Source Enumeration for Large Array Using Shrinkage-Based Detectors With Small Samples

LEI HUANG, Senior Member, IEEE
College of Information Engineering
Shenzhen University
China

CHENG QIAN, Student Member, IEEE
Harbin Institute of Technology Shenzhen Graduate School
Shenzhen, China

HING CHEUNG SO, Fellow, IEEE
City University of Hong Kong
Hong Kong, China

JUN FANG, Member, IEEE
University of Electronic Science and Technology of China
Chengdu, China

It is interesting to determine the number of signals impinging upon a large array with small samples. We tackle this problem by using linear shrinkage coefficients of signal and noise subspaces, ending up with two shrinkage coefficient–based detectors (SCDs) for source enumeration. It is proved that the noise shrinkage coefficients are asymptotically Gaussian distributed as the number of antennas and number of samples tend to infinity at the same rate. Moreover, the noise shrinkage coefficients almost surely converge to one while the signal shrinkage coefficients are almost surely less than one as $m,n \to \infty$ and $m/n \to c$. With these properties, the threshold-like and heuristic SCD algorithms for source number detection are devised. Simulation results are included to illustrate their effectiveness.

Manuscript received August 31, 2013; revised June 16, 2014; released for publication August 18, 2014.

DOI. No. 10.1109/TAES.2014.130579

Refereeing of this contribution was handled by W. Blanding.

The work described in this paper was in part supported by a grant from the Joint Research Scheme sponsored by the Research Grants Council of Hong Kong and the National Natural Science Foundation of China (NSFC; Project No. N_CityU 104/11, 61110229/61161160564), by the NSFC (Grants No. 61222106 and No. 61171187), and by the Shenzhen Kongqie talent program (Grant No. KQC201109020061A).

Authors' addresses: L. Huang, College of Information Engineering, Shenzhen University, Shenzhen, China, E-mail: (dr.lei.huang@ieee.org); C. Qian, Department of Electronic and Information Engineering, Harbin Institute of Technology Shenzhen Graduate School, Shenzhen, China; H. C. So, Department of Electronic Engineering, City University of Hong Kong, Hong Kong, China; J. Fang, Department of Electronic and Information Engineering, University of Electronic Science and Technology of China, Chengdu, China

I. INTRODUCTION

Most high-resolution methods, such as multiple signal classification (MUSIC) [1], estimation of signal parameters via rotational invariance techniques (ESPRIT) [2], method of direction estimation (MODE) [3], and principal-singular-vector utilization for modal analysis (PUMA) [4, 5], for direction-of-arrival (DOA) estimation require knowledge of the source number. However, this number is unknown *a priori* and instead needs to be estimated. As a result, source enumeration has been an active and fundamental topic in array processing [6–9]. Source enumeration algorithms vary from hypothesis testing [10] to information theoretic criterion (ITC) [11–14]. Basically, these methods have been developed for the classical asymptotic situation in which the number of antennas is fixed while the number of samples or snapshots tends to infinity. The Gaussian model well suits this condition, and the likelihood-rooted ITCs are the standard schemes for source number detection because the sample size is large enough. In the general asymptotic scenario, however, the number of antennas m tends to infinity at the same rate as the number of snapshots n, meaning that m can be comparable to or even larger than n, posing a challenge to estimate the source number accurately and efficiently. This general asymptotic case is relevant to real-world applications. As an example, by properly employing the waveform diversity in multiple-input, multiple-output (MIMO) radar [15–17], we can obtain a virtual array with an extended aperture in which the number of antennas is considerably increased, probably close to or even larger than the number of snapshots. However, it has been pointed out in [18] that the general asymptotic case is able to provide a more accurate description for practical scenarios in which the number of snapshots and the number of antennas are finite with comparable values.

