Signal Processing 114 (2015) 265–276

Contents lists available at ScienceDirect

Signal Processing

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

Fast and efficient DOA estimation method for signals with known waveforms using nonuniform linear arrays

Jian-Feng Gu [a], Wei-Ping Zhu [b,*], M.N.S. Swamy [b]

[a] Department of Electrical Engineering, Ecole Polytechnique de Montreal, QC, Canada H3C 3A7
[b] Department of Electrical and Computer Engineering, Concordia University, Montreal, QC, Canada H3G 1M8

ARTICLE INFO

Article history:
Received 24 April 2014
Received in revised form
1 September 2014
Accepted 21 February 2015
Available online 11 March 2015

Keywords:
DOA estimation
Known waveform
Sparse linear array
Maximum likelihood problem
Regression analysis

ABSTRACT

In this paper, a new approach is proposed to estimate the direction of arrival (DOA) of multiple non-coherent source signals with known waveforms but unknown gains based on a nonuniform linear sensor array. Unlike some previous methods, which estimate the DOA using spatial signatures of the signals with known waveforms, the proposed method first uses the known waveforms and mutually independent sensor measurement noises to establish a maximum likelihood estimation problem corresponding to multiple linear regression models, each containing the DOA and the gain information of all the source signals. Then, regression analysis is performed to estimate the coefficients of each linear regression model, and the well-known generalized least squares is used to obtain the estimates of the angles and gains from the estimated regression coefficients. The proposed method does not require a search over a large region of the parameter space, which is normally needed in ML-based DOA estimation methods. The effect of correlated sources on the performance of the parameter estimation is also studied. It is shown that the DOA and gain estimates are asymptotically optimal as the sources tend to be uncorrelated. Finally, simulation results that demonstrate the estimation performance of the proposed method are given.

1. Introduction

Direction of arrival estimation of multiple narrow-band sources by a linear array has been of great interest in a variety of commercial and military applications such as communications, sonar, air traffic control and electronic reconnaissance areas [1–3]. Many algorithms, such as beamforming-based techniques [4], subspace-based techniques [5–7], and sparsity-based techniques [8,9], have been proposed to solve the problem of estimating closely spaced source directions in the presence of noise. These techniques are mainly based on a common assumption that the received signals are non-cooperative signals, i.e., these signals are either unknown deterministic signals or Gaussian processes type of stochastic signal sources with unknown covariance. However, in active radar, active sonar, communication systems, and many other multisensor applications, the signals of interest may contain useful temporal information, such as training symbols, that can be exploited to effectively eliminate background noises and significantly improve the estimation performance [10,11]. In addition, the capacity of DOA estimation can be larger than the number of antenna elements [12–18].

Only a few techniques have been developed so far to handle the DOA estimation problem by making use of the waveforms of signal sources. Li and Compton [12] are among the very first researchers to improve the accuracy of DOA estimation with known waveforms. They obtained initial angle estimates using an iterative quadratic maximum likelihood (IQML) algorithm, and then used the alternating projection (AP) or the expectation maximization

* Corresponding author.
E-mail address: weiping@ece.concordia.ca (W.-P. Zhu).

http://dx.doi.org/10.1016/j.sigpro.2015.02.018
0165-1684/

(EM) algorithm to estimate the DOAs. Later on, a large sample decoupled ML estimator (DEML) was proposed to estimate the DOAs of incoherent signals with known waveforms [13]. The DEML estimator is computationally efficient, since it decouples the multidimensional minimization problem into a set of 1-D minimization problems. However, this estimator encounters the difficulty when the signals impinging on the array are coherent. To lift this constraint, Cedervall and Moses [14] extended the DEML estimator to decorrelate the coherency of incident signals and developed the coherent decoupled maximum likelihood (CDEML) algorithm. Both DEML and CDEML belong to the family of large sample ML algorithms, which do not work well in difficult scenarios such as when the SNR is low or the number of snapshots is small. To improve the accuracy and spatial resolution capacity of the DOA estimation for signals with known waveforms, Li et al. [15] proposed a white decoupled maximum likelihood estimator (WDEML) under the assumption that the observed noise is spatially white. Recently, Najjar Atallah and Marcos [16] have presented a parallel decomposition (PADEC) algorithm that attains comparable performance but with a lower complexity than that of the ML-based algorithms. The idea behind the PADEC algorithm is to obtain spatial signature of the signals using the least-squares (LS) error criterion, and to decorrelate the coherence of the signals by applying spatial smoothing techniques. However, for large size subarrays, the computational burden of PADEC may be unacceptably high, since the eigendecomposition is required to obtain the orthogonal projector on the noise subspace or the signal subspace. A computationally simpler and more efficient DOA estimation technique has been proposed in [17], where the DOA of known signal waveforms is computed based on the phase shift between two subarrays. This technique requires that signals from different sources be uncorrelated with one another; thus, it does not perform well when the signals are partially or completely correlated. More recently, Gu et al. [18] have suggested a fast linear operator to deal with DOA estimation of uncorrelated or coherent signal sources based on their waveforms. This method does not require the reconstruction of orthogonal projector in the noise subspace or the signal subspace, but its performance approaches to that of the ML-based methods.

Although the above mentioned techniques can provide more accurate DOA estimation and handle more signal sources as compared to the conventional methods without using the waveforms of source signals, they suffer from the restriction of array configuration such as uniform linear array (ULA) [19]. To cope with this weakness, in this paper, a new algorithm is proposed to deal with the problem of DOA estimation for signals with known waveforms but unknown complex gains based on a nonuniform or sparse linear array (SLA) [20]. By using linear regression analysis [21], the proposed algorithm is presented as an optimal estimator for simultaneous DOA and complex gain estimation. The output of each sensor of the antenna array, as a combination of the received signals of interest, is expressed as a linear regression model where each regression coefficient contains the information of DOA and the corresponding complex gain. Since the LS solution for linear regression model with zero-mean Gaussian noise is identical to the ML one [22], which is the best linear unbiased estimator (BLUE), we employ an LS-based method to estimate the coefficients for each sensor. Moreover, the well-known generalized LS technique is used to obtain asymptotically optimal estimate of DOAs without requiring heavy computation. Finally, simulations are carried out, showing that the proposed method needs less computational load to achieve the same performance as the ML-based method and obtains better estimation performance than the fast algorithm does.

The rest of this paper is organized as follows. In Section 2, the signal model along with some necessary assumptions is given. In Section 3, a ML estimation problem is transformed into the LS problem for dealing with multiple linear regression models and then regression analysis is also provided. In Section 4, an asymptotic optimal estimator for DOA is derived based on the WLS technique followed by a statistical performance analysis. In addition, a case study for the theoretical performance and the computational complexity is provided in the end. Several simulation experiments showing the estimation performance of the proposed method are provided in Section 5. Finally, we give our conclusions in Section 6.

2. Signal model

As shown in Fig. 1, suppose that K narrowband source signals of wavelength λ from directions $\theta_1, \theta_2, \ldots, \theta_K$ impinge onto a SLA with omnidirectional sensors whose inter-element spacings $d_1, d_2, \ldots, d_M$ are normalized in terms of $\lambda/2$, where the sensor element at the origin is used for referencing. At time t, $y_m(t)$ $(m = 0, 1, \ldots, M)$ denotes the complex signal at the mth sensor of the linear array with $y_0(t)$ being the data received by the reference element. Then, the complex signals observed at the outputs of the $(M+1)$ sensors $\mathbf{y}(t) \in \mathbb{C}^{(M+1)}$ can be written as

$$\mathbf{y}(t) = [y_0(t), y_1(t), \ldots, y_M(t)]^T \triangleq \mathbf{A}\mathbf{s}(t) + \mathbf{e}(t). \tag{1}$$

Here, the superscript T denotes matrix transpose, $\mathbf{s}(t) = \boldsymbol{\Gamma}\mathbf{x}(t) \in \mathbb{C}^K$ is the source signal in which $\mathbf{x}(t) = [x_1(t), x_2(t), \ldots, x_K(t)]^T$ denotes K known signal waveforms and $\boldsymbol{\Gamma} = \operatorname{diag}[\gamma_1, \gamma_2, \ldots, \gamma_K]$ with the diagonal elements denoting the unknown complex gains of the K signals, and $\mathbf{e}(t) = [e_0(t), e_2(t), \ldots, e_M(t)]^T$ is an $(M+1)$-dimensional vector representing the complex white Gaussian measurement noise with zero mean and unknown covariance matrix $\mathbf{R} \triangleq E\{\mathbf{e}(t)\mathbf{e}^H(t)\} = \operatorname{diag}\{\sigma_0^2, \sigma_1^2, \ldots, \sigma_M^2\}$, the superscript H denotes the conjugate transpose of a vector or matrix. Note that the real part $\operatorname{Re}(\mathbf{e}(t))$ and the imaginary part $\operatorname{Im}(\mathbf{e}(t))$ are

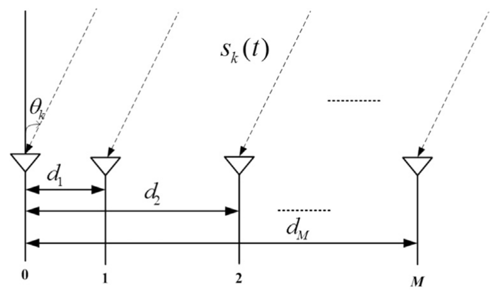

Fig. 1. Considered sensor array geometry.

two real Gaussian random vectors of the same distribution $\mathbb{N}(\mathbf{0}, \mathbf{R}/\mathbf{2})$. Finally, $\mathbf{A}(\boldsymbol{\theta})$ is the $(M+1)\times K$ array manifold matrix of the whole array including the reference sensor, as given by

$$\mathbf{A}(\boldsymbol{\theta}) = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ e^{-j\phi_1^{(1)}} & e^{-j\phi_2^{(1)}} & \cdots & e^{-j\phi_K^{(1)}} \\ \vdots & \vdots & \vdots & \vdots \\ e^{-j\phi_1^{(M)}} & e^{-j\phi_2^{(M)}} & \cdots & e^{-j\phi_K^{(M)}} \end{bmatrix}, \tag{2}$$

where $\phi_k^{(m)} = d_m \varphi_k$ with $\varphi_k = \pi \sin\theta_k$ denotes the phase difference between the mth $(m=1,2,\ldots,M)$ sensor and the reference sensor for the kth $(k=1,2,\ldots,K)$ signal. Since the signal waveforms are known, we can describe the output $y_m(t)$ of the mth sensor of the array by the following linear regression model:

$$\begin{aligned} y_m(t) &= \sum_{k=1}^{K} x_k(t)\gamma_k e^{-j\phi_k^{(m)}} + e_m(t) \\ &= (\boldsymbol{\Gamma}\mathbf{x}(t))^T \mathbf{A}_m^T(\boldsymbol{\theta}) + e_m(t) \triangleq \mathbf{x}^T(t)\mathbf{b}_m + e_m(t), \end{aligned} \tag{3}$$

where $\mathbf{A}_m(\boldsymbol{\theta})$ is the mth row of the matrix $\mathbf{A}(\boldsymbol{\theta})$ and $\mathbf{b}_m \triangleq = \boldsymbol{\Gamma}\mathbf{A}_m^T(\boldsymbol{\theta})$ named the regression coefficients contain the entire information of directions and complex gains of all the K signals. Thus, using (3), (1) can be rewritten as

$$\mathbf{y}(t) = (\mathbf{I}_{M+1} \otimes \mathbf{x}^T(t))\,\mathrm{vec}(\mathbf{B}) + \mathbf{e}(t), \tag{4}$$

where the symbol $\otimes$ denotes the Kronecker product of matrices, and $\mathbf{I}_{M+1}$ is the identity matrix of dimension $M+1$, $\mathbf{B} \triangleq [\mathbf{b}_0, \mathbf{b}_1, \ldots, \mathbf{b}_M]$ is called the regression coefficient matrix, and $\mathrm{vec}(\mathbf{B})$ is the vectorization of $\mathbf{B}$.

