DOA estimation of incoherently distributed sources using importance sampling maximum likelihood

WU Tao[1,2], DENG Zhenghong[1,*], HU Xiaoxiang, LI Ao, and XU Jiwei[3]

1. School of Automation, Northwestern Polytechnical University, Xi'an 710072, China;
2. Equipment Management and UAV College, Air Force Engineering University, Xi'an 710051, China;
3. School of Information Engineering, Xi'an University of Posts and Telecommunications, Xi'an 710061, China

Abstract: In this paper, an importance sampling maximum likelihood (ISML) estimator for direction-of-arrival (DOA) of incoherently distributed (ID) sources is proposed. Starting from the maximum likelihood estimation description of the uniform linear array (ULA), a decoupled concentrated likelihood function (CLF) is presented. A new objective function based on CLF which can obtain a closed-form solution of global maximum is constructed according to Pincus theorem. To obtain the optimal value of the objective function which is a complex high-dimensional integral, we propose an importance sampling approach based on Monte Carlo random calculation. Next, an importance function is derived, which can simplify the problem of generating random vector from a high-dimensional probability density function (PDF) to generate random variable from a one-dimensional PDF. Compared with the existing maximum likelihood (ML) algorithms for DOA estimation of ID sources, the proposed algorithm does not require initial estimates, and its performance is closer to Cramer-Rao lower bound (CRLB). The proposed algorithm performs better than the existing methods when the interval between sources to be estimated is small and in low signal to noise ratio (SNR) scenarios.

Keywords: direction-of-arrival (DOA) estimation, incoherently distributed (ID) sources, importance sampling maximum likelihood (ISML), Monte Carlo random calculation.

1. Introduction

In the field of array signal processing, such as radar, sonar, and communication, one of the most important practical needs is to estimate the direction-of-arrival (DOA) of targets. There are two ways to model targets, one being point source model, the other being the distributed source model. DOA estimation for point source targets has been relatively mature. At present, research directions mainly include estimations in special cases [1,2], applications of advanced array structure [3−5] and machine learning in estimation [6]. As the distance between a receiving array and a target is shortened, spatial expansion of the target is observed. The distributed source model is more suitable to describe spatial characteristics of the target. Distributed sources can be divided into coherently distributed (CD) sources and incoherently distributed (ID) sources according to whether different scatterers of a target are correlated [7]. By extending the classical methods of point sources, DOA estimation of CD sources has been well solved [8−10]. However, for ID sources, the energy of signal is distributed throughout the observation space, and the covariance matrix of the signal receiver vector is full rank. If traditional algorithms of point source model are used to estimate the parameters of ID sources, the performance will decline [7,11].

DOA estimation of ID sources is one of the research hotspots in recent years. So far, DOA estimators of ID sources can be divided into five categories. One is subspace search algorithms. For example, distributed signal parameter estimator (DSPE) [7] and dispersed signal parameter estimator (DISPARE) [11] algorithms are extended from the classic point source estimator multiple signal classification (MUSIC). These technologies need spectral peak search and effective measurement with respect to dimensions of pseudo signal subspace. The subspace rotation algorithms which are widely studied at present [12−20] show great advantages in low complexity estimation. These algorithms mostly use the first order Taylor series approximation of the array manifold vectors to construct the generalized array manifold (GAM) vectors and the rotation invariant relationship between GAM vectors to estimate parameters. Since such methods are based on the estimate signal parameters via rotational

invariance techniques (ESPRIT) framework, we call such methods ESPRIT-like algorithms. In such methods, [12−14] considered one dimensional ID source while [15,16] dealt with two-dimensional ID sources based on L-shaped array and circular array respectively. References [17,18] considered two-dimensional ID sources under multiple input multiple output (MIMO) systems background. Another class of low complexity algorithms is based on beamspace shift invariance structure via designing appropriate beamforming matrices such as [19,20] which have considered the noncircular ID sources. There are a class of nonparametric methods, such as beamforming and Capon spectral methods [21,22], which require high-dimensional searches and are ordinary in estimation accuracy. Covariance matching estimation techniques (COMET) methods [23−26] have also achieved good performance. However, [23,24] can only estimate a single source. In addition, initial estimates are required for iterative calculation in [25,26]. The maximum likelihood (ML) estimation has advantages in estimation accuracy, whereas the global optimal solution of the ML function needs multi-dimensional grid search, the computational complexity increases geometrically with the number of sources; instead, several iterative methods have been proposed for ML approach. In theory, iterative methods cannot guarantee the convergence to the global optimal solution, and it needs to provide initial iterative values. Trump et al. [27] presented the ML estimation of ID source for the first time, which requires complex nonlinear optimization and initial estimates. An approximate ML method for joint estimation of DOA and angular spread considering multiple sources was proposed in [28], which also needs to provide initial estimates. Ghogho et al. [29] proposed an approximate ML estimation method for a simplified ID source model. Based on the same source model in [29], two approximate ML methods were proposed in [30], one of which needs initial estimates, and the other has a mediocre performance in the case of low signal to noise ratio (SNR); however, both algorithms handle single-source estimation. As it has significant advantages in estimation accuracy, ML method is widely used in DOA estimation for point sources. Aiming at joint estimation of DOA and other parameters as well as DOA estimation under special array, [31−35] achieved good performance by ML methods. References [36−38] proposed to use importance sampling to conduct Monte Carlo estimation of parameters, such as DOA and Doppler frequency. The estimation accuracy of these methods is significantly improved, and it does not depend on the initial estimation; nevertheless, the above methods can only deal with point source.

This paper presents a non-iterative ML estimation for

DOA of ID sources. For ID sources, ML functions depend on DOAs and covariance matrices of receive vectors. In order to reduce parameters in process of likelihood function maximization, a concentrated likelihood function (CLF) is introduced to decouple the unnecessary parameters. To compute the optimal value of the CLF, we propose to apply the Pincus theorem [39] which can converge to the global optimal value but requires a multidimensional integral operation. Elaborating relationship between means of random vectors and the estimated values, the Monte Carlo random calculation technology [40,41] can be used to realize the multidimensional integral. The key solution then is designing an appropriate importance function as a probability density function (PDF) from which random vectors are realized. The ML method proposed in this paper is not limited to estimation for DOA of ID sources under uniform linear arrays (ULAs), and it can also be applied to different arrays by modifying the likelihood function. To show the contributions of this paper clearly, the main differences between the state-of-the-art methods and our work are listed as follows:

(i) Reference [12−20] used low complexity calculation methods, while this paper uses the ML method based on importance sampling, random calculations is utilized to find the optimal estimation.

(ii) Although [36−38] used the importance sampling method to calculate DOA, their research objects are all point sources, while this paper focuses on ID source.

(iii) Refences [27−28,30] used ML methods to estimate DOA of ID sources, they are all iterative methods and need to set initial values; the method proposed in this paper does not require initial values and can be applied to a wider range of array geometry types. The proposed method performs well in low SNR estimation and small snapshots size, especially for the signal source with relatively close distance.

This paper is organized as follows: In Section 2, we introduce the model of signal and array. In Section 3, the importance sampling maximum likelihood (ISML) estimator for DOA of ID sources is elaborated. Section 4 discusses the simulation results. Section 5 presents conclusions.

2. Signal and array model

As shown in Fig. 1, assume that the number of elements in a ULA is M and the spacing is d. There are K targets with nominal DOAs $\boldsymbol{\theta}=[\theta_1, \theta_2, \cdots, \theta_k]$ incident into the array, $\theta_k \in [-\pi/2, \pi/2]$. Different from the assumption of point source, ID source model assumes that the kth target is composed of L_k scatterers. Then the signal vector received by the array at time t can be written as

$$\boldsymbol{r}(t) = \sum_{k=1}^{K} s_k(t) \sum_{l=1}^{L_k} \gamma_{k,l}(t) \boldsymbol{a}(\theta_{k,l}(t)) + \boldsymbol{n}(t) \qquad (1)$$

where $s_k(t)$ is signal of the kth ID source, $\theta_{k,l}(t)$ is direction of the lth scatterer in kth ID source at time t, $\gamma_{k,l}(t)$ is the random complex gain of the lth scatterer, $\boldsymbol{n}(t)$ is the additive white Gaussian noise with variance σ^2, and $\boldsymbol{a}(\theta_{k,l}(t))$ denotes the array manifold of the ULA with respect to the point source. The ID source model assume that signal from different scatterers are uncorrelated, which means that the random complex gain of lth scatterer from kth ID sources and l'th scatterer from k'th ID sources [14,16,18] satisfies the following relationship:

$$E[\gamma_{k,l}(t)\gamma^*_{k',l'}(t')] = \frac{\sigma^2_{\gamma k}}{L_k}\delta(k-k')\delta(l-l')\delta(t-t') \qquad (2)$$

where $\delta(\cdot)$ is the Kronecker delta function, and $E[\cdot]$ is the expectation operation. $\sigma^2_{\gamma k}$ represents the total power gain of the kth source. The equation manifests that the power gain of a scatterer is $1/L_k$ of the total power gain of the kth source.
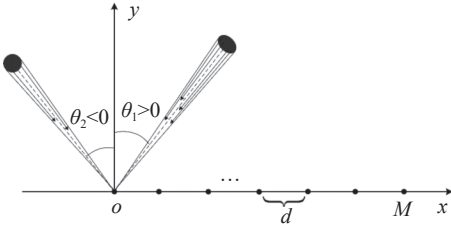