To correctly enumerate the source signals using a large antenna array with small samples, several methodologies have been developed, including Akaike's information criterion (AIC) in light of the random matrix theory (RMT) [7], the threshold testing (TT) derived from the RMT [19], and the minimum description length (MDL) based on the linear shrinkage (LS) of noise subspace components [9], which are called the RMT-AIC, RMT-TT, and LS-MDL, respectively. As addressed in [19], although RMT-AIC [7] can enumerate the sources for the general asymptotic region, its consistency cannot be guaranteed. Unlike the RMT-AIC approach, LS-MDL [9] is able to accurately detect the source number and offers consistency in the general asymptotic scenario. Nevertheless, the ITC-like methods have a relatively large detection threshold, thereby requiring a higher signal-to-noise ratio (SNR) than that of the threshold-like enumerators to obtain the same detection performance. As a result, the RMT-TT method [19] becomes an alternative candidate for the general asymptotic region. The RMT-TT method is based on testing the significance of a single

weakest signal sample eigenvalue. Given the noise variance, the RMT-TT approach is able to attain the asymptotic limit of detection. Unfortunately, the noise variance cannot be known *a priori* and instead needs to be estimated. The uncertainty in the noise variance estimation inevitably degrades the performance of the RMT-TT algorithm, which can be confirmed in the simulation results of Section V. Therefore, it is worthwhile to devise an effective testing method that is free of the uncertainty in the estimated noise variance for the general asymptotic region.

In this paper, we address this challenging source enumeration problem by means of the property of the LS coefficients. The LS technique can be used to efficiently improve the estimate of the covariance matrix, which has been widely investigated in the literature [20–22]. We show that the LS coefficients associated with the noise subspace almost surely converge to one while the coefficients corresponding to the signal subspace are almost surely strictly lower than one in the general asymptotic situation. Moreover, the noise shrinkage coefficients turn out to be asymptotically Gaussian distributed, paving the way for accurate computation of the theoretical threshold for source number detection. Using the properties of the LS coefficients, along with the asymptotic Gaussian distribution of the noise shrinkage coefficients, the threshold-like testing and heuristic schemes are derived for accurate source number detection. The threshold-like approach is based on the distribution of the noise shrinkage coefficients to determine the decision threshold, whereas the heuristic method is heuristically developed by using the consistent equality among the noise shrinkage coefficients, which are significantly larger than the signal shrinkage coefficients. We refer to them as shrinkage coefficient–based detectors (SCDs). Because the threshold-based SCD (SCD_{thre}) is able to accurately determine the theoretical threshold without the need for noise variance, its detection performance is not affected by the uncertainty in the noise power estimation. Consequently, the property of the constant false-alarm rate (CFAR) can be achieved. Moreover, the iterative procedure in [19] for estimating the noise variance is avoided, significantly reducing its computational complexity. However, the noise shrinkage coefficients are usually significantly larger than the signal shrinkage coefficients. By means of heuristic testing, the noise shrinkage coefficients can be accurately separated from the signal shrinkage coefficients, leading to a new heuristic shrinkage coefficient–based detector (SCD_{heur}).

The remainder of the paper is organized as follows. The data model is presented in Section II. The motivation of the proposed methods, calculation of the LS coefficients, and SCD approaches are provided in Section III. Simulation results are given in Section IV. Finally, conclusions are drawn in Section V.

Throughout this paper, we use boldface uppercase letters to denote matrices, boldface lowercase letters for column vectors, and lowercase letters for scalar quantities. Superscripts $(\cdot)^T$ and $(\cdot)^H$ represent transpose and conjugate transpose, respectively. In this paper, $\hat{a}$ denotes the estimate of a, $\mathbb{E}\{a\}$ is the expected value of a, and $\text{tr}(\mathbf{A})$ and $\|\mathbf{A}\|$ are the trace and Frobenius norm of $\mathbf{A}$, respectively. Furthermore, $\mathbf{I}_m$ is the $m \times m$ identity matrix and $\mathbf{0}_m$ is the $m \times 1$ zero vector. We use $\mathrm{x} \sim \mathcal{CN}(\boldsymbol{\mu}, \Sigma)$ ($\mathcal{N}(\boldsymbol{\mu}, \Sigma)$) to indicate that x follows a complex (real) Gaussian distribution with mean $\boldsymbol{\mu}$ and covariance matrix Σ.