It is important to stress that in the conventional methods, the spatial signatures are exploited to estimate the DOAs, while herein we pay special attention to the coefficient matrix $\mathbf{B}$ or its column vectors $\mathbf{b}_m$ $(m=0,1,\ldots,M)$ of the linear regression model. In the following section, an LS estimator for $\mathbf{b}_m$ is derived based on the maximum likelihood principle and the signal model obtained above.

3. Estimation of regression coefficient matrix

In this section, we will first state the ML estimation problem for the linear regression models and then convert it into multiple LS problems to estimate the matrix $\mathbf{B}$. The estimation performance of the regression coefficients is also analyzed for finite snapshots based on regression analysis.

3.1. ML estimator

Suppose the received array signals during the time period T are sampled as

$$\mathbf{y}(nT_s) = (\mathbf{I}_{M+1} \otimes \mathbf{x}^T(nT_s))\,\mathrm{vec}(\mathbf{B}) + \mathbf{e}(nT_s), \quad 1 \le n \le N \tag{5}$$

where $T_s = T/N$ is sampling interval and $\mathbf{e}(n)$ represents the samples of the noise $\mathbf{e}(t)$. The sampled version of the received data is often called "snapshots". By omitting T_s for convenience in sequel, (5) can be rewritten as

$$\mathbf{y}(n) = (\mathbf{I}_{M+1} \otimes \mathbf{x}^T(n))\,\mathrm{vec}(\mathbf{B}) + \mathbf{e}(n), \quad 1 \le n \le N \tag{6}$$

Since the signals are assumed to have known waveforms with unknown complex gains and $\mathbf{e}(n)$ is zero-mean i.i.d Gaussian, each snapshot has a complex Gaussian probability density function (pdf) with a different mean but the same covariance, i.e., $\mathbf{y}(n) \sim \mathbb{CN}\,((\mathbf{I}_{M+1} \otimes \mathbf{x}^T(n))\,\mathrm{vec}(\mathbf{B}), \mathbf{R})$. The joint pdf of the N independent snapshots $\mathbf{y}(n)$ $(n=1,2,\ldots,N)$ can be written as

$$f(\mathbf{Y}; \mathbf{B}, \mathbf{R}) = \prod_{n=1}^{N} \frac{e^{(-\mathbf{e}^H(n)\mathbf{R}^{-1}\mathbf{e}(n))}}{\pi^{M+1}\det(\mathbf{R})}, \tag{7}$$

where $\mathbf{Y} = [\mathbf{y}(1), \mathbf{y}(2), \ldots, \mathbf{y}(N)]$ is an $(M+1)\times N$ measurement matrix of the whole array, and $\det(\bullet)$ denotes the determinant of the matrix involved or the absolute value of a scalar. Assume that all the waveforms of K signals are known and the array measurement noises are uncorrelated with each other. Then, the ML estimate of the noise variance σ_m^2 is given by

$$\hat{\sigma}_m^2 = \frac{\mathbf{y}_m - \mathbf{X}^T\mathbf{b}_m}{N}, \tag{8}$$

and the ML estimate of $\mathbf{B}$ in (7) can be represented by the LS estimates for the $(M+1)$ linear regression coefficient vectors, namely,

$$\hat{\mathbf{b}}_m = \arg\min_{\mathbf{b}_m} \|\mathbf{y}_m - \mathbf{X}^T\mathbf{b}_m\|^2 \quad (m=0,1,\ldots,M), \tag{9}$$

where $\mathbf{b}_m$ denotes the coefficients of the mth regression model of order K, $\mathbf{y}_m = [y_m(1), y_m(2), \ldots, y_m(N)]$ is N measurements of the mth sensor and $\mathbf{X} = [\mathbf{x}(1), \mathbf{x}(2), \ldots, \mathbf{x}(N)]$ is the $K\times N$ matrix according to the known waveform.

If the signals are not coherent and there are sufficient snapshots for each interval such that matrix $\mathbf{X}^*\mathbf{X}^T$ is nonsingular, the solution to (9) is then given by

$$\hat{\mathbf{b}}_m = \mathbf{R}_{xx}^{-1}\mathbf{R}_{xy}^{(m)}, \tag{10}$$

where

$$\mathbf{R}_{xx} \triangleq \frac{\sum_{n=1}^{N}[\mathbf{x}^*(n)\mathbf{x}^T(n)]}{N} = \frac{\mathbf{X}^*\mathbf{X}^T}{N} \tag{11}$$

and

$$\mathbf{R}_{xy}^{(m)} \triangleq \frac{\sum_{n=1}^{N}[\mathbf{x}^*(n)y_m(n)]}{N} = \frac{\mathbf{X}^*\mathbf{y}_m}{N}. \tag{12}$$

Note that when the signals are highly correlated with one another or only a small number of snapshots are available, $\mathbf{R}_{xx}$ might be singular or nearly singular. In that case, we cannot obtain an accurate estimate of $\mathbf{b}_m$ and need some modifications for the rank-deficient matrix $\mathbf{R}_{xx}$. For instance, ridge regression or principal component regression methods [23,24] can be utilized to obtain the generalized inverse of $\mathbf{R}_{xx}$. It is also noted that the previous methods have made use of the known waveforms to obtain the spatial signatures for estimating DOAs and complex gains with the traditional subspace techniques such as the MUSIC and ESPRIT. In this paper, however, we exploit the LS techniques and the projection matrix of the known waveforms to obtain the coefficients for each sensor, which contain the entire information of DOAs and complex gains. Therefore, the next subsection is devoted to the regression analysis of statistical properties of the estimated coefficients.

3.2. Regression analysis

Using (3) and letting $\mathbf{b}_m \triangleq \hat{\mathbf{b}}_m - \Delta\mathbf{b}_m$, where $\Delta\mathbf{b}_m = [\Delta b_1^{(m)}, \Delta b_2^{(m)}, \ldots, \Delta b_K^{(m)}]^T$ is the estimation error vector of the coefficients $\mathbf{b}_m$, we get

$$\mathbf{y}_m = \mathbf{X}^T\hat{\mathbf{b}}_m - \mathbf{X}^T\Delta\mathbf{b}_m + \mathbf{e}_m = \hat{\mathbf{y}}_m + \boldsymbol{\varepsilon}_m, \tag{13}$$

where $\mathbf{e}_m = [e_m(1), e_m(2), \ldots, e_m(N)]^T$ is the N-dimensional white Gaussian noise vector for the mth sensor, $\hat{\mathbf{y}}_m = \mathbf{X}^T\hat{\mathbf{b}}_m$ is the estimate of $\mathbf{y}_m$ and $\boldsymbol{\varepsilon}_m = -\mathbf{X}^T\Delta\mathbf{b}_m + \mathbf{e}_m = [\varepsilon_m(1), \varepsilon_m(2), \ldots, \varepsilon_m(N)]^T$ is defined as the LS residual. Obviously, $\Delta\mathbf{b}_m$ can be written as

$$\Delta\mathbf{b}_m = (\mathbf{X}^*\mathbf{X}^T)^{-1}\mathbf{X}^*(\mathbf{e}_m - \boldsymbol{\varepsilon}_m). \tag{14}$$

On the other hand, using (10) and (13), and recalling that $\hat{\mathbf{y}}_m = \mathbf{X}^T\hat{\mathbf{b}}_m$ and $\mathbf{y}_m = \mathbf{X}^T\mathbf{b}_m + \mathbf{e}_m$ we can obtain

$$\begin{aligned}\boldsymbol{\varepsilon}_m &= \mathbf{y}_m - \hat{\mathbf{y}}_m \\ &= (\mathbf{I}_N - \mathbf{X}^T(\mathbf{X}^*\mathbf{X}^T)^{-1}\mathbf{X}^*)\mathbf{y}_m \\ &= (\mathbf{I}_N - P_{\mathbf{X}^T})\mathbf{e}_m,\end{aligned} \tag{15}$$

where $P_\bullet$ is the projection operator onto the space spanned by the columns of the matrix $\bullet$.

Then, substituting (15) into (14) and recalling that $\mathbf{e}(n) = [e_0(n), e_1(n), \ldots, e_M(n)]^T$ has zero mean, the mean of the estimation error $\Delta\mathbf{b}_m$ is easily obtained as

$$\begin{aligned}E\{\Delta\mathbf{b}_m\} &= (\mathbf{X}^*\mathbf{X}^T)^{-1}\mathbf{X}^*E\{\mathbf{e}_m - (\mathbf{I}_N - P_{\mathbf{X}^T})\mathbf{e}_m\} \\ &= (\mathbf{X}^*\mathbf{X}^T)^{-1}\mathbf{X}^*E[\mathbf{e}_m] = \mathbf{0},\end{aligned} \tag{16}$$

implying that the estimate $\hat{\mathbf{b}}_m$ is unbiased or $E\{\hat{\mathbf{b}}_m\} = \mathbf{b}_m$. It is easy to show that the covariance matrix of $\hat{\mathbf{b}}_m$ is given by

$$\begin{aligned}\mathbf{C}_{\hat{\mathbf{b}}_m} &= E\{\Delta\mathbf{b}_m\Delta\mathbf{b}_m^H\} \\ &= (\mathbf{X}^*\mathbf{X}^T)^{-1}\mathbf{X}^*E\{\mathbf{e}_m\mathbf{e}_m^H\}\mathbf{X}^T(\mathbf{X}^*\mathbf{X}^T)^{-1} \\ &= \sigma_m^2[\mathbf{X}^*\mathbf{X}^T]^{-1}.\end{aligned} \tag{17}$$

In obtaining (17), we have used the fact that $E\{\mathbf{e}_m\mathbf{e}_m^H\} = \sigma_m^2\mathbf{I}_N$. By using (10), the above equation can be rewritten as

$$\mathbf{C}_{\hat{\mathbf{b}}_m} = \frac{\sigma_m^2\mathbf{R}_{xx}^{-1}}{N}. \tag{18}$$

Clearly, the covariance matrix $\mathbf{C}_{\hat{\mathbf{b}}_m}$ is inversely proportional to the number of snapshots. When N approaches infinity, $\mathbf{C}_{\hat{\mathbf{b}}_m}$ tends to zero. Note that when the incident signals are uncorrelated with one another in the time interval, i.e., the elements in (11) satisfy $\sum_{n=1}^{N} x_i^*(n)x_j(n) = 0$, for $i \neq j$, and $i, j \in [1, 2, \ldots, K]$, we have $\mathbf{R}_{xx} = \text{diag}([\rho_1, \rho_2, \ldots, \rho_K])$, where ρ_k $(k = 1, 2, \ldots, K)$ is the power of the kth known waveform, since the estimation error $\Delta b_k^{(m)}$ of $b_k^{(m)}$ has a zero-mean white Gaussian distribution with variance $\text{Var}(\Delta b_k^{(m)}) = 1/(N \cdot \text{WNR}_k^{(m)})$, $\text{WNR}_k^{(m)} \triangleq \rho_k/\sigma_m^2$ being the waveform to noise ratio (WNR) of the kth signal on the mth sensor.

