_\alpha \tag{46}$$
for any complex constants b_1 and b_2.
3. If X and Y are independent $S\alpha S$, then $[X,Y]_\alpha = 0$.

References

[1] D. Johnson, D. Dudgeon, Array Signal Processing: Concepts and Techniques, Prentice-Hall, Englewood Cliffs, NJ, 1993.
[2] R.O. Schmidt, Multiple emitter location and signal parameter estimation, IEEE Trans. Antennas Propag. 34 (1986) 276–280.
[3] R.R. Mohler, A second-order eigenstructure array processor, in: Proceedings of the Workshop Higher Order Spectral Ana, Vail, CO, 1989, pp. 152–157.
[4] H.H. Chiang, C.L. Nikias, The ESPRIT algorithm with high order statistics, in: Proceedings of the Workshop Higher Order Spectral Ana, Vail, CO, 1989, pp. 168–168.
[5] B. Porat, B. Friedlander, Direction finding algorithms based on high-order statistics, IEEE Trans. Signal Process. 39 (9) (1991) 2016–2023.
[6] C.L. Nikias, M. Shao, Signal Processing with α-Stable Distribution and Applications, John Wiley&Sons, New York, 1995.
[7] K.L. Blackard, T.S. Rappaport, C.W. Bostian, Measurements and models of radio frequency impulsive noise for indoor wireless communications, IEEE J. Sel. Areas Commun. 11 (7) (1993) 991–1001.

[8] M.D. Button, J.G. Gardiner, I.A. Glover, Measurement of the impulsive noise environment for satellite-mobile radio systems at 1.5 GHz, IEEE Trans. Veh. Technol. 51 (3) (2002) 551–560.
[9] A. Ahandra, Measurements of radio impulsive noise from various sources in an indoor environment at 900 MHz and 1800 MHz, in: Proceedings of the 13th IEEE International Symposium on Personal, Indoor and Mobile Radio Communications, 2, 2002, pp. 639–643.
[10] P. Tsakalides, C.L. Nikias, Robust covariation-based MUSIC (ROC-MUSIC) algorithm for bearing estimation in impulsive noise environments, IEEE Trans. Signal Process. 44 (1996) 1623–1633.
[11] T.H. Liu, J.M. Mendel, A subspace-based direction finding algorithm using fractional lower order statistics, IEEE Trans. Signal Process. 49 (2001) 1605–1613.
[12] H. Belkacemi, S. Marcos, Robust subspace-based algorithms for joint angle/Doppler estimation in non-Gaussian clutter, Signal Process. 87 (2007) 1547–1558.
[13] W.F. Liu, P.P. Pokharel, J.C. Principe, Correntropy: properties and applications in non-Gaussian signal processing, IEEE Trans. Signal Process. 55 (2007) 5286–5298.
[14] I. Santamaria, P.P. Pokharel, J.C. Principe, Generalized correlation function: definition, properties, and application to blind equalization, IEEE Trans. Signal Process. 54 (2006) 2187–2197.
[15] E. Parzen, Statistical methods on time series by ilbert space methods, Applied Mathematics and Statistics Laboratory, Stanford University, Stanford, CA, Technical Report 23, 1959.
[16] J.C. Principe, D. Xu, J. Fisher, Information theoretic learning, in: S. Haykin (Ed.), Unsupervised Adaptive Filtering, Wiley, New York, 2000.
[17] V. Vapnik, The Nature of Statistical Learning Theory, Springer Verlag, New York, 1995.
[18] A. Gretton, R. Herbrich, A. Smola, O. Bousquet, B. Scholkopf, Kernel methods for measuring independence, J. Mach. Learn. Res. 6 (2005) 2075–2129.
[19] S.C. Chan, Y.X. Zou, A recursive least M-estimate algorithm for robust adaptive filtering in impulsive noise: fast algorithm and convergence performance analysis, IEEE Trans. Signal Process. 52 (4) (2004) 975–991.
[20] Q.T. Zhang, Probability of resolution of the MUSIC algorithm, IEEE Trans. Signal Process. 43 (1995) 978–987.
[21] P. Gounon, C. Adnet, J. Galy, Localization angulaire de signaux non circulaires, Traitement du Signal 15 (3) (1998) 17–23.
[22] H. Abeida, J.P. Delmas, MUSIC-like estimation of direction of arrival for noncircular sources, IEEE Trans. Signal Process. 54 (7) (2006) 2678–2690.