Fig. 1   Array structure of ULA with two ID sources

Let $\mu=2\pi d/\lambda$, where λ is wavelength of signal, then the mth element of $\boldsymbol{a}(\theta_{k,l}(t))$ can be expressed as follows:

$$[\boldsymbol{a}(\theta_{k,l}(t))]_m = \exp\{j\mu(m-1)\sin(\theta_{k,l}(t))\}. \qquad (3)$$

The direction of the lth scatterer within the kth ID source can be expressed as a sum of a deviation angle from nominal DOA of the kth source and nominal DOA, which can be expressed as follows:

$$\theta_{k,l}(t) = \theta_k(t) + \varphi_{k,l}(t) \qquad (4)$$

where $\theta_k(t)$ is the nominal DOA of kth ID source and $\varphi_{k,l}(t)$ is the deviation from nominal DOA with zero mean and variance σ^2_k. σ_k can be called angular spread of the kth source. Then we can obtain

$$\boldsymbol{a}(\theta_{k,l}(t)) \approx \boldsymbol{a}(\theta_k) + \boldsymbol{a}'(\theta_k)\varphi_{k,l}(t) \qquad (5)$$

where $\boldsymbol{a}'(\theta_k)$ is the partial derivative of $\boldsymbol{a}(\theta_k)$ at θ_k. According to the ID source model [7], scatterers of each source are assumed to be enormous, which may obey Gaussian, uniform, or other form distributions according to spatial geometry characteristics of targets.

According to (5), (1) can be expressed as follows:

$$\boldsymbol{r}(t) \approx \sum_{k=1}^{K} s_k(t) \sum_{l=1}^{L_k} \gamma_{k,l}(t)(\boldsymbol{a}(\theta_k) + \boldsymbol{a}'(\theta_k)\varphi_{k,l}(t)) + \boldsymbol{n}(t). \quad (6)$$

Consequently, the receiver vector of ULA can be written as follows:

$$\boldsymbol{r}(t) \approx \sum_{k=1}^{K} (\boldsymbol{a}(\theta_k)\upsilon_{k,0}(t) + \boldsymbol{a}'(\theta_k)\upsilon_{k,1}(t)) + \boldsymbol{n}(t) \qquad (7)$$

where

$$\upsilon_{k,0}(t) = s_k(t)\sum_{l=1}^{L_k}\gamma_{k,l}(t), \qquad (8)$$

$$\upsilon_{k,1}(t) = s_k(t)\sum_{l=1}^{L_k}\gamma_{k,l}(t)\varphi_{k,l}(t). \qquad (9)$$

Apparently

$$E\{\upsilon_{k,0}(t)\upsilon^*_{k,0}(t)\} =$$
$$E\{|s_k(t)|^2\}E\left\{\sum_{l=1}^{L_k}\gamma_{k,l}(t)\sum_{l=1}^{L_k}\gamma^*_{k,l}(t)\right\} = P_k \qquad (10)$$

where $(\cdot)^*$ means the conjugate of a complex variable, and $P_k = E\{|s_k(t)|^2\}\sigma^2_{\gamma k}$ is power of the kth ID source. Based on distribution characteristics of $\varphi_{k,l}(t)$ and signal propagation characteristics described by (2), we have

$$E\{\upsilon_{k,1}(t)\upsilon^*_{k,1}(t)\} = P_k\sigma^2_k. \qquad (11)$$

Equation (1) can be overwritten as a linear form approximately

$$\boldsymbol{r}(t) \approx \boldsymbol{B}(\theta)\boldsymbol{g}(t) + \boldsymbol{n}(t) \qquad (12)$$

where

$$\boldsymbol{B}(\theta) = [\boldsymbol{A}(\theta_1), \boldsymbol{A}(\theta_2), \cdots, \boldsymbol{A}(\theta_K)], \qquad (13)$$

$$\boldsymbol{A}(\theta_k) = [\boldsymbol{a}(\theta_k), \boldsymbol{a}'(\theta_k)], \qquad (14)$$

$$\boldsymbol{g}(t) = [\boldsymbol{g}_1^T, \boldsymbol{g}_2^T, \cdots, \boldsymbol{g}_K^T]^T, \qquad (15)$$

$$\boldsymbol{g}_k = [\upsilon_{k,0}(t), \upsilon_{k,1}(t)]^T, \qquad (16)$$

where $(\cdot)^T$ means the transpose of a matrix, $\boldsymbol{B}(\theta)$ is $M\times 2K$ dimensional and is regarded as the GAM of the ULA; $\boldsymbol{g}(t)$ is $2K\times 1$ dimensional and is generalized signal vector. The covariance of $\boldsymbol{g}(t)$ can be expressed as follows:

$$\boldsymbol{\Lambda} = E\{\boldsymbol{g}(t)\boldsymbol{g}^H(t)\} = \mathrm{diag}\{\boldsymbol{\Lambda}_1, \boldsymbol{\Lambda}_2, \cdots, \boldsymbol{\Lambda}_K\} \qquad (17)$$

where $(\cdot)^H$ means Hermitian transpose of a matrix, and $\boldsymbol{\Lambda}_k = P_k[1, \sigma^2_k]$.

Therefore, the covariance of the received signal $\boldsymbol{r}(t)$ given by (12) can be derived as follows:

$$\boldsymbol{R} \approx E\{\boldsymbol{r}(t)\boldsymbol{r}^H(t)\} = \boldsymbol{B}(\theta)\boldsymbol{\Lambda}\boldsymbol{B}^H(\theta) + \sigma^2\boldsymbol{I}_M. \qquad (18)$$

3. Proposed algorithm

3.1 ML estimation

The receiver vector of the ULA obeys a mean of $\mathbf{0}$ and a variance of $\boldsymbol{B}(\boldsymbol{\theta})\boldsymbol{\Lambda}\boldsymbol{B}(\boldsymbol{\theta})^{\mathrm{H}}+\sigma^2\boldsymbol{I}_M$ complex Gaussian distribution, which can be expressed as

$$\boldsymbol{r}(t) \sim \mathrm{CN}(\mathbf{0}, \boldsymbol{B}(\boldsymbol{\theta})\,\boldsymbol{\Lambda}\,\boldsymbol{B}(\boldsymbol{\theta})^{\mathrm{H}} + \sigma^2\boldsymbol{I}_M). \qquad (19)$$

Suppose that the ULA receives N snapshot data, define vector $\bar{\boldsymbol{r}}=[\boldsymbol{r}\,(1)^{\mathrm{T}}, \boldsymbol{r}\,(2)^{\mathrm{T}},\cdots, \boldsymbol{r}\,(N)^{\mathrm{T}}]^{\mathrm{T}}$. Then the joint PDF of N samples can be written as

$$f(\bar{\boldsymbol{r}};\, \boldsymbol{\theta},\boldsymbol{\Lambda}) = \frac{1}{\pi^{NM}|\boldsymbol{B}(\boldsymbol{\theta})\boldsymbol{\Lambda}\boldsymbol{B}(\boldsymbol{\theta})^{\mathrm{H}}+\sigma^2\boldsymbol{I}_{\mathrm{M}}|^N} \cdot$$
$$\exp\left[-\sum_{t=1}^{N}\boldsymbol{r}(t)^{\mathrm{H}}(\boldsymbol{B}(\boldsymbol{\theta})\boldsymbol{\Lambda}\boldsymbol{B}(\boldsymbol{\theta})^{\mathrm{H}}+\sigma^2\boldsymbol{I}_{\mathrm{M}})^{-1}\boldsymbol{r}(t)\right]. \qquad (20)$$