It is also clear from (18) that when the incident signals are coherent, one cannot get the covariance matrix $\mathbf{C}_{\hat{\mathbf{b}}_m}$ due to the fact that the matrix $\mathbf{R}_{xx}$ becomes singular. When the signals are partly correlated with one another, the estimation error $\Delta b_k^{(m)}$ of $b_k^{(m)}$ has a zero-mean white Gaussian distribution with variance $\text{Var}(\Delta b_k^{(m)}) = \sigma_m^2[\mathbf{R}_{xx}^{-1}]_{kk}/N$. As the variance of $\hat{b}_k^{(m)}$ depends on the unknown noise variance σ_m^2 of the mth sensor, we show below how to get an estimate of σ_m^2.

Using (15), the square sum of the LS residual $\boldsymbol{\varepsilon}_m$ is given by

$$\boldsymbol{\varepsilon}_m^H\boldsymbol{\varepsilon}_m = \mathbf{e}_m^H(\mathbf{I}_N - P_{\mathbf{X}^T})\mathbf{e}_m = \text{Tr}(P_{\mathbf{X}^T}^{\perp}\mathbf{e}_m\mathbf{e}_m^H), \tag{19}$$

where $P_\bullet^{\perp}$ is the orthogonal projection of the matrix $\bullet$. Thus, the expectation of (19) is derived as

$$E\{\boldsymbol{\varepsilon}_m^H\boldsymbol{\varepsilon}_m\} = E\{\text{Tr}(P_{\mathbf{X}^T}^{\perp}\mathbf{e}_m\mathbf{e}_m^H)\} = (N-K)\sigma_m^2, \tag{20}$$

where we have used the property of an idempotent matrix [25] that the trace equals the rank of the matrix, namely, $\text{Tr}(P_{\mathbf{X}^T}^{\perp}) = \text{Rank}(\mathbf{I}_N) - \text{Rank}(P_{\mathbf{X}^T})$. Finally, we can obtain an unbiased ML estimate of σ_m^2 from

$$\begin{aligned}\hat{\sigma}_m^2 &= \frac{\|\mathbf{y}_m - \hat{\mathbf{y}}_m\|^2}{N-K} = \frac{\sum_{n=1}^{N}\varepsilon_m(n)\varepsilon_m^*(n)}{N-K} \\ &= \frac{\|\mathbf{y}_m - \mathbf{X}^T\hat{\mathbf{b}}_m + \mathbf{X}^T\Delta\mathbf{b}_m\|^2}{N-K},\end{aligned} \tag{21}$$

where the number $N-K$ is called the degree of freedom (DOF) for the noise estimation, which is the number of snapshots minus the number of source signals. Note that when the number of snapshots is much larger than the number of source signals, which is quite common in array signal processing, the ML estimate of σ_m^2 approaches the average of the squared sum of the residuals (errors) $\varepsilon_m(n)$, $n = 1, 2, \ldots, N$. Having obtained the properties of the regression coefficients $\mathbf{b}_m$, which contain the whole information of the DOA and complex gain, we will present an optimal estimator for the DOA and complex gain in the next section.

4. DOA and complex gain estimation

In this section, we will estimate the DOA and complex gain using the estimated regression coefficient matrix and study its theoretical performance.

4.1. The estimation algorithm and its performance

Recalling that $\mathbf{b}_m \triangleq \boldsymbol{\Gamma}\mathbf{A}_m^T(\boldsymbol{\theta}) = \hat{\mathbf{b}}_m - \Delta\mathbf{b}_m$, we have

$$\hat{b}_k^{(0)} = \gamma_k + \Delta b_k^{(0)}, \tag{22}$$

$$\hat{b}_k^{(m)} = \gamma_k e^{-j\phi_k^{(m)}} + \Delta b_k^{(m)} \quad (m = 1, 2, \ldots, M), \tag{23}$$

where $\Delta b_k^{(m)} \triangleq \text{Re}(\Delta b_k^{(m)}) + j\,\text{Im}(\Delta b_k^{(m)})$ is the estimation error which is described by a complex Gaussian process with zero mean and variance of $\sigma_m^2[\mathbf{R}_{xx}^{-1}]_{k,k}/N$. Note that $\text{Re}(\Delta b_k^{(m)})$ and $\text{Im}(\Delta b_k^{(m)})$ are uncorrelated with one another, each having zero mean and variance of $\sigma_m^2[\mathbf{R}_{xx}^{-1}]_{k,k}/2N$ (see Appendix A for proof). When both sides of (23) are multiplied by the complex conjugate of (22) and then divided by the amplitude square of the $|\gamma_k|^2$, we get

$$\hat{a}_k^{(m)} \triangleq \frac{\hat{b}_k^{(m)}\hat{b}_k^{(0)*}}{|\gamma_k|^2} = e^{-j\phi_k^{(m)}} + \Delta a_k^{(m)}, \tag{24}$$

which can be considered as the estimated value of the mth element of the kth steering vector $\boldsymbol{a}(\theta_k)$, where $\Delta a_k^{(m)}$ is the estimation error with zero mean and variance of

$\mathrm{Var}(\Delta a_k^{(m)}) = (\sigma_m^2+\sigma_0^2)\left[\mathbf{R}_{xx}^{-1}\right]_{k,k}/N|\gamma_k|^2$ and the covariance of $\mathrm{Cov}(\Delta a_k^{(p)}, \Delta a_k^{(q)}) = \sigma_0^2\left[\mathbf{R}_{xx}^{-1}\right]_{k,k}/N|\gamma_k|^2$, $p \neq q \in [1,\ldots,M]$ (see Appendix B for derivation). Obviously, the unknown power gain $|\gamma_k|^2$ has no effect on the estimation of DOA due to the phase normalization using the amplitude of $\hat{a}_k^{(m)}$. As a result, it is easy to verify that the estimation error vector $\Delta\hat{\mathbf{a}}_k = \left[\Delta\hat{a}_k^{(1)}, \Delta\hat{a}_k^{(2)}, \ldots, \Delta\hat{a}_k^{(M)}\right]^T$ $(k=1,\ldots,K)$ has a distribution $\mathbb{CN}(\mathbf{0}, \mathbf{\Omega}_k)$ with

$$\mathbf{\Omega}_k = \eta_k^{(0)}\begin{bmatrix} \eta_k^{(1)} & & \mathbf{1}_{M-1}^T \\ \mathbf{1}_1 & \eta_k^{(2)} & \mathbf{1}_{M-2}^T \\ \cdots & \ddots & \cdots \\ \mathbf{1}_{M-1}^T & & \eta_k^{(M)} \end{bmatrix}, \tag{25}$$

where $\eta_k^{(0)} \triangleq = \sigma_0^2\left[\mathbf{R}_{xx}^{-1}\right]_{k,k}/N|\gamma_k|^2$, $\eta_k^{(m)} = (\sigma_m^2/\sigma_0^2)+1$ $(m=1,2,\ldots,M)$, and $\mathbf{1}_m$ denotes a m-dimensional column vector whose elements are all unity. If the noise of each sensor is i.i.d. with the same variance, which is often assumed in conventional methods, (25) can be simplified as

$$\mathbf{\Omega}_k = \eta_k^{(0)}\begin{bmatrix} 2 & & \mathbf{1}_{M-1}^T \\ \mathbf{1}_1 & 2 & \mathbf{1}_{M-2}^T \\ \cdots & \ddots & \cdots \\ \mathbf{1}_{M-1}^T & & 2 \end{bmatrix}. \tag{26}$$

Now, let us introduce M-dimensional vector,

$$\mathbf{\Phi}_k = \left[\phi_k^{(1)}, \phi_k^{(2)}, \ldots, \phi_k^{(M)}\right]^T. \tag{27}$$

Based on the above discussions from (22)–(26), the unbiased estimate $\hat{\mathbf{\Phi}}_k$ of $\mathbf{\Phi}_k$ is a Gaussian random vector with the covariance matrix as given by (see the proof in Appendix C)

$$\mathbf{\Sigma}_k \triangleq \mathrm{Cov}\left(\hat{\mathbf{\Phi}}_k\right) \approx \mathbf{\Omega}_k/2. \tag{28}$$

From $\phi_k^{(m)} = d_m\varphi_k = d_m\pi\sin\theta_k$ in (2) and the unbiased estimate $\hat{\phi}_k^{(m)}$ of $\phi_k^{(m)}$, the parameter φ_k can be obtained by solving the following generalized LS problem [26]:

$$\hat{\varphi}_k = \arg\min_{\varphi} \Upsilon_k(\varphi) = \left(\hat{\mathbf{\Phi}}_k - \varphi\mathbf{d}\right)^T \mathbf{\Sigma}_k^{-1}\left(\hat{\mathbf{\Phi}}_k - \varphi\mathbf{d}\right), \tag{29}$$

where $\mathbf{d} = [d_1, d_2, \ldots, d_M]^T$. The solution to the above problem is given by

$$\hat{\varphi}_k = \frac{\mathbf{d}^T\mathbf{\Sigma}_k^{-1}\hat{\mathbf{\Phi}}_k}{\mathbf{d}^T\mathbf{\Sigma}_k^{-1}\mathbf{d}}, \tag{30}$$

which leads to the estimated DOA $\hat{\theta}_k = \arcsin(\hat{\varphi}_k/\pi)$. Note that $\hat{\varphi}_k$ is an unbiased estimate with a Gaussian distribution, since $\Delta\mathbf{\Phi}_k \triangleq \mathbf{\Phi}_k - \hat{\mathbf{\Phi}}_k$ is a Gaussian random vector with zero mean. It is easy to get the variance of this estimate as

$$\mathrm{Var}(\hat{\varphi}_k) = \frac{\mathbf{d}^T\mathbf{\Sigma}_k^{-1}\,\mathrm{Cov}\left(\hat{\mathbf{\Phi}}_k\right)\mathbf{\Sigma}_k^{-1}\mathbf{d}}{\mathbf{d}^T\mathbf{\Sigma}_k^{-1}\mathbf{d}} = \frac{1}{\mathbf{d}^T\mathbf{\Sigma}_k^{-1}\mathbf{d}}. \tag{31}$$

Also, the variance of $\hat{\theta}_k$ can be obtained as

$$\mathrm{Var}\left(\hat{\theta}_k\right) = \frac{1}{[\pi\cos\theta_k]^2}\mathrm{Var}(\hat{\varphi}_k). \tag{32}$$

We now present an estimator for the complex gain. Although, we could estimate the complex gain using (22) directly, yet it is not optimal due to the use of one sensor only. Here, we would like to use the estimated DOAs obtained from all the sensors to determine complex gain.