II. PROBLEM FORMULATION

Consider an array of m antennas receiving d narrowband incoherent source signals $\{s_1(t), \ldots, s_d(t)\}$ from distinct directions $\{\theta_1, \ldots, \theta_d\}$. Assume the sources and the array are in the same plane. The tth snapshot vector of the array output is written as

$$\mathbf{x}_t = \mathbf{A}s_t + \mathbf{w}_t, t = 1, \cdots, n \tag{1}$$

where

$$\begin{aligned}\mathbf{x}_t &= [x_1(t), \cdots, x_m(t)]^T \\ \mathbf{A} &= [\mathbf{a}(\theta_1), \cdots, \mathbf{a}(\theta_d)] \\ \mathbf{s}_t &= [s_1(t), \cdots, s_d(t)]^T \\ \mathbf{w}_t &= [w_1(t), \cdots, w_m(t)]^T.\end{aligned}$$

are the observed snapshot vector, steering matrix, signal vector, and noise vector, respectively. Here, $\mathbf{a}(\theta_i)$, $i = 1, \ldots, d$, is the steering vector, where θ_i is the DOA of the ith source, d is the unknown number of sources, m is the number of antennas with $d < m$, and n is the number of snapshots. Unless stated otherwise, the signals are assumed to be independent and identically distributed (IID) complex Gaussian, i.e., $\mathbf{s}_t \sim \mathcal{CN}(\mathbf{0}_d, \Sigma_s)$, in which $\Sigma_s \triangleq \mathbb{E}[\mathbf{s}_t\mathbf{s}_t^H] \in \mathbb{C}^{d\times d}$ is of full rank. Furthermore, the noise $\mathbf{w}_t$ is assumed to be an IID complex Gaussian vector with a mean of zero and the covariance $\tau\mathbf{I}_m$, i.e., $\mathbf{w}_t \sim \mathcal{CN}(\mathbf{0}_m, \tau\mathbf{I}_m)$, which is independent of $\mathbf{s}_t$.

With the assumptions given earlier, the observed samples can be taken as an IID Gaussian vector, i.e., $\mathbf{x}_t \sim \mathcal{CN}(\mathbf{0}_m, \Sigma)$. Here, Σ as the population covariance matrix, which is calculated as

$$\Sigma = \mathbb{E}[\mathbf{x}_t\mathbf{x}_t^H] = \mathbf{A}\Sigma_s\mathbf{A}^H + \tau\mathbf{I}_m. \tag{2}$$

Recall that the signals are incoherent and $d < m$, meaning that Σ_s is nonsingular and $\mathbf{A}$ is of full column rank. In practice, however, only the sample covariance matrix (SCM) is accessible, which is calculated by $\mathbf{S} = (1/n)\sum_{t=1}^{n}\mathbf{x}_t\mathbf{x}_t^H$. Consequently, our task in this work is to infer the source number d from the noisy observations $\{\mathbf{x}_1, \ldots, \mathbf{x}_n\}$ for $m, n \to \infty$ and $m/n \to c$, with $c \in (0, \infty)$ as a constant number.

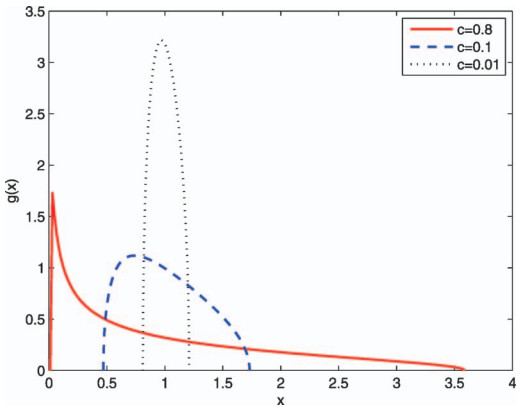


Fig. 1. Marčenko-Pastur distribution for eigenvalues of noise-only SCM with unity noise variance.