Then the logarithmic likelihood function of the observed data can be written as

$$L(\boldsymbol{\theta},\boldsymbol{\Lambda}) = -N\ln|\boldsymbol{B}(\boldsymbol{\theta})\boldsymbol{\Lambda}\boldsymbol{B}(\boldsymbol{\theta})^{\mathrm{H}} + \sigma^2\boldsymbol{I}_{\mathrm{M}}|-$$
$$\sum_{t=1}^{N}\boldsymbol{r}(t)^{\mathrm{H}}(\boldsymbol{B}(\boldsymbol{\theta})\boldsymbol{\Lambda}\boldsymbol{B}(\boldsymbol{\theta})^{\mathrm{H}} + \sigma^2\boldsymbol{I}_{\mathrm{M}})^{-1}\boldsymbol{r}(t). \qquad (21)$$

It can be seen that the unknown variables in above likelihood function include $\boldsymbol{\Lambda}$ and noise variance σ^2 in addition to $\boldsymbol{\theta}$. Reference [42] decouples the likelihood function from $\boldsymbol{\Lambda}$ and derives a CLF. The CLF with respect to $\boldsymbol{\theta}$ can be rewritten as

$$L_C(\boldsymbol{\theta}) = \frac{1}{N\sigma^2}\sum_{t=1}^{N}\boldsymbol{r}(t)^{\mathrm{H}}\boldsymbol{G}(\boldsymbol{\theta})\boldsymbol{r}(t)-$$
$$\ln|\boldsymbol{G}(\boldsymbol{\theta})\boldsymbol{S}\boldsymbol{G}(\boldsymbol{\theta}) + \sigma^2(\boldsymbol{I}_{\mathrm{M}} - \boldsymbol{G}(\boldsymbol{\theta}))| \qquad (22)$$

where

$$\boldsymbol{S} = \mathrm{E}[\boldsymbol{r}(t)\boldsymbol{r}(t)^{\mathrm{H}}],$$
$$\boldsymbol{G}(\boldsymbol{\theta}) = \boldsymbol{B}(\boldsymbol{\theta})\left[\boldsymbol{B}(\boldsymbol{\theta})^{\mathrm{H}}\boldsymbol{B}(\boldsymbol{\theta})\right]^{-1}\boldsymbol{B}(\boldsymbol{\theta})^{\mathrm{H}}.$$

Noise variance σ^2 can be estimated by eigendecomposition of $\mathbf{R}$, taking an average of the $2K+1$ to M eigenvalues of $\mathbf{R}$ [13].

3.2 Global maximization of the likelihood function

In order to obtain the ML estimate of $\boldsymbol{\theta}$, we need to maximize (22) with respect to $\boldsymbol{\theta}$. Actually, (22) is high-dimensional and nonlinear, the iterative optimization approach cannot guarantee the convergence to the global optimal solution theoretically. To obtain the optimal value, the direct realization requires multi-dimensional grid search, and the computational complexity increases exponentially with the number of sources. For multivariable functions with multiple local maxima, Pincus [39] proposed a closed expression of the global maximum.

According to Pincus theorem [39], it is assumed that $\boldsymbol{x}$ is an n-dimensional vector, and $f(\boldsymbol{x})$ is a continuous function on a bounded domain closure in n-dimensional Euclidean space, and the global maximum of $f(\boldsymbol{x})$ is obtained at $[x_1, x_2, \cdots, x_n]$.

Then, as the weight parameter ρ reaches infinity, the estimated value of the variable x_i can be obtained as

$$\hat{x}_i = \lim_{\rho\to\infty}\frac{\int\cdots\int x_i\exp(\rho f(\boldsymbol{x}))\mathrm{d}\boldsymbol{x}}{\int\cdots\int \exp(\rho f(\boldsymbol{x}))\mathrm{d}\boldsymbol{x}},\; i = 1,2,\cdots,n. \qquad (23)$$

The degree of integration in (23) is the dimension of vector $\boldsymbol{x}$. A normalized function can be defined as follows:

$$\bar{f}(\boldsymbol{x}) = \lim_{\rho\to\infty}\frac{\exp(\rho f(\boldsymbol{x}))}{\int\cdots\int \exp(\rho f(\boldsymbol{x}))\mathrm{d}\boldsymbol{x}}. \qquad (24)$$

The above equation is nonnegative, and it satisfies $\int\cdots\int \bar{f}(\boldsymbol{x})\mathrm{d}\boldsymbol{x} = 1$. Thus, $\bar{f}(\boldsymbol{x})$ can be considered as a pseudo PDF of $\boldsymbol{x}$. Thus, (24) can be rewritten as

$$\hat{x}_i = \lim_{\rho\to\infty}\int\cdots\int x_i\bar{f}(\boldsymbol{x})\mathrm{d}\boldsymbol{x},\; i = 1,2,\cdots,n. \qquad (25)$$

In this case, Pincus theorem can be regarded as finding the expectation of $\boldsymbol{x}$ which satisfies the PDF of $\bar{f}(\boldsymbol{x})$.

Therefore, we can obtain the estimated DOA of the kth source by Pincus theorem as follows:

$$\hat{\theta}_k = \lim_{\rho\to\infty}\int\cdots\int \theta_k\bar{L}(\boldsymbol{\theta})\mathrm{d}\boldsymbol{\theta},\quad k = 1,2,\cdots,K \qquad (26)$$

where

$$\bar{L}(\boldsymbol{\theta}) = \frac{\exp(\rho L_C(\boldsymbol{\theta}))}{\int\cdots\int \exp(\rho L_C(\boldsymbol{\theta}))\mathrm{d}\boldsymbol{\theta}}. \qquad (27)$$

Equation (26) can be considered as the solving expectation of the random vector $\boldsymbol{\theta}$ whose pseudo PDF is $\bar{L}(\boldsymbol{\theta})$. Unfortunately, $\bar{L}(\boldsymbol{\theta})$ is a nonlinear high-dimensional function. The Monte Carlo random calculation technology [40,41] can be applied effectively. Its basic principle can be considered as $\hat{\boldsymbol{\theta}}$ is the unbiased estimate of (26), which has the following expression:

$$\hat{\boldsymbol{\theta}} = \frac{1}{Q}\sum_{q=1}^{Q}\boldsymbol{\theta}_q \qquad (28)$$

where $\boldsymbol{\theta}_q$ is the qth sample generated by the pseudo PDF $\bar{L}(\boldsymbol{\theta})$. The sample mean converges to the optimal estimated value when Q goes to infinity. Thereafter, the key question is how to get samples of random vector $\boldsymbol{\theta}$ obeying the high-dimensional nonlinear distribution $\bar{L}(\boldsymbol{\theta})$.

3.3 Importance sampling method

The importance sampling method is a random calculation approach to realize a complex integral by a designed and easily implemented pseudo PDF.

Integral $\int h(\boldsymbol{x})\mathcal{F}(\boldsymbol{x})\mathrm{d}\boldsymbol{x}$ can be expressed as follows:

$$\int h(\boldsymbol{x})\mathcal{F}(\boldsymbol{x})\mathrm{d}\boldsymbol{x} = \int h(\boldsymbol{x})\frac{\mathcal{F}(\boldsymbol{x})}{p(\boldsymbol{x})}p(\boldsymbol{x})\mathrm{d}\boldsymbol{x} \qquad (29)$$

where $\mathcal{F}(\boldsymbol{x})$ and $p(\boldsymbol{x})$ both satisfy the requirement of PDF, that is, they are non-negative, and the integral in domain is equal to 1. Equation (29) can be regarded as expectation of $h(\boldsymbol{x})\mathcal{F}(\boldsymbol{x})/p(\boldsymbol{x})$ as the PDF of $\boldsymbol{x}$ is $p(\boldsymbol{x})$. $p(\boldsymbol{x})$ is named as importance function. When the number of samples is large enough, according to the Monte Carlo random calculation principle, the expected value can be obtained by sample mean, which can be written as follows:

$$\hat{\mathrm{E}}\left[h(\boldsymbol{x})\frac{\mathcal{F}(\boldsymbol{x})}{p(\boldsymbol{x})}\right] = \frac{1}{Q}\sum_{q=1}^{Q}h(\boldsymbol{x})\frac{\mathcal{F}(\boldsymbol{x})}{p(\boldsymbol{x})}. \qquad (30)$$