We first define

$$\hat{\mathbf{\Gamma}}_k = \left[\hat{\gamma}_k^{(0)}, \hat{\gamma}_k^{(1)}, \ldots, \hat{\gamma}_k^{(M)}\right]^T \tag{33}$$

with $\hat{\gamma}_k^{(0)} = \hat{b}_k^{(0)}$ and $\hat{\gamma}_k^{(m)} = \hat{b}_k^{(m)} e^{j\hat{\phi}_k^{(m)}} \triangleq = \gamma_k^{(m)} + \Delta\gamma_k^{(m)}$ $(m=1,2,\ldots,M)$, where $\Delta\gamma_k^{(m)}$ are zero mean Gaussian random variables with the covariance matrix $\mathbf{H}_k$ (see Appendix D for derivation). In a manner similar to deriving (29)–(31), the optimal estimator for γ_k can be obtained as

$$\hat{\gamma}_k = \frac{\mathbf{1}_{M+1}^T\mathbf{H}_k^{-1}\hat{\mathbf{\Gamma}}_k}{\mathbf{1}_{M+1}^T\mathbf{H}_k^{-1}\mathbf{1}_{M+1}}. \tag{34}$$

Also, it is easy to get the variance of this estimator as

$$\mathrm{Var}(\hat{\gamma}_k) = \frac{1}{\mathbf{1}_{M+1}^T\mathbf{H}_k^{-1}\mathbf{1}_{M+1}}. \tag{35}$$

It is worth noting that we have used the estimate of the vector $\mathbf{\Phi}_k$ obtained from the regression coefficient matrix $\mathbf{B}$. Although each element of $\mathbf{\Phi}_k$ in (27) can be estimated from (24) by calculating the phase of $a_k^{(m)}$, ambiguity may arise due to its inability to deal with the 2π ambiguity. In the case of $d_1 > 1$, leading to $\left|\phi_k^{(m)}\right| > \pi$, where $\phi_k^{(m)} = -\angle\left(a_k^{(m)}\right)$. If the assumption $d_1 \le 1$ is made, there is no ambiguity for $\phi_k^{(1)}$, i.e., $\left|\phi_k^{(1)}\right| = d_1\pi|\sin\theta_k| \le \pi$. Then, one can use unambiguous $\phi_k^{(1)}$ as a reference to deal with potential ambiguities for $d_m \ge 1$, $m=2,\ldots,M$, according to the disambiguity method proposed in [27,28]. However, for the case of $d_1 > 1$, there is not an unambiguous value to be used to overcome the problem of the ambiguities, we should exploit the technique suggested in [19] to obtain these unambiguities.

4.2. A case study for ULA

In this subsection, we would like to consider a special case where the source signals are uncorrelated with equal-power, and the uniform linear array (ULA) is of half wavelength inter-element spacing, i.e., $\mathbf{d} = [1,2,\ldots,M]^T$. Note that the additive noises at the ULA are i.i.d Gaussian random processes with $\mathrm{SNR} = (\rho/\sigma^2)|\gamma_k|^2 = \mathrm{WNR}|\gamma_k|^2$, the covariance matrix is then given by

$$[\mathbf{\Sigma}_k]_{p,q} = \begin{cases} \dfrac{1}{N\cdot\mathrm{SNR}}, & 1 \le p = q \le M \\ \dfrac{1}{2N\cdot\mathrm{SNR}} & \text{otherwise,} \end{cases} \tag{36}$$

which can be rewritten in the matrix form as

$$\mathbf{\Sigma}_k = \frac{\mathbf{\Lambda}}{2N\cdot\mathrm{SNR}}, \tag{37}$$

where $\mathbf{\Lambda} = \mathbf{I}_M + \mathbf{1}_M \mathbf{1}_M^T$. By using the Sherman–Morrison formula [29], we have

$$\mathbf{\Lambda}^{-1} = \mathbf{I}_M - \frac{\mathbf{1}_M \mathbf{1}_M^T}{1+M}. \tag{38}$$

Thus

$$\mathbf{\Sigma}_k^{-1} = 2N \cdot \text{SNR}\left(\mathbf{I}_M - \frac{\mathbf{1}_M \mathbf{1}_M^T}{1+M}\right) \tag{39}$$

and

$$\mathbf{d}^T \mathbf{\Sigma}_k^{-1} \mathbf{d} = \frac{(M^2+M)(M+2)}{6} N \cdot \text{SNR} \tag{40}$$

Hence, the variance of the DOA estimate with known noise covariance matrix can be calculated from (31), (32) and (40) as

$$\text{Var}\left(\hat{\theta}_k\right)_{\text{KNOWN}} = \frac{6}{(M^2+M)(M+2)N \cdot \text{SNR}} \left[\frac{1}{\pi \cos \theta_k}\right]^2. \tag{41}$$

Interestingly, using the Cramer–Rao bound (CRB) expression in [12,13], we can find that (41) gives exactly the CRB of θ_k, i.e., $\text{CRB}(\theta_k) = \text{Var}\left(\hat{\theta}_k\right)_{\text{KNOWN}}$. Therefore, our estimator is a MVUE estimator when the additive noise has i.i.d Gaussian distribution. On the other hand, when the noise covariance matrix is unknown, we can also get the variance by the result of (21), namely,

$$\text{Var}\left(\hat{\theta}_k\right)_{\text{UNKNOWN}} = \frac{6}{(M^2+M)(M+2)(N-K) \cdot \text{SNR}} \left[\frac{1}{\pi \cos \theta_k}\right]^2. \tag{42}$$

The CRB to variance ratio is then found from (32) and (42) as

$$\frac{\text{Var}\left(\hat{\theta}_k\right)_{\text{KNOWN}}}{\text{Var}\left(\hat{\theta}_k\right)_{\text{UNKNOWN}}} = 1 - \frac{K}{N}. \tag{43}$$

Usually, the number of snapshots is much larger than the number of signals, i.e., $N \gg K$. Thus the performance of our method approaches CRB when N is large enough. In other words, our proposed method is an asymptotic optimal estimator for DOA estimation for uncorrelated sources.

4.3. Implementation and complexity analysis

Based on the above analysis, the implementation of the proposed method for DOA estimation with known waveforms can be summarized as follows:

1. Compute the regression coefficient matrix using (10).
2. Construct the M-dimensional vectors $\mathbf{\Phi}_k$, $k = 1, 2, \ldots, K$, based on the matrix $\mathbf{B}$.
3. Obtain the phase φ_k using (30) and then the DOA estimate $\theta_k = \arcsin(\varphi_k/\pi)$.
4. Estimate the complex gain γ_k using (34).

The implementation of the proposed method requires two major steps: one is to obtain the matrix $\mathbf{B}$ using (10) and the other is to compute the phase φ_k using (30) and the complex gain γ_k using (34). The number of flops needed to obtain $\mathbf{B}$ is $2N(M+1)K + 2K^2(N+M+1) + o(K^3)$, because it requires approximately $2NK$ flops to obtain each $\mathbf{R}_{xy}^{(m)}$ and takes about $2K^2N + o(K^3)$ to compute $\mathbf{R}_{xx}^{-1}$ according to the algorithm suggested in [30]. The flop is defined as a complex floating-point addition/multiplication operation. The calculation of $\hat{\varphi}_k$ requires roughly $o(M^3)$ flops which is almost the same as that of computing the complex gain. Thus the number of flops required by the proposed algorithm is roughly $2N(M+1+K)K$ in total when $N \gg M$ and K, which occurs often in practical applications of DOA estimation. In comparison, the subarray-based (SB) method suggested in [17] requires nearly $2N(M+1)K$ flops which is the same as the computational burden of the fast linear propagator (LP) method in [18]. Although the DEML [13] needs about $2N(M+1+K)K$ flops to obtain the spatial features, which is the same as that needed for the implementation of our proposed method, yet it requires additional expensive computation to get the DOA. This is because the DEML method needs to find the roots or search the parameter spaces of interest based on the estimated spatial features. Therefore, even if our method requires a little more flops than the LP and SB methods do, it has the asymptotic optimal property as the DEML method does but with much less computational burden.

5. Simulation results

In this section, simulation results are presented to show the performance of the proposed DOA estimation techniques as compared to some of the existing methods. In the first two examples, two signals of equal power with angles $\theta_1 = -5^\circ$ and $\theta_2 = 5^\circ$ impinge onto the sparse linear array (SLA) with 3 sensors separated by $d_1 = 1$ and $d_2 = 3$. A 3-sensor ULA with a half wavelength antenna spacing is also considered for the performance study of the proposed method. The additive noise is uniform white noise, i.e., $\mathbf{R} = \sigma_0^2 \mathbf{I}_{M+1}$, and the complex gains are set to $\mathbf{\Gamma} = \text{diag}(e^{-j\pi/4}, e^{j\pi/4})$ and the SNR is defined as the ratio of the power of the source signal to that of the additive noise at each sensor, i.e., $\text{SNR} = \rho/\sigma_0^2$. Each example contains 1000 independent trials to obtain the root mean square error (RMSE). The LP method [18] and the SB method [17] along with the theoretical RMSE and CRB [12,13] are plotted for performance comparison. In the last two examples, we evaluate the performance of our method for different numbers of sources and sensors.

Example 1. Performance of DOA and complex gain estimation with respect to number of snapshots.

The performance of the proposed DOA estimation method versus the number of snapshots is now assessed for two uncorrelated signals. The number of snapshots is varied from $N = 10$ to $N = 150$, and the SNR is set at 5 dB. Both theoretical and simulated RMSEs of the estimated angles are plotted in Fig. 2, along with the CRBs for ULA and SLA. We see that when the ULA is used, the proposed method yields an estimation performance similar to that of the LP and SB methods, and the three methods are consistent with the theoretical RMSE and the CRBs, especially at a large number of snapshots. More importantly, the performance of the

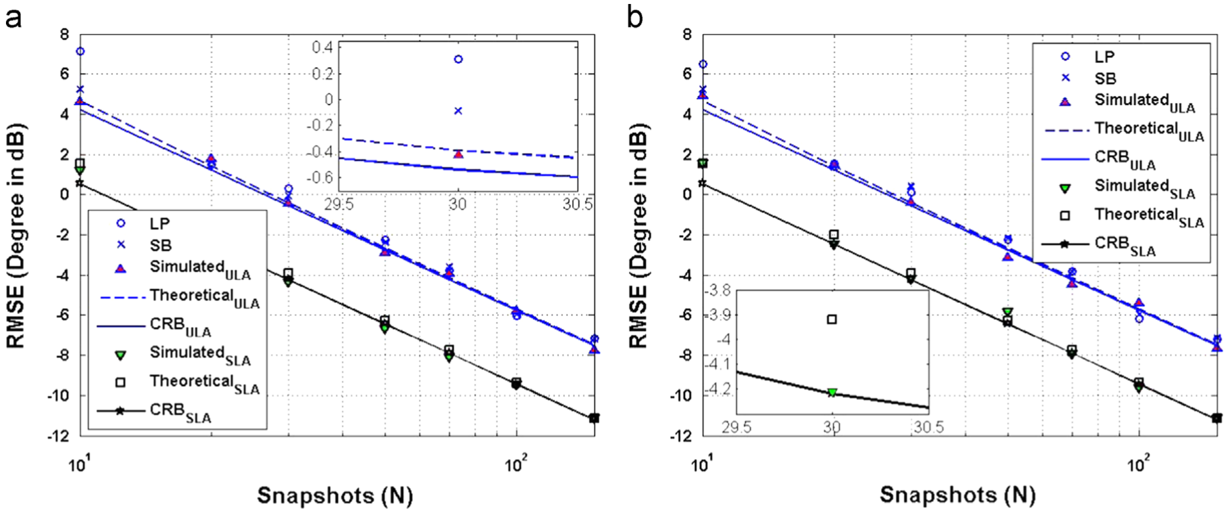

Fig. 2. RMSE of DOA estimation for two uncorrelated sources versus snapshots from 10 to 150 at (a) $\theta_1 = -5^\circ$ and (b) $\theta_2 = 5^\circ$ with SNR=5 dB.

proposed method with SLA is much better than that of ULA due to the larger aperture of the array of SLA. Furthermore, we note that the simulated RMSEs of our method agree very well with the theoretical RMSEs in (32) and CRBs, which decrease almost linearly with the number of snapshots for both ULA and SLA. Note that in these figures "Simulated$_{SLA}$", "Theoretical$_{SLA}$", and "CRB$_{SLA}$" denote the simulated RMSE, and the theoretical RMSE of the proposed method and the CRB for the sparse linear array, while "Simulated$_{ULA}$", "Theoretical$_{ULA}$", and "CRB$_{ULA}$" denote the corresponding results for the uniform linear array. In addition, the performance of our method for complex gain is also compared with the SB method shown in Fig. 3, along with the theoretical performance for ULA. From the figures, we see that the RMSE curve of our method is identical to that of the SB method and the theoretical results.