III. SOURCE ENUMERATION BASED ON THE LS COEFFICIENT

A. Motivation

Carrying out the eigenvalue decomposition (EVD) on the population covariance matrix Σ, we obtain the population eigenvalues as

$$\lambda_1 \geq \cdots \geq \lambda_d > \lambda_{d+1} = \cdots = \lambda_m = \tau. \tag{3}$$

Let $\mathbf{u}_1, \ldots, \mathbf{u}_m$ be the corresponding population eigenvectors, which can be correctly decomposed into two orthogonal subspaces; i.e., $\{\mathbf{u}_1, \ldots, \mathbf{u}_d\}$ spans the signal subspace, while $\{\mathbf{u}_{d+1}, \ldots, \mathbf{u}_m\}$ spans the noise subspace provided that the source number d is known *a priori.* The subspace-based estimators, such as MUSIC [1], ESPRIT [2], MODE [3] and PUMA [4, 5], can be subsequently applied for high-resolution DOA estimation. However, d is unknown and needs to be estimated. To facilitate this decomposition, we use the presumed source number $k \in [0, m-1]$ instead. With the multiplicity of the smallest eigenvalues, i.e.,

$$\lambda_{d+1} = \cdots = \lambda_m \tag{4}$$

it is easy to determine the source number d. However, the noise population eigenvalues cannot be accessed in practice. Instead, only the noise sample eigenvalues of $\mathbf{S}$ can be obtained, and they satisfy

$$\ell_{d+1} \geq \cdots \geq \ell_m \tag{5}$$

where the inequalities hold with a probability of one. For $n \to \infty$, while m remains unchanged, the maximum likelihood (ML) estimates $\ell_i (i = 1, \cdots, m)$ are the efficient estimates of λ_i $(i = 1, \ldots, m)$. Nevertheless, these sample eigenvalues are usually rather inaccurate for the general asymptotic situation in which $m,n \to \infty$ and $m/n \to c$. It was revealed in [23] that the noise sample eigenvalues asymptotically obey the half-circle distribution or Marčenko-Pastur distribution, which is plotted in Fig. 1, where $g(x)$ is the probability density function of random variable x. Hence, enumerating the sources via the eigenvalues of the SCM cannot lead to satisfactory detection performance in this scenario. One possible choice is to accurately determine the distribution of the sample eigenvalues by means of the Marčenko-Pastur distribution and then use the threshold-like testing for source enumeration, e.g., RMT-TT [19]. However, this method suffers performance degradation because of the uncertainty in the noise variance estimation. Another possible choice is the ITCs, including the AIC and MDL approaches, which are good candidates for adaptive source enumeration. Although the ITC schemes can properly work in the classical asymptotic condition, they may be inefficient in the general asymptotic case because the probability model cannot be accurately approximated by the relatively finite sample size, which is possibly comparable with the number of antennas. It thereby calls for proposals that are able to provide reliable detection of the source number in the general asymptotic scenario at low SNRs. To achieve this goal, we employ the LS technique to accurately compute a sequence of shrinkage coefficients by means of the identity structure of the noise covariance matrix. The shrinkage coefficients are then used to determine the source number.