Now we consider designing the importance function $p(\boldsymbol{x})$. On one hand, the ideal $p(\boldsymbol{x})$ should be as similar to $\mathcal{F}(\boldsymbol{x})$ in shape as possible to reduce the variance of estimation by (30); on the other hand, the structure of $p(\boldsymbol{x})$ should be simple to facilitate the generation of random vector $\boldsymbol{x}$. Therefore, there is a trade-off between the two aspects when designing $p(\boldsymbol{x})$. We consider the first term of CLF shown in (22) as follows:

$$\boldsymbol{L}'_C(\boldsymbol{\theta}) = \frac{1}{N\sigma^2}\sum_{t=1}^{N}\boldsymbol{r}(t)^{\mathrm{H}}\boldsymbol{B}(\boldsymbol{\theta})\boldsymbol{\Phi}^{-1}\boldsymbol{B}(\boldsymbol{\theta})^{\mathrm{H}}\boldsymbol{r}(t). \qquad (31)$$

In (31), entries in $\boldsymbol{\Phi}=[\boldsymbol{B}(\boldsymbol{\theta})^{\mathrm{H}}\ \boldsymbol{B}(\boldsymbol{\theta})]$ can be written as follows:

$$\begin{cases} \boldsymbol{\Phi}_{2k-1,2k-1} = M \\ \boldsymbol{\Phi}_{2k-1,2k} = A\mathrm{j}\mu\cos\theta_k \\ \boldsymbol{\Phi}_{2k,2k} = B\mu^2\cos^2\theta_k \\ \boldsymbol{\Phi}_{2k,2k-1} = -A\mathrm{j}\mu\cos\theta_k \end{cases} \qquad (32)$$

where $A=M(M-1)/2$, $B=-(M-1)M(2M-1)/6$. If $k\neq k'$, the following relationships can be proven:

$$\begin{cases} |\boldsymbol{\Phi}_{2k-1,2k'-1}| \leqslant M \\ |\boldsymbol{\Phi}_{2k-1,2k'}| \leqslant |A\mathrm{j}\mu\cos\theta_k| \\ |\boldsymbol{\Phi}_{2k,2k'}| \leqslant |B\mu^2\cos^2\theta_k| \\ |\boldsymbol{\Phi}_{2k,2k'-1}| \leqslant |A\mathrm{j}\mu\cos\theta_k| \end{cases} \qquad (33)$$

Then $\boldsymbol{\Phi}$ can be approximately equal to the following expression:

$$\boldsymbol{\Phi} \approx \mathrm{diag}[\boldsymbol{\Lambda}_1\ \boldsymbol{\Lambda}_2\ \cdots\ \boldsymbol{\Lambda}_K] \qquad (34)$$

where

$$\boldsymbol{\Lambda}_k = \begin{bmatrix} M & v_{k1} \\ v_{k1}^* & v_{k2} \end{bmatrix} = \begin{bmatrix} M & A\mathrm{j}\mu\cos\theta_k \\ -A\mathrm{j}\mu\cos\theta_k & B\mu^2\cos^2\theta_k \end{bmatrix}. \qquad (35)$$

Therefore, if normalization is not considered, the importance function is selected as

$$\widetilde{\boldsymbol{L}}'_C(\boldsymbol{\theta}) = \exp\left(\rho_1\frac{1}{N\sigma^2}\sum_{t=1}^{N}\boldsymbol{r}(t)^{\mathrm{H}}\boldsymbol{B}(\boldsymbol{\theta})\boldsymbol{\Phi}^{-1}\boldsymbol{B}(\boldsymbol{\theta})^{\mathrm{H}}\boldsymbol{r}(t)\right) \qquad (36)$$

where ρ_1 is another weight parameter.

According to (34),

$$\sum_{t=1}^{N}\boldsymbol{r}(t)^{\mathrm{H}}\boldsymbol{B}(\boldsymbol{\theta})\boldsymbol{\Phi}^{-1}\boldsymbol{B}(\boldsymbol{\theta})^{\mathrm{H}}\boldsymbol{r}(t) \approx$$

$$\sum_{t=1}^{N}\sum_{k=1}^{K}\left\{|F_1(\theta_k,t,m)|^2a_{k1} + |F_2(\theta_k,t,m)|^2a_{k3}+\right.$$

$$\left. F_1(\theta_k,t,m)F_2^*(\theta_k,t,m)a_{k2}^* + F_2(\theta_k,t,m)F_1^*(\theta_k,t,m)a_{k2}\right\} \qquad (37)$$

where

$$\begin{cases} a_{k1} = v_{k2}^2\xi_k \\ a_{k2} = v_{k1}\xi_k \\ a_{k3} = M\xi_k \\ \xi_k = (Mv_{k2} - v_{k1}v_{k1}^*)^{-1} \end{cases} \qquad (38)$$

$F_1(\theta_k,t,m)$ and $F_2(\theta_k,t,m)$ can be expressed as

$$F_1(\theta_k,t,m) = \sum_{m=1}^{M}r_m^*(t)\exp(\mathrm{j}m\mu\sin\theta_k), \qquad (39)$$

$$F_2(\theta_k,t,m) = \sum_{m=1}^{M}m\mathrm{j}\mu\cos\theta_k r_m^*(t)\exp(\mathrm{j}m\mu\sin\theta_k). \qquad (40)$$

Then the importance function (36) can be written as

$$\widetilde{\boldsymbol{L}}'_C(\boldsymbol{\theta}) = \prod_{k=1}^{K}\exp(\rho_1 I(\theta_k))$$

where

$$I(\theta_k) = \frac{1}{N\sigma^2}\sum_{t=1}^{N}\left\{|F_1(\theta_k,t,m)|^2a_{k1}+\right.$$

$$|F_2(\theta_k,t,m)|^2a_{k3} + F_1(\theta_k,t,m)F_2^*(\theta_k,t,m)a_{k2}+$$

$$\left. F_2(\theta_k,t,m)F_1^*(\theta_k,t,m)a_{k2}^*\right\}. \qquad (41)$$

Considering its normalization, the importance function can be written as

$$p(\boldsymbol{\theta}) = \prod_{k=1}^{K}\frac{\exp(\rho_1 I(\theta_k))}{\int\exp(\rho_1 I(\theta'_k))\mathrm{d}\theta'_k}. \qquad (42)$$

Thus, each θ_k can be considered independent and has the same PDF as follows:

$$p(\boldsymbol{\theta}) = \frac{\exp(\rho_1 I(\theta))}{\int\exp(\rho_1 I(\theta'))\mathrm{d}\theta'}. \qquad (43)$$

It can be seen that the importance sampling method can transform the problem of generating random vectors from a complex high-dimensional PDF into generating random variables from a number of one-dimensional PDFs

in the context of calculating the estimated value through a large sample of random numbers.

3.4 DOA estimation

Circular mean can effectively reduce bias and computational complexity [43]. Since θ_k is a cyclic random variable, we introduce the circular mean to calculate the estimated DOA. The circular mean of a random variable α with PDF expressed by (44) and period $[-\pi,\pi]$ is defined as

$$\mathrm{E}(\alpha) = \angle \int_{-\pi}^{\pi} \exp(\mathrm{j}\alpha) p(\alpha) \mathrm{d}\alpha \qquad (44)$$

where $\angle$ represents the angle of a complex variable. Define variable $\omega_k = \sin\theta_k$. The range of the variable ω_k is $[-1,1]$. Then the circular mean form of $h(\boldsymbol{x})$ in (29) can be expressed as follows:

$$h(\boldsymbol{\omega}) = \exp(\mathrm{j}\pi\boldsymbol{\omega}) \qquad (45)$$

where $\boldsymbol{\omega}=[\omega_1,\omega_2,\cdots,\omega_K]^{\mathrm{T}}$. Variable substitution for (27) and (42) are made as follows:

$$F(\boldsymbol{\omega}) = \frac{\exp(\rho L_c(\boldsymbol{\omega}))}{\int \exp(\rho L_c(\boldsymbol{\omega}')) \mathrm{d}\boldsymbol{\omega}'} \qquad (46)$$

$$p(\boldsymbol{\omega}) = \prod_{k=1}^{K} \frac{\exp(\rho_1 I(\omega_k))}{\int \exp(\rho_1 I(\omega'_k)) \mathrm{d}\omega'_k}. \qquad (47)$$

In the same way, each ω_k can be considered independent and has the same PDF as follows:

$$p(\omega) = \frac{\exp(\rho_1 I(\omega))}{\int \exp(\rho_1 I(\omega')) \mathrm{d}\omega'}. \qquad (48)$$

From the definition of the circular mean, according to importance sampling approach described by (29), we get the estimated ω_k as follows:

$$\hat{\omega}_k = \frac{1}{\pi} \angle \frac{1}{Q} \sum_{q=1}^{Q} \exp(\mathrm{j}\pi[\boldsymbol{\omega}_q]_k) \frac{\mathcal{F}(\boldsymbol{\omega}_q)}{p(\boldsymbol{\omega}_q)}. \qquad (49)$$