Example 2. DOA estimation performance for correlated signals.

We study the performance of the proposed method against the correlation between two incident signals. The correlation factor (CF) between the two signals is varied from 0 to 0.9 with SNR=10 dB and the number of snapshots is 200. Both simulated and theoretical RMSEs of the estimated angles against the CF are shown in Fig. 4, together with the CRB. The performance of the SB methods degrades severely at medium and strong correlation of signals, while the RMSE of the proposed method is very close to the theoretical value and that of CRB, especially for weak and moderately correlated signals. The reason for this phenomena lies in the fact that the SB method exploits the spatial signatures of the signals to estimate the DOA, while our method uses the linear regression model at each sensor. The spatial signatures corresponding to known waveforms lead to a leakage of each other because of non-orthogonal property when the signals have medium to high correlation due to the fact that the spatial signatures are obtained by the correlation matrix between the received data and the known waveforms, while our method makes use of the LS-based technique in which the orthogonal projection matrix of the known waveforms is employed to obtain the coefficients. Therefore, our method is very suitable for source signals with medium or high correlation provided that $\mathbf{R}_{xx}$ is a nonsingular matrix.

Example 3. Performance of DOA estimation with respect to the number of sensors.

Now, the performance of the proposed method versus the number of sensors with the same array aperture is assessed. The assumptions for source signals and DOAs are similar to those in Example 1, and the number of snapshots is fixed at 100 and the SNR ranges from 0 dB to 12 dB. The configurations of SLA are set as 3 sensors ($d_1 = 1, d_2 = 4$), 4 sensors ($d_1 = 1, d_2 = 3, d_3 = 1$), 5 sensors ($d_1 = 1, d_2 = 2, d_3 = 1, d_4 = 1$), and 6 sensors ($d_1 = 1, d_2 = 1, d_3 = 1, d_4 = 1, d_5 = 1$). The simulated RMSEs of the estimated DOAs using the proposed method are plotted in Fig. 5, where the theoretical result and CRB for the 6-sensor ULA are also provided. It is true that we observe better performance when adding sensors at the beginning, however, because the aperture is fixed, after adding even more sensors the performance will not change. This is due to the fact that the spatial samples become more and more correlated, hence contain less and less additional information about the signal.

Example 4. Performance of DOA estimation with respect to the number of sources.

In the previous examples, the estimation performance of the proposed method for two sources based on three sensors ($d_1 = 1, d_2 = 5$) is tested. Here, we evaluate its estimation performance in terms of multiple uncorrelated sources: two sources (the same as in the previous examples), three sources DOAs $= [-5^\circ, 5^\circ, 10^\circ]$, $\boldsymbol{\Gamma} = \mathrm{diag}(e^{-j\pi/4}, e^{j\pi/4}, 1))$, four sources (DOAs $= [-5^\circ, 5^\circ, 10^\circ, 15^\circ]$, $\boldsymbol{\Gamma} = \mathrm{diag}(e^{-j\pi/4}, e^{j\pi/4}, 1, -1))$ and five sources (DOAs $= [-5^\circ, 5^\circ, 10^\circ, 15^\circ, 20^\circ]$, $\boldsymbol{\Gamma} = \mathrm{diag}(e^{j\pi/4}, e^{-j\pi/4}, 1, -1, e^{j\pi/6})$). Fig. 6

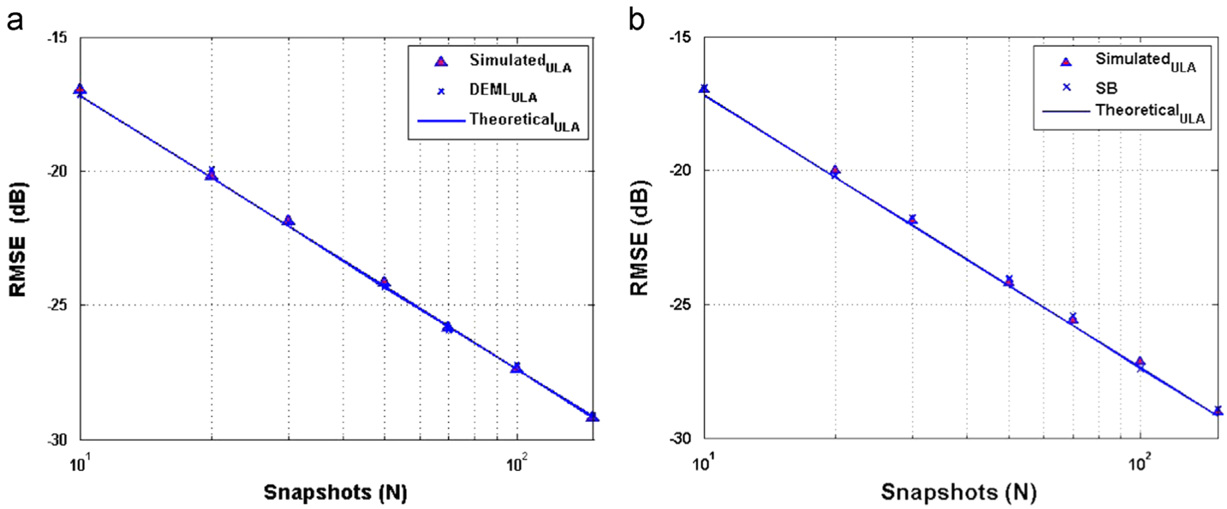

Fig. 3. RMSE of complex gain estimation for two uncorrelated sources versus snapshots from 10 to 150 at (a) $\theta_1=-5^\circ$ and (b) $\theta_2=5^\circ$ with SNR=5 dB.

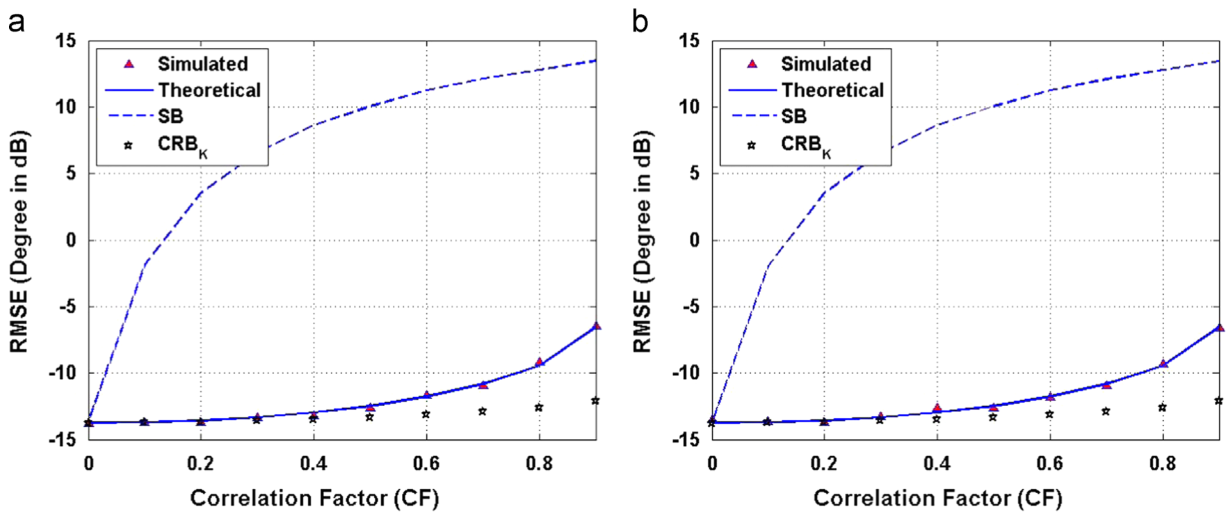

Fig. 4. RMSE of DOA estimation for two sources versus correlation factor (CF) from 0 to 0.9 at (a) $\theta_1=-5^\circ$ and (b) $\theta_2=5^\circ$ with 10 dB SNR and 200 snapshots.

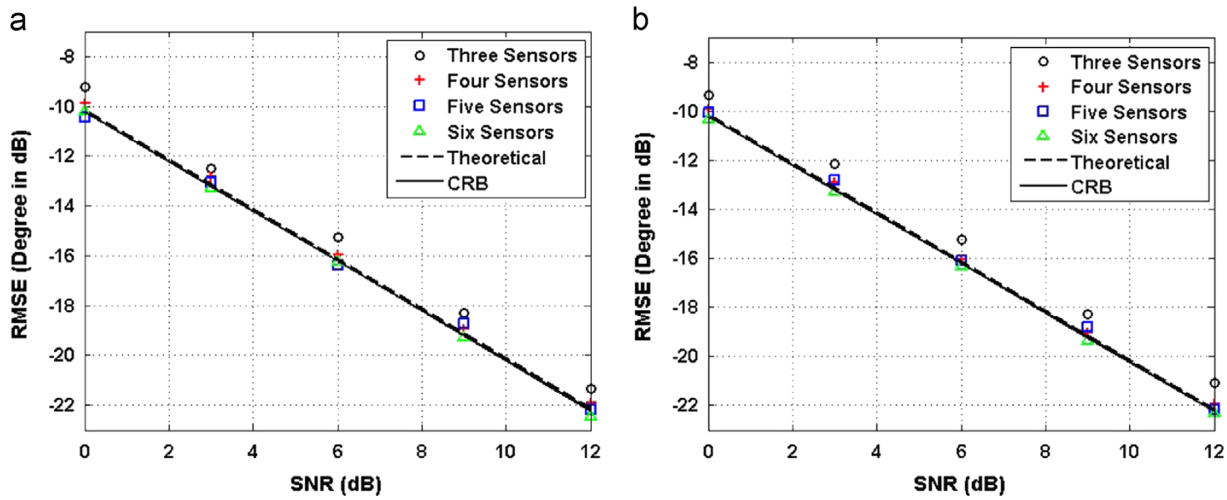

Fig. 5. RMSE of DOA estimation for two sources using multiple sensors versus SNR from 0 to 12 dB at (a) $\theta_1=-5^\circ$ and (b) $\theta_2=5^\circ$ with 100 snapshots.