$\boldsymbol{\omega}_q$ is the qth random vector realized by $p(\boldsymbol{\omega})$. $[\boldsymbol{\omega}_q]_k$ is the kth variable in $\boldsymbol{\omega}_q$. Since the circular mean is the mean of angles of complex variable, the estimate of ω_k can be equivalent to the following expression:

$$\hat{\omega}_k = \frac{1}{\pi} \angle \frac{1}{Q} \sum_{q=1}^{Q} \exp(\mathrm{j}\pi[\boldsymbol{\omega}_q]_k) y(\boldsymbol{\omega}_q) \qquad (50)$$

$$y(\boldsymbol{\omega}) = \frac{\exp(\rho L_c(\boldsymbol{\omega}))}{\prod_{k=1}^{K} \exp(\rho_1 I(\omega_k))}. \qquad (51)$$

$y(\boldsymbol{\omega})$ is given by dividing (46) and (47) but omitting the constants $\int \exp(\rho L_c(\boldsymbol{\omega}')) \mathrm{d}\boldsymbol{\omega}'$ and $\prod_{k=1}^{K} \int \exp(\rho_1 I(\omega'_k)) \mathrm{d}\omega'_k$ in their denominators. This is because a complex angle after summation operation of complex variables does not

change when the magnitude of each complex variable cancels out a constant term together. In this way, using the circular mean can reduce the complexity of the estimation calculation. It is worth noting that both the numerator and denominator in (51) are of an exponential form, which is prone to overflow during calculation. Hence, considering

$$y(\boldsymbol{\omega}) = \exp(\rho L_c(\boldsymbol{\omega}) - \rho_1 \sum_{k=1}^{K} I(\omega_k)), \qquad (52)$$

we use $y'(\boldsymbol{\omega}_q)$ as a substitute, which can be written as follows:

$$y'(\boldsymbol{\omega}_q) = \exp\left[\rho L_c(\boldsymbol{\omega}_q) - \rho_1 \sum_{k=1}^{K} I([\boldsymbol{\omega}_q]_k) - \max_{1 \leqslant q \leqslant Q}(\rho L_c(\boldsymbol{\omega}_q) - \rho_1 \sum_{k=1}^{K} I[\boldsymbol{\omega}_q]_k)\right]. \qquad (53)$$

$\hat{\omega}_k$ denotes the estimated value of ω_k, which can be obtained as follows:

$$\hat{\omega}_k = \frac{1}{\pi} \angle \frac{1}{Q} \sum_{q=1}^{Q} \exp(\mathrm{j}\pi[\boldsymbol{\omega}_q]_k) y'(\boldsymbol{\omega}_q). \qquad (54)$$

Due to the nature of the circular mean, $y'(\boldsymbol{\omega}_q)$ only changes the amplitude of a complex variable and does not affect the complex angle.

3.5 Computational procedure and complexity analysis

The ML algorithm proposed in this paper can be summarized as follows:

Step 1 Calculate $F_\omega(\alpha)$, which is the approximate probability distribution function of $p(\omega)$. Take a sufficient number of discrete points to approximate the probability distribution function. Set 2 000 discrete points. $\omega_z = -1 + z/1\,000$ ($z=1, 2, \cdots, 2\,000$). Then the probability distribution function of $p(\omega)$ can be expressed as follows:

$$F_\omega(\alpha) = \int_{-1}^{\alpha} p(\omega) \mathrm{d}\omega \approx \frac{1}{\sum_{z=1}^{2\,000} \exp(\rho_1 I(\omega_z))} \sum_{z=1}^{-1+z/1\,000} \exp(\rho_1 I(\omega_z)). \qquad (55)$$

Step 2 Generate random vectors $\boldsymbol{\omega}$ from $p(\boldsymbol{\omega})$.
Substep 2.1 With the previous analysis, realize a single vector $\boldsymbol{\omega}$ from PDF $p(\boldsymbol{\omega})$ which is equivalent to generate K random variables from PDF $p(\boldsymbol{\omega})$. To do this, according to the inverse probability integration method in [41], first operation is generating random number u_k ($k=1,2,\cdots,K$) from $[0,1]$ uniform distribution, then calculate

$$\omega_k = F_\omega^{-1}(u_k) \qquad (56)$$

Substep 2.2　Repeat Substep 2.1 Q times to generate random vectors $\omega_q (q=1,2,\cdots,Q)$.

Step 3　Estimate the value of DOA.

Substep 3.1　Calculate an estimate of the circular mean $\hat{\omega}_k$ $(k=1,2,\cdots,K)$ from (54).

Substep 3.2　Estimated DOA of the kth ID source as θ_k =arcsin ω_k. The flow of the algorithm is shown in Fig. 2.
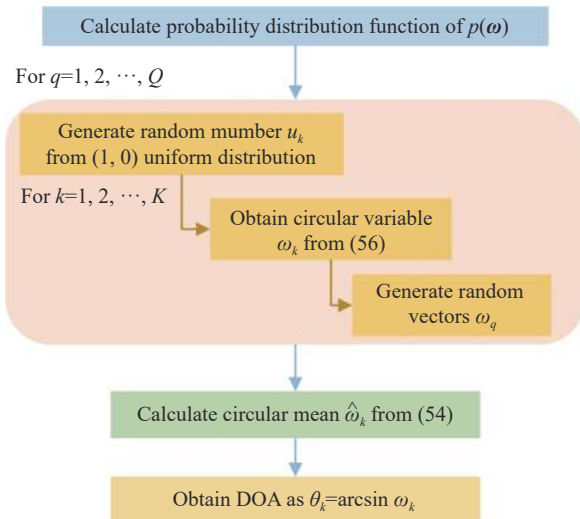

Fig. 2　Flow of the proposed algorithm

We analyze the computational complexity of the method proposed in comparison with representative ML methods [27,28] and a representative ESPRIT-like algorithm [14]. Reference [27] solves the optimal value of likelihood function through Newton search, which is a high-dimensional optimization process. Complexity of the algorithm are as follows: complexity of calculating the covariance matrix is $O(MN^2)$ and complexity of the optimization procedure is $O(80L_NKM^3)$ where L_N is number of iterations of Newton search. Reference [28] separates the source power and noise power from the likelihood function, and seeks the optimal value of likelihood function through a $2K$ dimensional search. The main calculation processes include calculating the covariance matrix, separating the power and noise power from the likelihood function and searching operation. Complexity of the three parts is $O(MN^2)$, $O(M^6+3KM^3)$, and $O(2KL_\theta L_\sigma M^3)$ respectively where L_θ and L_σ are the search points of DOA and angular spread. Reference [14] is a low complexity estimation method using the rotational invariant relationship of signal subspaces. The main calculation processes consist of calculating the covariance matrix, obtaining the signal subspace from the covariance matrix, conducting the polynomial root operation, and solving DOA from a linear equation. Complexity of the four parts is $O(MN^2)$,

$O(M^3)$, $O(8K^3M^3+8K^3)$ and $O(4MK^2+2KM^2)$ respectively. The computational cost of the proposed method mostly lies in calculating the approximate probability distribution function $F_\omega(\alpha)$, generating a large number of random vectors ω and calculating DOA. Complexity of the three parts is $O(6L_zNM)$ where L_z is number of discrete points of $F_\omega(\alpha)$, $O(QK)$ and $O(KM^2+6KM)$ respectively. In light of above analysis, compared with [14], the computational complexity of ML algorithms is relatively high on the whole. The computational complexity of [27] depends on the number of iterations L_N, and that of [28] depends on the grid density of $2K$ dimensional search L_θ and L_σ. Nevertheless, both methods need to provide initial estimates, consequently, the scale of iterations L_N and grid density L_θ and L_σ depend on the quality of the initial value. If the initial value is estimated well, the computational complexity is low, on the contrary, the computational complexity is high. The method proposed in this paper mainly depends on the number Q of random vectors and the fineness of the constructed distribution function $F_\omega(\alpha)$.

4. Numerical simulation

In this part, the performance of the proposed method is investigated by numerical simulations. Firstly, the influence of parameters of the ISML algorithm proposed in this paper is discussed. Secondly, two ML algorithms in [27,28], an ESPRIT-like algorithm [14] are compared with the proposed algorithm. The array structure of ULA is with nine elements, although the method in this paper is not limited to ULA. The SNR is defined as $10\lg(1/\sigma^2)$. In order to evaluate the estimation effect of different algorithms, we compare the root mean square error (RMSE) of the estimated results with the Cramer-Rao lower bound (CRLB) value in [27].

For the proposed method, ρ and ρ_1 need to be manually selected. In the following experiment, the influence of the two parameters will be examined: Two ID sources to be estimated are set as $-4°$ and $9°$, angular spreads are all $2°$, $Q=2\,000$, SNR= 5 dB, the number of snapshots is set as 200. RMSE takes average of the two sources by 100 Monte Carlo simulations. The estimation results under different parameters are shown in Fig. 3. As can be seen from Fig. 3(a), when ρ exceeds 8, the estimated result does not change significantly. However, as can be seen from Fig. 3(b), with the increase of ρ_1, the RMSE changes slowly at the beginning, then increases, and tends to remain unchanged when it achieves a certain extent. As parameter ρ tends to infinity, (27) takes the shape of a K-dimensional Dirac-delta function at the global optimal value point. There is a significant difference between other extreme points and global extreme

points. In fact, when the value of ρ exceeds a certain range, the global maximum points can already be separated from other extreme points. For ρ_1, the requirement is satisfied as long as some similarity between the importance function and the CLF is ensured. Since the importance function is a simplification of CLF, its extreme value point is different from that of the CLF. If ρ_1 is too large, the random samples will be generated around some points which deviate from the real extreme value, consequently, there will be errors in estimation.
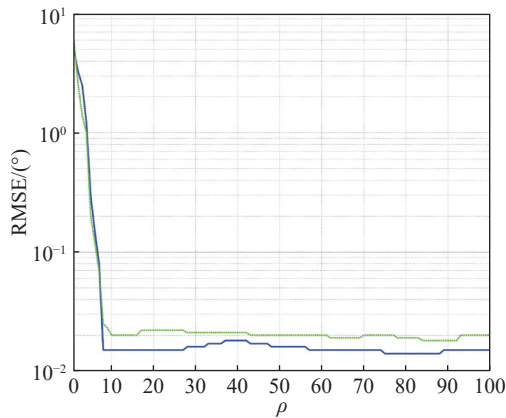
(a) RMSE versus parameter ρ with ρ_1=0.7
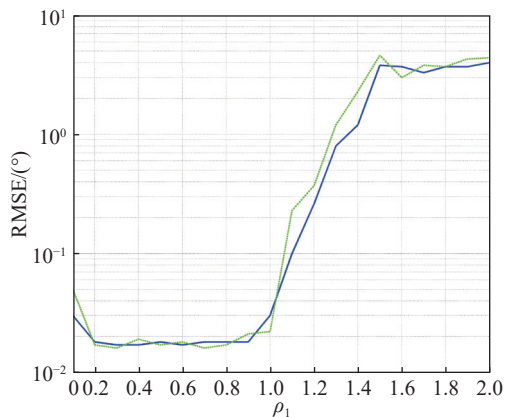
(b) RMSE versus parameter ρ_1 with ρ=10

——— : Source 1;   ········· : Source 2.

Fig. 3   RMSE versus parameter ρ and ρ_1

Theoretically, the larger Q is, the closer it is to the true value, but more computation is required. In this experiment, the parameters of sources to be estimated are the same as the first, SNR=5 dB, the number of snapshots is equal to 200. Fig. 4 shows the RMSE estimated by 100 Monte Carlo simulations. When Q is insufficient, the estimation is invalid, whereas the error will decrease with Q increasing. In practice, when the difference between several successive estimates is less than a given value, generating of random vectors will stop. When Q is insufficient, there is a risk that not all sources to be estimated will be identified. Therefore, the number of unknown

sources should be determined first, and then the decision conditions can be used to determine whether the random vector continues to be generated.
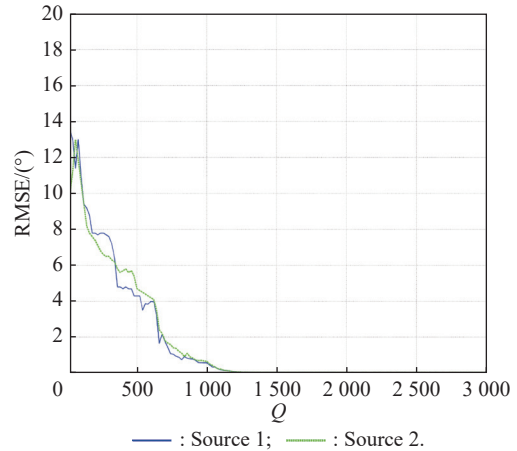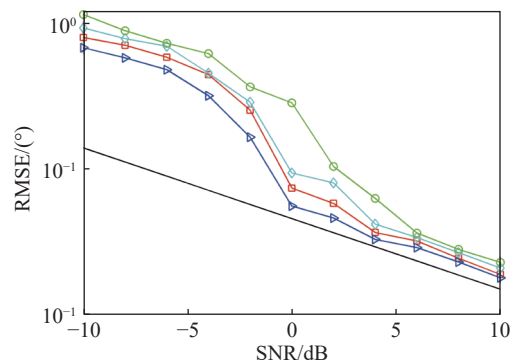

——— : Source 1;   ········· : Source 2.

Fig. 4   RMSE versus parameter Q with ρ=10 and ρ_1=0.7

We discuss the performance of different estimators varies with SNR and number of snapshots when the sources are all distinguishable. The DOA of the two sources to be estimated are −8° and 9°. Angular spreads are all set as 2°. For ULA with nine array elements, the 3 dB beam width is approximately 12.8°, so the two sources are considered to be well apart. We choose ρ=10, ρ_1=0.7, and we use 100 Monte Carlo simulations to calculate RMSE. ML algorithms [27,28] take the initial estimates as −11° and 13°. Fig. 5 shows the estimated results when the number of snapshots is 100, Q is selected as 2 000, the SNR varies from −10 dB to 10 dB. Fig. 5 also shows that estimation results of the proposed method are closer to CRLB compared with other methods. Especially, the method presented in this paper shows superiority under the condition of low SNR. Fig. 6 shows the RMSE estimated with SNR of 0 dB, where experiment conditions remains the same as before and snapshots changes from 20 to 200. The results indicate that the proposed method shows advantages in the case of small snapshots.
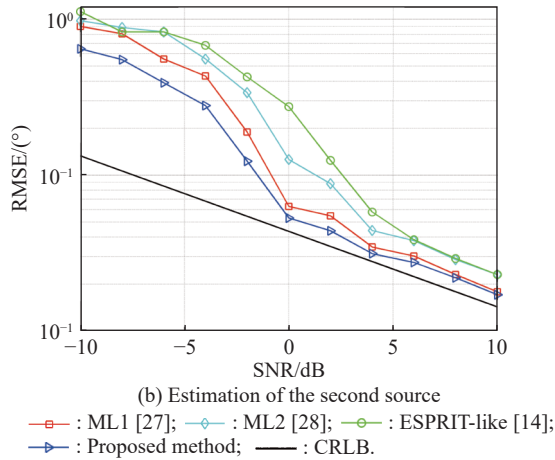

(a) Estimation of the first source

(b) Estimation of the second source

—□— : ML1 [27];   —◇— : ML2 [28];   —○— : ESPRIT-like [14];
—▷— : Proposed method;   —— : CRLB.

Fig. 5   RMSE versus SNR with number of snapshots equal to 100

(a) Estimation of the first source

(b) Estimation of the second source

—□— : ML1 [27];   —◇— : ML2 [28];   —○— : ESPRIT-like [14];
—▷— : Proposed method;   —— : CRLB.
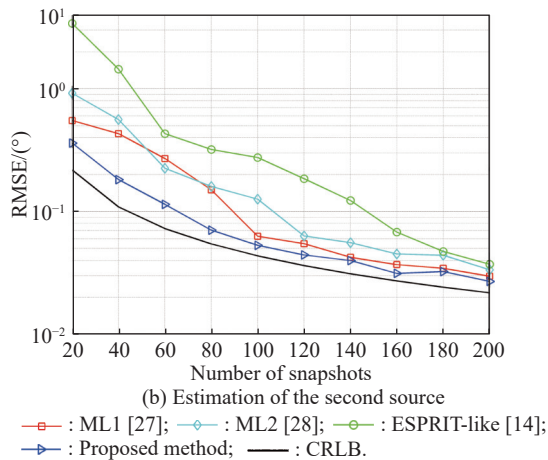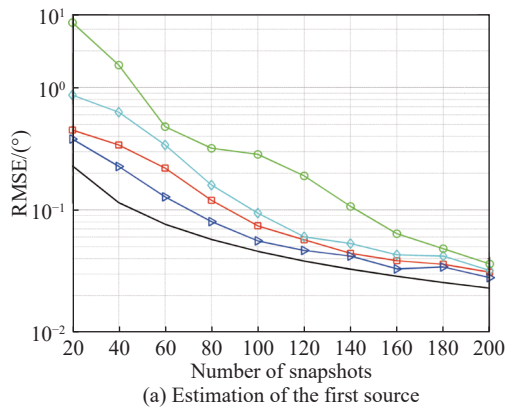
Fig. 6   RMSE versus number of snapshots with SNR equal to 0 dB

In next experiment, we investigate the computational complexity of different algorithms simultaneously. All the codes are written in Matlab, and the processor is the 11th Gen Inter(R) Core (TM) i7-1165G7. Table 1 shows the average estimated time for the different algorithms. It can be seen that ESPRIT-like algorithm [14] has higher computational efficiency than all other algorithms. Among the ML estimation methods, the calculation time of the proposed method is longer than algorithms [27,28].