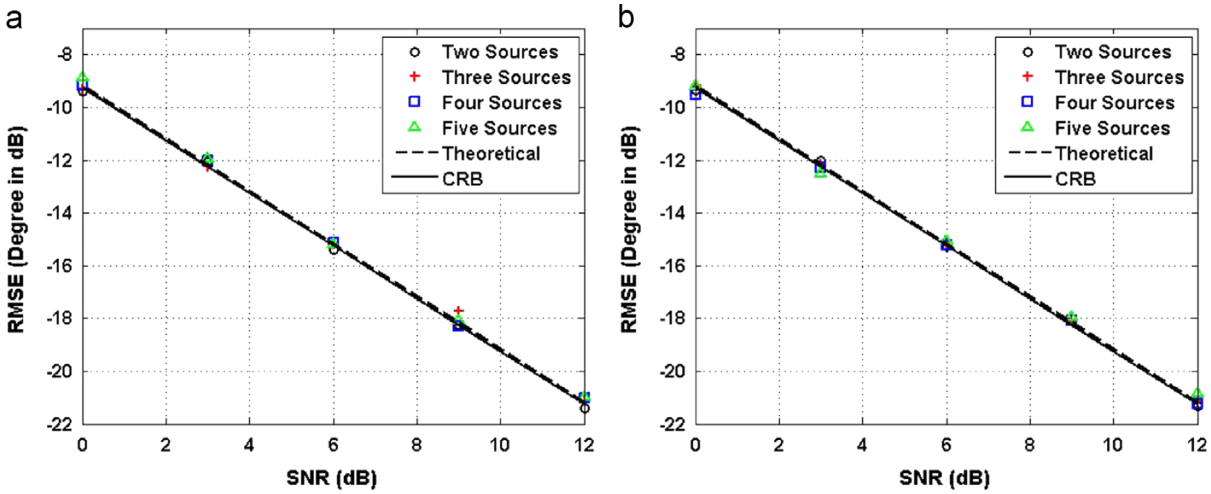

Fig. 6. RMSE of DOA estimation for multiple sources versus SNR from 0 to 12 dB at (a) $\theta_1 = -5^\circ$ and (b) $\theta_2 = 5^\circ$ with 100 snapshots.

shows the performance of the proposed method as a function of the SNR for multiple sources at DOA $= -5^\circ$ and DOA $= 5^\circ$, respectively. We have also plotted the curves of the theoretical results and CRB of the two-source situation for comparison. We note that the proposed method is robust to the number of sources, even when the number of sources is more than the number of sensors, where most of the subspace-based methods with unknown waveforms do not work.

6. Conclusion

In this paper, we have proposed a very efficient DOA estimation method for signals with known waveforms based on sparse linear array. Different from the traditional methods of estimating DOA with known waveforms, our method splits the ML estimation problem into several LS problems, each for a linear regression model representing one sensor, where each coefficient of the linear regression model includes a pair of angle and gain for one source signal. An asymptotically optimal estimator for the DOA and gain has been derived along with its statistical performance analysis for uncorrelated sources. Finally, a number of simulations have been conducted, showing that our proposed method provides a better DOA estimation performance than the SB method, especially for correlated signals, and is close to the CRB.

Appendix A. Statistics of Re($\Delta b_k^{(m)}$) and Im($\Delta b_k^{(m)}$)

From (14) and (15), we have the estimation error vector $\Delta \mathbf{b}_m = [\Delta b_1^{(m)}, \Delta b_2^{(m)}, \ldots, \Delta b_K^{(m)}]^T = \mathbf{\Pi} \mathbf{e}_m$, where $\mathbf{\Pi} \triangleq (\mathbf{X}^* \mathbf{X}^T)^{-1} \mathbf{X}^*$ is a deterministic complex matrix and $\mathbf{b}_m$ is a zero mean complex Gaussian random vector with covariance $\sigma_m^2 \mathbf{R}_{xx}^{(-1)}/N$. We can rewrite the estimation error $\Delta b_k^{(m)}$ as

$$\begin{aligned}\Delta b_k^{(m)} &= \mathrm{Re}(\Delta b_k^{(m)}) + j\,\mathrm{Im}(\Delta b_k^{(m)}) \\ &= [\mathrm{Re}(\mathbf{\Pi}(k,:)) + j\,\mathrm{Im}(\mathbf{\Pi}(k,:))][\mathrm{Re}(\mathbf{e}_m) + j\,\mathrm{Im}(\mathbf{e}_m)] \\ &= \begin{bmatrix} \mathrm{Re}\left(\mathbf{\Pi}^T(k,:)\right) \\ \mathrm{Im}\left(\mathbf{\Pi}^T(k,:)\right) \end{bmatrix}^T \begin{bmatrix} \mathrm{Re}(\mathbf{e}_m) \\ -\mathrm{Im}(\mathbf{e}_m) \end{bmatrix} \\ &\quad + j \begin{bmatrix} \mathrm{Re}\left(\mathbf{\Pi}^T(k,:)\right) \\ \mathrm{Im}\left(\mathbf{\Pi}^T(k,:)\right) \end{bmatrix}^T \begin{bmatrix} \mathrm{Im}(\mathbf{e}_m) \\ \mathrm{Re}(\mathbf{e}_m) \end{bmatrix}. \end{aligned} \tag{A.1}$$

Since $\mathbf{e}_m$ is i.i.d. complex Gaussian noise, it is easy to verify that $\Delta b_k^{(m)}$, as a linear combination of $\mathbf{e}_m$, is a complex Gaussian random variable. Hence, the real part

$$\mathrm{Re}\left(\Delta b_k^{(m)}\right) = \left(\begin{bmatrix} \mathrm{Re}\left(\mathbf{\Pi}^T(k,:)\right) \\ \mathrm{Im}\left(\mathbf{\Pi}^T(k,:)\right) \end{bmatrix}^T \begin{bmatrix} \mathrm{Re}(\mathbf{e}_m) \\ -\mathrm{Im}(\mathbf{e}_m) \end{bmatrix} \right)$$

is also Gaussian and its expectation can be obtained as

$$E\left[\mathrm{Re}\left(\Delta b_k^{(m)}\right)\right] = [\mathrm{Re}(\mathbf{\Pi}(k,:))\ \mathrm{Im}(\mathbf{\Pi}(k,:))]\, E \begin{bmatrix} \mathrm{Re}(\mathbf{e}_m) \\ -\mathrm{Im}(\mathbf{e}_m) \end{bmatrix} = 0 \tag{A.2}$$

where we have used the assumptions $E[\mathrm{Re}(\mathbf{e}_m)] = \mathbf{0}$ and $E[\mathrm{Im}(\mathbf{e}_m)] = \mathbf{0}$. The variance of $\mathrm{Re}\left(\Delta b_k^{(m)}\right)$ can then be computed as

$$\begin{aligned} &\mathrm{Var}\left(\mathrm{Re}\left(\Delta b_k^{(m)}\right)\right) \\ &\triangleq = E\left[\left(\mathrm{Re}\left(\Delta b_k^{(m)}\right)\right)^2\right] \\ &= \begin{bmatrix} \mathrm{Re}\left(\mathbf{\Pi}^T(k,:)\right) \\ \mathrm{Im}\left(\mathbf{\Pi}^T(k,:)\right) \end{bmatrix}^T E\left[\begin{bmatrix} \mathrm{Re}(\mathbf{e}_m) \\ -\mathrm{Im}(\mathbf{e}_m) \end{bmatrix} \begin{bmatrix} \mathrm{Re}(\mathbf{e}_m) \\ -\mathrm{Im}(\mathbf{e}_m) \end{bmatrix}^T\right] \begin{bmatrix} \mathrm{Re}\left(\mathbf{\Pi}^T(k,:)\right) \\ \mathrm{Im}\left(\mathbf{\Pi}^T(k,:)\right) \end{bmatrix} \\ &= \frac{\sigma_m^2}{2} \begin{bmatrix} \mathrm{Re}\left(\mathbf{\Pi}^T(k,:)\right) \\ \mathrm{Im}\left(\mathbf{\Pi}^T(k,:)\right) \end{bmatrix}^T \begin{bmatrix} \mathrm{Re}\left(\mathbf{\Pi}^T(k,:)\right) \\ \mathrm{Im}\left(\mathbf{\Pi}^T(k,:)\right) \end{bmatrix} \\ &= \frac{\sigma_m^2}{2}\left(\mathbf{\Pi}(k,:)\mathbf{\Pi}^H(k,:)\right) = \frac{\sigma_m^2}{2N}\left[\mathbf{R}_{xx}^{-1}\right]_{k,k}. \end{aligned} \tag{A.3}$$

Similarly, we can obtain $E\left[\mathrm{Im}\left(\Delta b_k^{(m)}\right)\right]=0$ and $\mathrm{Var}\left(\mathrm{Im}\left(\Delta b_k^{(m)}\right)\right)=(\sigma_m^2/2N)\left[\mathbf{R}_{xx}^{-1}\right]_{k,k}$. Therefore, with the estimated noise variance $\hat{\sigma}_m^2$ in (21), the statistics of the real and imaginary parts of $\Delta b_k^{(m)}$ have been obtained. Finally, we show that $\mathrm{Re}(\Delta b_k^{(m)})$ and $\mathrm{Im}(\Delta b_k^{(m)})$ are uncorrelated, namely,

$$\begin{aligned}
&\mathrm{Cov}\left(\mathrm{Re}\left(\Delta b_k^{(m)}\right),\mathrm{Im}\left(\Delta b_k^{(m)}\right)\right)\\
&\quad\triangleq=E\left[\mathrm{Re}\left(\Delta b_k^{(m)}\right)\mathrm{Im}\left(\Delta b_k^{(m)}\right)\right]\\
&\quad=\begin{bmatrix}\mathrm{Re}\left(\boldsymbol{\Pi}^T(k,:)\right)\\ \mathrm{Im}\left(\boldsymbol{\Pi}^T(k,:)\right)\end{bmatrix}^T E\left[\begin{bmatrix}\mathrm{Re}(\mathbf{e}_m)\\ -\mathrm{Im}(\mathbf{e}_m)\end{bmatrix}\begin{bmatrix}\mathrm{Im}(\mathbf{e}_m)\\ \mathrm{Re}(\mathbf{e}_m)\end{bmatrix}^T\right]\begin{bmatrix}\mathrm{Re}\left(\boldsymbol{\Pi}^T(k,:)\right)\\ \mathrm{Im}\left(\boldsymbol{\Pi}^T(k,:)\right)\end{bmatrix}\\
&\quad=\frac{\sigma_m^2}{2}\begin{bmatrix}\mathrm{Re}\left(\boldsymbol{\Pi}^T(k,:)\right)\\ \mathrm{Im}\left(\boldsymbol{\Pi}^T(k,:)\right)\end{bmatrix}^T\begin{bmatrix}\mathbf{0} & \boldsymbol{I}_N\\ -\boldsymbol{I}_N & \mathbf{0}\end{bmatrix}\begin{bmatrix}\mathrm{Re}\left(\boldsymbol{\Pi}^T(k,:)\right)\\ \mathrm{Im}\left(\boldsymbol{\Pi}^T(k,:)\right)\end{bmatrix}=\frac{\sigma_m^2}{2}0=0.
\end{aligned}\tag{A.4}$$