Therefore, the application scenarios of calculation need to be considered with respect to these estimators for ID sources. If the requirement for accuracy is high, the method proposed in this paper has advantages; if the requirement for real-time computing is relatively high, the algorithm proposed needs to be further studied to improve the computational efficiency. It should be noted that in estimation implementation, the algorithms [27,28] needs initial estimates, calculation of initial estimates also requires the calculation time; however, the process is not included in the statistics in this experiment.

Table 1   Estimation time of different algorithms

Method	Average computing time	Relative speed
ESPRIT-like [14]	0.010 1	1.00
ML1 [27]	0.143 0	14.16
ML2 [28]	0.097 0	9.60
Proposed method	0.155 0	15.35

Since the distances of sources to be estimated have a great influence on estimation, we investigate the performance of estimators when the distance between sources is different. The DOA of the first source θ_1 is $-8°$. The DOA of the second source is $(\theta_1+\Delta)$, where Δ is the interval of the two sources. Angular spreads are all 2°. SNR=5 dB, snapshots number=100, Q=4 000, ρ=10, and ρ_1=0.7. The initial estimate of first source of ML algorithms [27,28] is $-11°$ and the second is 4° apart from the true value. Fig. 6 shows RMSE with angles interval varying from 1° to 20°. As can be seen from Fig. 7, the different estimators perform well beyond 9°. However, as the interval falls within 8°, the ESPRIT-like method [14] begins to deteriorate. Comparatively, ML algorithms [27,28] are robust to the variation of intervals between sources. It can be seen that the algorithm proposed in this paper can distinguish two sources even when the interval between two sources is 5°. Therefore, we can conclude that the method proposed has significant advantages compared with other methods in the context of near distance between two ID sources.
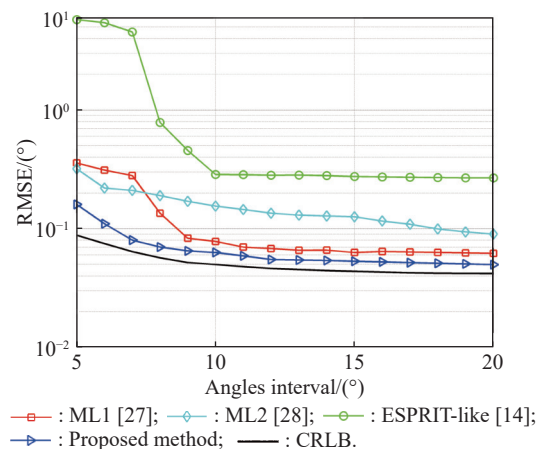

—□— : ML1 [27];   —◇— : ML2 [28];   —○— : ESPRIT-like [14];
—▷— : Proposed method;   —— : CRLB.

Fig. 7   RMSE versus angles interval

In above experiments, the ML methods [27,28] require initial estimates. In order to investigate the effect of initial estimates on estimation, we discuss the relationship between initial estimates and RMSE. To get the best performance, we set SNR as 5 dB and the number of snapshots as 200. The DOAs of the two sources are −8° and 9°. Angular spreads are all 2°. Assume that the initial estimate for the first source is −11°. Fig. 8 shows the RMSE curve as the initial estimate of the second source changes from 1° to 20°. It can be seen from Fig. 8 that for algorithm in [27,28], if the initial estimate is significantly different from the real value, the estimated performance will be poor, whereas the algorithm in [28] is slightly better than the algorithm in [27]. The ESPRIT-like algorithm does not need the initial estimation. Among ML algorithms, only the proposed algorithm can guarantee the robustness without relying on the initial estimation.
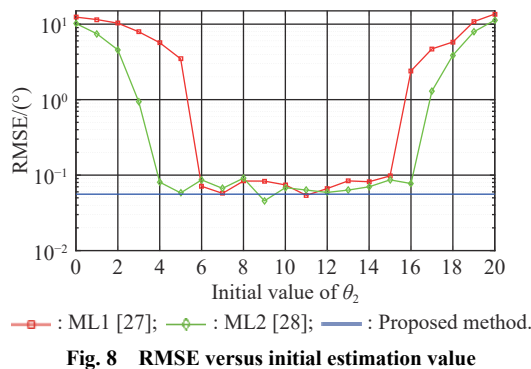

: ML1 [27];   : ML2 [28];   : Proposed method.

Fig. 8   RMSE versus initial estimation value

5. Conclusions

In this paper, an importance sampling ML method for DOA estimation of ID sources has been developed, an objective function is constructed from the decoupled CLF which can achieve global optimal points according to the Pincus theorem. In the light of importance sampling, a selected important function is designed to realize random vectors in Monte Carlo random calculation while the circular mean concept is used to reduce the complexity. The method proposed does not require setting initial values and can be applied to a wider range of array geometry types. The proposed method performs well in low SNR estimation and small snapshots size, especially for the sources with relatively close distance.

References

[1] CHEN P, CHEN Z M, CAO Z X, et al. A new atomic norm for DOA estimation with gain-phase errors. IEEE Trans. on Signal Processing, 2020, 68(7): 4293–4306.
[2] ZHENG Z, FU M C, WANG W Q, et al. Localization of mixed near-field and far-field sources using symmetric double-nested arrays. IEEE Trans. on Antennas and Propagation, 2019, 67(11): 7059–7070.
[3] WANG S X, ZHAO Y, LAILA I, et al. Joint 2D DOA and Doppler frequency estimation for L-shaped array using compressive sensing. Journal of Systems Engineering and Elec-

tronics, 2020, 31(1): 32–40.
[4] SEDIGHI S, RAO B S, OTTERSTEN B. An asymptotically efficient weighted least squares estimator for co-array-based DOA estimation. IEEE Trans. on Signal Processing, 2020, 68(5): 589–604.
[5] MA H H, TAO H H. Joint 2D-DOA and polarization estimation for sparse nonuniform rectangular array composed of spatially spread electromagnetic vector sensor. Journal of Systems Engineering and Electronics, 2020, 31(6): 1116–1127.
[6] PAPAGEORGIOU G K, SELLATHURAI M, ELDAR Y C. Deep networks for direction-of-arrival estimation in low SNR. IEEE Trans. on Signal Processing, 2021, 69(1): 3714–3729.
[7] VALAEE S, CHAMPAGNE B, KABAL P. Parametric localization of distributed sources. IEEE Trans. on Signal Processing, 1995, 43(9): 2144–2153.
[8] YANG X M, ZHENG Z, KO C C, et al. Low-complexity 2D parameter estimation of coherently distributed noncircular signals using modified propagator. Multidimensional Systems and Signal Processing, 2017, 28(2): 407–426.
[9] ZHENG Z, LI G J, TENG Y L. 2D DOA estimator for multiple coherently distributed sources using modified propagator. Circuits Systems and Signal Processing, 2010, 31(1): 255–270.
[10] XIONG W M, JOSE P, SYLVIE M. Performance analysis of distributed source parameter estimator (DSPE) in the presence of modeling errors due to the spatial distributions of sources. Signal Processing, 2018, 143(2): 146–151.
[11] MENG Y, STOICA P, WONG K M. Estimation of direction-of-arrival of spatially dispersed signals in array processing. IEE Proceedings-Radar Sonar and Navigation, 1996, 143(1): 1–9.
[12] ZHUANG J, XIONG H, WANG W, et al. Application of manifold separation to parametric localization for incoherently distributed sources. IEEE Trans. on Signal Processing. 2018, 66(11): 2849–2860.
[13] SHAHBAZPANAHI S, VALAEE S, BASTANI M H. Distributed source localization using esprit algorithm. IEEE Trans. on Signal Processing, 2001, 49(10): 2169–2178.
[14] CAO R Z, GAO F F, ZHANG X F. An angular parameter estimation method for incoherently distributed sources via generalized shift invariance. IEEE Trans. on Signal Processing, 2016, 64(17): 4493–4503.
[15] CHEN H, LIU Y H, WANG Q, et al. Two-dimensional angular parameter estimation for noncircular incoherently distributed sources based on an L-shaped array. IEEE Sensors Journal, 2020, 20(22): 13704–13715.
[16] GU Q Y, WANG H G, SUN W T, et al. A sequential ESPRIT algorithm based on a novel UCSA configuration for parametric estimation of two-dimensional incoherently distributed source. IEEE Trans. on Vehicular Technology, 2021, 70(1): 356–370.
[17] HU A Z, LYU T J, GAO H, et al. An ESPRIT-based approach for 2-D localization of incoherently distributed sources in massive MIMO systems. IEEE Journal of Selected Topics in Signal Processing, 2014, 8(5): 996–1011.
[18] ZHENG Z, WANG W Q, MENG H P, et al. Efficient beamspace-based algorithm for two-dimensional DOA estimation of incoherently distributed sources in massive MIMO systems. IEEE Trans. on Vehicular Technology, 2018, 67(12): 11776–11789.
[19] LIU Y H, CHEN H, WANG Q, et al. Noncircularity-based generalized shift invariance for estimation of angular parameters of incoherently distributed sources. Signal Processing, 2021, 183(5): 107989.
[20] ZHENG Z, LU J, WANG W Q, et al. An efficient method for angular parameter estimation of incoherently distributed sources via beamspace shift invariance. Digital Signal Processing, 2018, 83: 261–270.
[21] HASSANIEN A, SHAHBAZPANAHI S, GERSHMAN A B. A generalized capon estimator for localization of multiple

spread sources. IEEE Trans. on Signal Processing, 2004, 52(1): 280–283.

[22] ZOUBIR A, WANG Y. Robust generalised Capon algorithm for estimating the angular parameters of multiple incoherently distributed sources. IET Signal Processing, 2008, 2(2): 163–168.

[23] BESSON O, STOICA P. Decoupled estimation of DOA and angular spread for a spatially distributed source. IEEE Trans. on Signal Processing, 2000, 48(7): 1872–1882.

[24] ZOUBIR A, WANG Y, CHARGE P. On the ambiguity of COMET-EXIP algorithm for estimating a scattered source. Proc. of the International Conference on Acoustics, Speech, and Signal Processing, 2005: 941–944.

[25] SHAHBAZPANAHI S, VALAEE S, GERSHMAN A B. A covariance fitting approach to parametric localization of multiple incoherently distributed sources. IEEE Trans. on Signal Processing, 2004, 52(3): 592–600.

[26] SHAHBAZPANAHI S, VALAEE S, GERSHMAN A B. Parametric localization of multiple incoherently distributed sources using covariance fitting. Proc. of the IEEE Workshop Sensor Array Multichannel Signal Processing, 2002: 332–336.

[27] TRUMP T, OTTERSTEN B. Estimation of nominal direction-of-arrival and angular spread using an array of sensors. Signal Processing, 1996, 50(1): 57–69.

[28] SIESKUL B. An asymptotic maximum likelihood for joint estimation of nominal angles and angular spreads of multiple spatially distributed sources. IEEE Trans. on Vehicular Technology, 2010, 59(3): 1534–1538.

[29] GHOGHO M, BESSON O, SWAMI A. Estimation of directions of arrival of multiple scattered sources. IEEE Trans. on Signal Processing, 2001, 49(11): 2467–2480.

[30] BESSON O, VINCENT F, STOICA P, et al. Approximate maximum likelihood estimators for array processing in multiplicative noise environments. IEEE Trans. on Signal Processing, 2000, 48(9): 2506–2518.

[31] LEITE W S, DELAMARE R C. List-based OMP and an enhanced model for DOA estimation with nonuniform arrays. IEEE Trans. on Aerospace and Electronic Systems, 2021, 57(6): 4457–4464.

[32] BLASONE G P, COLONE F, LOMBARDO P, et al. Passive radar STAP detection and DOA estimation under antenna calibration errors. IEEE Trans. on Aerospace and Electronic Systems, 2021, 57(5): 2725–2742.

[33] LIU S Y, ZHANG Z, GUO Y. 2D DOA estimation with imperfect L-shaped array using active calibration. IEEE Communications Letters, 2021, 25(4): 1178–1182.

[34] QIU S, SHENG W X, MA X F, et al. A maximum likelihood method for joint DOA and polarization estimation based on manifold separation. IEEE Trans. on Aerospace and Electronic Systems, 2021, 57(4): 2481–2500.

[35] GEETHA P, NANDA S J, YADAV R P. A parallel chaotic sailfish optimization algorithm for estimation of DOA in wireless sensor array. Physical Communication, 2022, 51(1): 101536.

[36] SAHA S, KAY S M. Maximum likelihood parameter estimation of superimposed chirps using Monte Carlo importance sampling. IEEE Trans. on Signal Processing, 2002, 50(2): 224–230.

[37] WANG H G, KAY S M. Maximum likelihood angle-Doppler estimator using importance sampling. IEEE Trans. on Aerospace and Electronic Systems, 2010, 46(2): 610–622.

[38] WANG H G, KAY S M, SAHA S. An importance sampling maximum likelihood direction-of-arrival estimator. IEEE Trans. on Signal Processing, 2008, 56(10): 5082–5092.

[39] PINCUS M. A closed form solution for certain programming problems. Operations Research, 1968, 16: 690–694.

[40] SMITH A F, GELFAND A. Bayesian statistics without tears: a sampling-resampling framework. American Statistician, 1992, 46: 84–88.

[41] KAY S. Intuitive probability and random processes using MATLAB. New York: Springer, 2006.

[42] JAFFER A G. Maximum likelihood direction finding for stochasticsources: a separable solution. Proc. of the International Conference on Acoustics, Speech, and Signal Processing, 1988: 2893–2896.

[43] KAY S. Comments on frequency estimation by linear prediction. IEEE Trans. on Acoustics Speech and Signal Processing, 1979, 27(2): 198–199.

Biographies

WU Tao was born in 1984, He obtained his B.S. degree in electronic engineering in 2005 and M.S. degree in system engineering in 2008. He is currently pursuing his Ph.D. degree in Automation School of Northwest Polytechnic University. He is a lecturer and has been engaged in teaching and scientific research of array signal processing and system modelling and simulation. He has published more than 10 papers. His main research interests are array signal processing and complex system modelling and simulation.
E-mail: taowu_nwpu@126.com
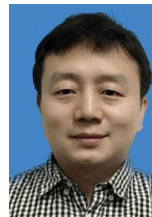
DENG Zhenghong was born in 1974. He received his M.S. degree in computer science and engineering in 1999 and Ph.D. degree from Northwestern Polytechnical University in 2002, respectively. Currently, he is a professor at Automation School of Northwest Polytechnic University. In 2013, he was invited to University of Victoria in Canada as a short-term research fellow. His main research interests include statistical signal processing, array signal processing, underwater acoustical image processing, and acoustic localization.
E-mail: dthree@nwpu.edu.cn

HU Xiaoxiang was born in 1982. He received his M.S. degree in guidance, navigation, and control, and Ph.D. degree in control science and engineering from the Xi'an Research Institute of High-Tech, Xi'an, China, in 2008 and 2012, respectively. He is currently an associate professor in the School of Automation, Northwestern Polytechnical University. His current research interests include fuzzy control, sliding mode control, and hypersonic flight vehicles.
E-mail: xxhu@nwpu.edu

LI Ao was born in 1993. He received his B.S. degree in electronic engineering and automation in 2014, and M.S. degree in electronic engineering in 2018. He is currently pursuing his Ph.D. degree in the School of Automation, Northwest Polytechnic University. His main research interests are model predict control and cooperative control of UAVs.
E-mail: ao.li@mail.nwpu.edu.cn
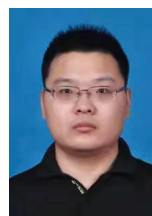
XU Jiwei was born in 1986. He obtained his M.S. degree in signal processing engineering in 2014 and Ph.D. degree from Northwestern Polytechnical University in 2019. He is an associate professor with Xi'an University of Posts and Telecommunications. His research interests include underwater signal processing and sensor network
E-mail: xu@xupt.edu.cn