Appendix B. Statistics of $\Delta a_k^{(m)}$

From (23) and (24), we get

$$\begin{aligned}
\hat{a}_k^{(m)}&\triangleq\frac{\hat{b}_k^{(m)}\hat{b}_k^{(0)*}}{|\gamma_k|^2}=\frac{(\gamma_k e^{-j\phi_k^{(m)}}+\Delta b_k^{(m)})(\gamma_k+\Delta b_k^{(0)})^*}{|\gamma_k|^2}\\
&=\frac{|\gamma_k|^2 e^{-j\phi_k^{(m)}}+\Delta b_k^{(m)}\gamma_k^*+\gamma_k e^{-j\phi_k^{(m)}}\Delta b_k^{(0)*}+\Delta b_k^{(m)}\Delta b_k^{(0)*}}{|\gamma_k|^2}\\
&\approx e^{-j\phi_k^{(m)}}+\frac{\Delta b_k^{(m)}\gamma_k^*+\gamma_k e^{-j\phi_k^{(m)}}\Delta b_k^{(0)*}}{|\gamma_k|^2}\\
&\triangleq e^{-j\phi_k^{(m)}}+\Delta a_k^{(m0)}+\Delta a_k^{(0m)},
\end{aligned}\tag{B.1}$$

where $\Delta a_k^{(0m)}\triangleq\Delta b_k^{(m)}\gamma_k^*/|\gamma_k|^2$ and $\Delta a_k^{(m0)}\triangleq\gamma_k e^{-j\phi_k^{(m)}}\Delta b_k^{(0)*}/|\gamma_k|^2$ are the first-order estimated errors. In obtaining (B.1), the second-order estimated error term has been ignored due to the fact that $\Delta b_k^{(m)}\Delta b_k^{(0)*}/|\gamma_k|^2$ is too small relative to $\hat{a}_k^{(m)}$ to affect the estimated result. By using the results of $\Delta b_k^{(m)}$ in Appendix A, it can easily be verified that both $\Delta a_k^{(0m)}$ and $\Delta a_k^{(m0)}$ are complex Gaussian random variables. The expectations of $\Delta a_k^{(0m)}$ and $\Delta a_k^{(m0)}$ are

$$E\left[\Delta a_k^{(0m)}\right]=\frac{E\left[\Delta b_k^{(m)}\right]}{\gamma_k}=0\quad\text{and}\quad E\left\{\Delta a_k^{(m0)}\right\}=\frac{E\left[\Delta b_k^{(0)*}\right]e^{-j\phi_k^{(m)}}}{\gamma_k^*}=0$$

and thus, the expected value of $\Delta a_m^{(k)}=\Delta a_k^{(0m)}+\Delta a_k^{(m0)}$ is also zero. Moreover, the variances of $\Delta a_k^{(0m)}$ and $\Delta a_k^{(m0)}$ are given by

$$\mathrm{Var}(\Delta a_k^{(0m)})=E\left[\Delta a_k^{(0m)}\Delta a_k^{(0m)*}\right]=\frac{E\left[\Delta b_k^{(m)}\Delta b_k^{(m)*}\right]}{|\gamma_k|^2},$$

$$\mathrm{Var}(\Delta a_k^{(m0)})=E\left[\Delta a_k^{(m0)}\Delta a_k^{(m0)*}\right]=\frac{E\left[\Delta b_k^{(0)}\Delta b_k^{(0)*}\right]}{|\gamma_k|^2}.$$

Therefore, the covariance of $\Delta a_k^{(m)}$ can be obtained as follows:

$$\begin{aligned}
\mathrm{Cov}\left(\Delta a_k^{(p)},\Delta a_k^{(q)}\right)&\triangleq E\left[\left(\Delta a_k^{(0p)}+\Delta a_k^{(p0)}\right)(\Delta a_k^{(0q)}+\Delta a_k^{(q0)})^*\right]\\
&=E\left[\Delta a_k^{(0p)}\Delta a_k^{(0q)*}\right]+E\left[\Delta a_k^{(p0)}\Delta a_k^{(0q)*}\right]\\
&\quad+E\left[\Delta a_k^{(p0)}\Delta a_k^{(q0)*}\right]+E\left[\Delta a_k^{(0p)}\Delta a_k^{(q0)*}\right]\\
&=\frac{E\left[\Delta b_k^{(q)}\Delta b_k^{(p)*}\right]+e^{-j\phi_k^{(m)}}E\left[\Delta b_k^{(q)}\Delta b_k^{(0)*}\right]}{|\gamma_k|^2}\\
&\quad+\frac{e^{j\phi_k^{(m)}}E\left[\Delta b_k^{(0)}\Delta b_k^{(p)*}\right]+E\left[\Delta b_k^{(0)}\Delta b_k^{(0)*}\right]}{|\gamma_k|^2}\\
&=\frac{\mathrm{Var}\left(\Delta b_k^{(0)}\right)+\delta_{p,q}\mathrm{Var}\left(\Delta b_k^{(p)}\right)}{|\gamma_k|^2}\\
&=\frac{(\sigma_0^2+\delta_{p,q}\sigma_p^2)\left[\mathbf{R}_{xx}^{-1}\right]_{k,k}}{N|\gamma_k|^2}
\end{aligned}\tag{B.2}$$

which gives directly the variance of $\Delta a_k^{(m)}$ as $(\sigma_0^2+\sigma_m^2)\left[\mathbf{R}_{xx}^{-1}\right]_{k,k}/N|\gamma_k|^2$.

Appendix C. Derivation of (28)

Since the ambiguities are solved by the techniques suggested in [19,27,28], the estimated value of $\boldsymbol{\Phi}_k$ can be obtained as

$$\hat{\boldsymbol{\Phi}}_k=\left[\hat{\phi}_k^{(1)},\hat{\phi}_k^{(2)},\ldots,\hat{\phi}_k^{(M)}\right]^T\tag{C.1}$$

where $\hat{\phi}_k^{(m)}=-\angle(\hat{a}_k^{(m)})+2\pi l_k^{(m)}=-\angle(a_k^{(m)})+2\pi l_k^{(m)}+\Delta\phi_k^{(m)}$ and $\Delta\phi_k^{(m)}\triangleq\hat{\phi}_k^{(m)}-\phi_k^{(m)}$. Using the first-order Taylor series ex-
pansion, we have

$$e^{-j\phi_k^{(m)}}+\Delta\phi_k^{(m)}\approx e^{-j\phi_k^{(m)}}-j\phi_k^{(m)}e^{-j\phi_k^{(m)}}=e^{-j\phi_k^{(m)}}+\Delta a_k^{(m)},\tag{C.2}$$

where $\Delta\phi_k^{(m)}=j\Delta a_k^{(m)}e^{j\phi_k^{(m)}}$. Considering that the estimated error $\Delta\phi_k^{(m)}$ is a real quantity, we let

$$\Delta\phi_k^{(m)}=\mathrm{Re}\left[j\Delta a_k^{(m)}e^{j\phi_k^{(m)}}\right]=-\mathrm{Re}\left[\Delta a_k^{(m)}\right]\ \sin(\phi_k^{(m)})-\mathrm{Im}\left[\Delta a_k^{(m)}\right]\ \cos(\phi_k^{(m)})\tag{C.3}$$

From Appendices A and B, it is easy to prove that both $\mathrm{Re}\left[\Delta a_k^{(m)}\right]$ and $\mathrm{Im}\left[\Delta a_k^{(m)}\right]$ have an identical distribution $\mathbb{N}(\mathbf{0},\frac{1}{2}\boldsymbol{\Omega}_k)$. Therefore, $\Delta\phi_k^{(m)}$ as a linear combination of $\mathrm{Re}\left[\Delta a_k^{(m)}\right]$ and $\mathrm{Im}\left[\Delta a_k^{(m)}\right]$ is a real-valued Gaussian random variable, and its expectation is

$$\begin{aligned}
E\left[\Delta\phi_k^{(m)}\right]&=E\left[-\mathrm{Re}\left(\Delta a_k^{(m)}\right)\sin(\phi_k^{(m)})-\mathrm{Im}\left(\Delta a_k^{(m)}\right)\cos(\phi_k^{(m)})\right]\\
&=-E\left[\mathrm{Re}\left(\Delta a_k^{(m)}\right)\right]\ \sin(\phi_k^{(m)})-E\left[\mathrm{Im}\left(\Delta a_k^{(m)}\right)\right]\cos(\phi_k^{(m)})=0.
\end{aligned}\tag{C.4}$$

The covariance of $\Delta\phi_k^{(m)}$ can be calculated as

$$\begin{aligned}
\mathrm{Cov}\left(\Delta\phi_k^{(p)},\Delta\phi_k^{(q)}\right)&\triangleq E\left[\Delta\phi_k^{(p)}\Delta\phi_k^{(q)}\right]\\
&=E\left[\mathrm{Re}\left(\Delta a_k^{(p)}\right)\mathrm{Re}\left(\Delta a_k^{(q)}\right)\right]\ \sin\left(\phi_k^{(p)}\right)\sin\left(\phi_k^{(q)}\right)\\
&\quad+E\left[\mathrm{Im}\left(\Delta a_k^{(p)}\right)\mathrm{Im}\left(\Delta a_k^{(q)}\right)\right]\ \cos\left(\phi_k^{(p)}\right)\cos\left(\phi_k^{(q)}\right)\\
&=\tfrac{1}{2}\ \mathrm{Cov}\left(\Delta a_k^{(p)},\Delta a_k^{(q)}\right),
\end{aligned}\tag{C.5}$$

where we have used the fact that $\sin(\phi_k^{(p)})=\sin(\phi_k^{(q)})$ and $\cos(\phi_k^{(p)})=\cos(\phi_k^{(q)})$. Clearly, the variance of $\Delta\phi_k^{(m)}$ is given by

$$\mathrm{Var}\left(\Delta\phi_k^{(m)}\right) \triangleq E\left[\left(\Delta\phi_k^{(m)}\right)^2\right] = \tfrac{1}{2}\mathrm{Var}\left(\Delta a_k^{(m)}\right). \tag{C.6}$$

Appendix D. Statistics of $\hat{\gamma}_k^{(m)}$

We rewrite (30) as

$$\hat{\varphi}_k \triangleq \varphi_k + \Delta\varphi_k = \frac{\mathbf{d}^T\mathbf{\Sigma}_k^{-1}\hat{\mathbf{\Phi}}_k}{\mathbf{d}^T\mathbf{\Sigma}_k^{-1}\mathbf{d}} \triangleq \frac{\mathbf{d}^T\mathbf{\Sigma}_k^{-1}\mathbf{\Phi}_k}{\mathbf{d}^T\mathbf{\Sigma}_k^{-1}\mathbf{d}} + \frac{\mathbf{d}^T\mathbf{\Sigma}_k^{-1}\Delta\mathbf{\Phi}_k}{\mathbf{d}^T\mathbf{\Sigma}_k^{-1}\mathbf{d}}, \tag{D.1}$$

where $\Delta\mathbf{\Phi}_k \triangleq = \left[\Delta\phi_k^{(1)}, \Delta\phi_k^{(2)}, \dots, \Delta\phi_k^{(M)}\right]^T$ and $\Delta\varphi_k = \sum_{m=1}^{M}\omega_m\Delta\phi_k^{(m)}$ with $\omega_m = \left(\mathbf{d}^T\mathbf{\Sigma}_k^{-1}/\mathbf{d}^T\mathbf{\Sigma}_k^{-1}\mathbf{d}\right)_{1,m}$ being the mth element of $\mathbf{d}^T\mathbf{\Sigma}_k^{-1}/\mathbf{d}^T\mathbf{\Sigma}_k^{-1}\mathbf{d}$. From Appendix B and (C.5), we can obtain

$$\begin{aligned} e^{jd_m(\varphi_k+\Delta\varphi_k)} &\approx e^{jd_m\varphi_k} + jd_m\Delta\varphi_k e^{jd_m\varphi_k} \\ &= e^{jd_m\varphi_k} + jd_m e^{jd_m\varphi_k}\sum_{i=1}^{M}\omega_i j\Delta a_k^{(i)} e^{jd_i\varphi_k} \\ &= e^{jd_m\varphi_k} - d_m\sum_{i=1}^{M}\omega_i\Delta a_k^{(i)} e^{j(d_i+d_m)\varphi_k} \end{aligned} \tag{D.2}$$

such that

$$\begin{aligned} \hat{\gamma}_k^{(m)} &= \hat{b}_k^{(m)} e^{jd_m\hat{\varphi}_k} \\ &= \left(\gamma_k e^{-jd_m\varphi_k} + \Delta b_m^{(k)}\right)\left(e^{jd_m\varphi_k} - d_m\sum_{i=1}^{M}\omega_i\Delta a_k^{(i)} e^{j(d_i+d_m)\varphi_k}\right) \\ &\approx \gamma_k + \gamma_k d_m\sum_{i=1}^{M}\omega_i\Delta a_k^{(i)} e^{jd_i\varphi_k} + e^{jd_m\varphi_k}\Delta b_k^{(m)} \\ &= \frac{\gamma_k}{|\gamma_k|^2} d_m\sum_{i=1}^{M}\omega_i\left(\Delta b_k^{(i)}\gamma_k^* + \gamma_k e^{-jd_i\varphi_k}\Delta b_k^{(0)*}\right)e^{jd_i\varphi_k} + \gamma_k + e^{jd_m\varphi_k}\Delta b_k^{(m)} \\ &= (d_m\omega_m+1)e^{jd_m\varphi_k}\Delta b_k^{(m)} + d_m\sum_{i\neq m}^{M}\omega_i\Delta b_k^{(i)} e^{jd_i\varphi_k} \\ &\quad + \gamma_k + \frac{\gamma_k^2}{|\gamma_k|^2} d_m\Delta b_k^{(0)*}\sum_{i=1}^{M}\omega_i \\ &\triangleq \gamma_k + \Delta\gamma_k^{(m)}, \end{aligned} \tag{D.3}$$

where the second or higher order estimation errors are omitted. Note that $\hat{\gamma}_k^{(0)} = \hat{b}_k^{(0)}$. (D.3) shows that is a zero-mean Gaussian random variable, whereby one can obtain its covariance as follows:

$$\begin{aligned} [\mathrm{H}_k]_{p,q} &\triangleq E\left\{\Delta\gamma_k^{(p)}\Delta\gamma_k^{(q)}\right\} \\ &= d_p d_q E\left\{\Delta b_k^{(0)}\Delta b_k^{(0)*}\right\}\left(\sum_{i=1}^{M}\omega_i\right)^2 + (d_p\omega_p+1)d_q\omega_p E\left\{\Delta b_k^{(p)}\Delta b_k^{(p)*}\right\} \\ &\quad + (d_q\omega_q+1)d_p\omega_q E\left\{\Delta b_k^{(q)}\Delta b_k^{(q)*}\right\} + d_p d_q\sum_{i\neq p,q}^{M}\omega_i^2 E\left\{\Delta b_k^{(i)}\Delta b_k^{(i)*}\right\} \\ &= \mathrm{Var}\left(\Delta b_k^{(0)}\right)\left(\sum_{i=1}^{M}\omega_i\right)^2 d_p d_q + \mathrm{Var}\left(\Delta b_k^{(p)}\right)(d_p\omega_p+1)d_q\omega_p \\ &\quad + \mathrm{Var}\left(\Delta b_k^{(q)}\right)(d_q\omega_q+1)d_p\omega_q + d_p d_q\sum_{i\neq p,q}^{M}\omega_i^2\,\mathrm{Var}\left(\Delta b_k^{(i)}\right), \quad p\neq q \end{aligned} \tag{D.4}$$

and it can easily be verified that

$$[\mathbf{H}_k]_{0,q} \triangleq E\left\{\Delta\gamma_k^{(0)}\Delta\gamma_k^{(q)}\right\} = 0 = [\mathbf{H}_k]_{p,0} \quad (p\neq q\neq 0). \tag{D.5}$$

Obviously, the variance of $\Delta\gamma_k^{(m)}$ is given by

$$\begin{aligned} [\mathrm{H}_k]_{m,m} &\triangleq \mathrm{Var}\left(\Delta\gamma_k^{(m)}\right) = E\left\{\Delta\gamma_k^{(m)}\Delta\gamma_k^{(m)*}\right\} \\ &= (d_m\omega_m+1)^2 E\left\{\Delta b_k^{(m)}\Delta b_k^{(m)*}\right\} + d_m^2\sum_{i\neq m}^{M}\omega_i^2 E\left\{\Delta b_k^{(i)}\Delta b_k^{(i)*}\right\} \\ &\quad + E\left\{\Delta b_k^{(0)}\Delta b_k^{(0)*}\right\}\left(d_m\sum_{i=1}^{M}\omega_i\right)^2 \\ &= d_m^2\sum_{i=1}^{M}\omega_i^2\,\mathrm{Var}\left(\Delta b_k^{(i)}\right) + \left(d_m\sum_{i=1}^{M}\omega_i\right)^2\mathrm{Var}\left(\Delta b_k^{(0)}\right) \\ &\quad + (2d_m\omega_m+1)\mathrm{Var}\left(\Delta b_k^{(m)}\right), \end{aligned} \tag{D.6}$$

and $[\mathbf{H}_k]_{0,0} = \mathrm{Var}\left(\Delta b_k^{(0)}\right)$.

References

[1] H.L. Van Trees, Detection, Estimation, and Modulation Theory, Optimum Array Processing, Wiley-Interscience, New York, 2004.
[2] H. Krim, M. Viberg, Two decades of array signal processing research: the parametric approach, IEEE Signal Process. Mag. 13 (4) (1996) 67–94.
[3] T.E. Tuncer, B. Friedlander, Classical and Modern Direction-of-Arrival Estimation, Elsevier Inc., Amsterdam, 2009.
[4] J. Capon, High-resolution frequency–wavenumber spectrum analysis, Proc. IEEE 57 (8) (1969) 1408–1418.
[5] R. Roy, T. Kailath, ESPRIT—estimation of signal parameters via rotational invariance techniques, IEEE Trans. Acoust. Speech Signal Process. 37 (7) (1989) 984–995.
[6] R. Schmidt, Multiple emitter location and signal parameter estimation, IEEE Trans. Antennas Propag. 34 (3) (1986) 276–280.
[7] A.-J. Van Der Veen, E.F. Deprettere, A.L. Swindlehurst, Subspace-based signal analysis using singular value decomposition, Proc. IEEE 81 (9) (1993) 1277–1308.
[8] D. Malioutov, M. Çetin, A.S. Willsky, A sparse signal reconstruction perspective for source localization with sensor arrays, IEEE Trans. Signal Process. 53 (8) (2005) 3010–3022.
[9] J.M. Kim, O.K. Lee, J.C. Ye, Compressive MUSIC: revisiting the link between compressive sensing and array signal processing, IEEE Trans. Inf. Theory 58 (1) (2012) 278–301.
[10] F. Wan, W.-P. Zhu, M.N.S. Swamy, Semi-blind most significant tap detection for sparse channel estimation of OFDM systems, IEEE Trans. Circuits Syst. I Regul. Pap. 57 (3) (2010) 703–713.
[11] A. Khabbazibasmenj, A. Hassanien, S.A. Vorobyov, M.W. Morency, Efficient transmit beamspace design for search-free based DOA estimation in MIMO radar, IEEE Trans. Signal Process. 62 (6) (2014) 1490–1500.
[12] J. Li, R. Compton Jr., Maximum likelihood angle estimation for signals with known waveforms, IEEE Trans. Signal Process. 41 (9) (1993) 2850–2862.
[13] J. Li, B. Halder, P. Stoica, M. Viberg, Computationally efficient angle estimation for signals with known waveforms, IEEE Trans. Signal Process. 43 (9) (1995) 2154–2163.
[14] M. Cedervall, R.L. Moses, Efficient maximum likelihood DOA estimation for signals with known waveforms in the presence of multipath, IEEE Trans. Signal Process. 45 (3) (1997) 808–811.
[15] H. Li, G. Liu, J. Li, Angle estimator for signals with known waveforms, Electron. Lett. 35 (23) (1999) 1992–1994.
[16] L.I. Najjar Atallah, S. Marcos, DOA estimation and association of coherent multipaths by using reference signals, Signal Process. 84 (6) (2004) 981–996.
[17] N. Wang, P. Agathoklis, A. Antoniou, A new DOA estimation technique based on subarray beamforming, IEEE Trans. Signal Process. 54 (9) (2006) 3279–3290.
[18] J.-F. Gu, P. Wei, H.-M. Tai, Fast direction-of-arrival estimation with known waveforms and linear operators, IET Signal Process. 2 (1) (2008) 27–36.
[19] J.-F. Gu, W.-P. Zhu, M.N.S. Swamy, Sparse linear arrays for estimating and tracking DOAs of signals with known waveforms, in:

Proceedings of the International Symposium on Circuits and Systems (ISCAS), 2013.
[20] D. Khan, K.L. Bell, Explicit Ziv–Zakai bound for analysis of doa estimation performance of sparse linear arrays, Signal Process. 93 (12) (2013) 3449–3458.
[21] S. Chatterjee, A.S. Hadi, Regression Analysis by Example, John Wiley and Sons, New York, 2013.
[22] S.M. Kay, Fundamentals of Statistical Signal Processing, Volume I: Estimation Theory, Prentice Hall, Upper Saddle River, 1993.
[23] P.C. Hansen, Rank-Deficient and Discrete Ill-Posed Problems: Numerical Aspects of Linear Inversion, vol. 4, SIAM, 1998.
[24] M.H. Gruber, Regression Estimators: A Comparative Study, JHU Press, Baltimore, 2010.
[25] D.A. Harville, Matrix Algebra from a Statistician's Perspective, Springer, Berlin, 2008.
[26] G.A. Seber, A.J. Lee, Linear Regression Analysis, John Wiley and Sons, New York, 2012.
[27] J.-F. Gu, S.C. Chan, W.-P. Zhu, M.N.S. Swamy, Joint DOA estimation and source signal tracking with Kalman filtering and regularized QRD RLS algorithm, IEEE Trans. Circuits Syst. II: Exp. Briefs 60 (1) (2013) 46–50.
[28] J.-F. Gu, W.-P. Zhu, M.N.S. Swamy, Minimum redundancy linear sparse subarrays for direction of arrival estimation without ambiguity, in: Proceedings of the International Symposium on Circuits and Systems (ISCAS), 2011, pp. 390–393.
[29] W.W. Hager, Updating the inverse of a matrix, SIAM Rev. 31 (2) (1989) 221–239.
[30] T.F. Chan, An improved algorithm for computing the singular value decomposition, ACM Trans. Math. Softw. (TOMS) 8 (1) (1982) 72–83.