B. LS Coefficient

Let $\sum_{\mathcal{N}}^{(k)} = \text{diag}(\lambda_{k+1}, \cdots, \lambda_m)$ and $\mathbf{S}_{\mathcal{N}}^{(k)} = \text{diag}(\ell_{k+1}, \cdots, \ell_m)$, which corresponds to the noise subspace covariance matrix under the assumption that k is the presumed source number. It follows from the results of Anderson [24] that $\mathbf{S}_{\mathcal{N}}^{(k)}$ is the unbiased ML estimate of $\Sigma_{\mathcal{N}}^{(k)}$ in the classical asymptotic situation when $n \to \infty$ and m is fixed. However, this unbiased estimate has a large variance and is usually ill conditioned for $m,n \to \infty$ and $m/n \to c$. In contrast to the ML estimate, a structured estimate $u_k \mathbf{I}_{m-k}$ with $\mu_k = 1/(m-k) \sum_{i=k+1}^{m} \lambda_i$ is able to reduce the variance at the expense of increasing the bias. However, $\Sigma_{\mathcal{N}}^{(k)}$ has the structure of the identity matrix for $k \geq d$ but is a diagonal matrix for $k < d$. Consequently, our aim is to derive an optimal estimate of $\Sigma_{\mathcal{N}}^{(k)}$ in the sense of the minimum mean square error (MSE) by means of the LS technique along with the IID Gaussian assumption of observations. To this end, we consider the following constrained minimization of the MSE:

$$\begin{aligned} \min_{\alpha^{(k)}} \quad & g\left(\alpha^{(k)}\right) \triangleq \mathbb{E}\left[\left\|\mathbf{R}^{(k)} - \Sigma_{\mathcal{N}}^{(k)}\right\|^2\right] \\ \text{s.t} \quad & \mathbf{R}^{(k)} = \alpha^{(k)} \mu_k \mathbf{I}_{m-k} + (1 - \alpha^{(k)}) \mathbf{S}_{\mathcal{N}}^{(k)} \end{aligned} \tag{6}$$

where $\alpha^{(k)} \in [0,1]$ denotes the shrinkage coefficient [20]. The LS estimate $\mathbf{R}^{(k)}$ is obtained by shrinking $\mathbf{S}_{\mathcal{N}}^{(k)}$ toward $\mu_k \mathbf{I}_{m-k}$, with the effect of a tradeoff between the bias and the variance. Thus, minimizing the MSE in (6) results in an accurate estimate of $\Sigma_{\mathcal{N}}^{(k)}$ for $m,n \to \infty$ and $m/n \to c$.

Using $\|\mathbf{A} + \mathbf{B}\|^2 = \|\mathbf{A}\|^2 + \|\mathbf{B}\|^2 + 2\text{tr}[\text{Re}(\mathbf{A}\mathbf{B}^H)]$ for $\mathbf{A}, \mathbf{B} \in \mathbb{C}^{m \times m}$, the MSE cost function in (6) can be

easily computed as

$$
\begin{aligned}
g\left(\alpha^{(k)}\right) &= \mathbb{E}\left[\left\|\alpha^{(k)}\mu_k\mathbf{I}_{m-k} + \left(1-\alpha^{(k)}\right)\mathbf{S}_{\mathcal{N}}^{(k)} - \Sigma_{\mathcal{N}}^{(k)}\right\|^2\right] \\
&= \mathbb{E}\left[\left\|\alpha^{(k)}\left(\mu_k\mathbf{I}_{m-k} - \Sigma_{\mathcal{N}}^{(k)}\right) + \left(1-\alpha^{(k)}\right)\left(\mathbf{S}_{\mathcal{N}}^{(k)} - \Sigma_{\mathcal{N}}^{(k)}\right)\right\|^2\right] \\
&= \left(\alpha^{(k)}\right)^2\mathbb{E}\left[\left\|\mu_k\mathbf{I}_{m-k} - \Sigma_{\mathcal{N}}^{(k)}\right\|^2\right] + \left(1-\alpha^{(k)}\right)^2\mathbb{E}\left[\left\|\mathbf{S}_{\mathcal{N}}^{(k)} - \Sigma_{\mathcal{N}}^{(k)}\right\|^2\right] \\
&\quad + \alpha^{(k)}\left(1-\alpha^{(k)}\right)2\mathbb{E}\left[\mathrm{tr}\left[\left(\mu_k\mathbf{I}_{m-k} - \Sigma_{\mathcal{N}}^{(k)}\right)\left(\mathbf{S}_{\mathcal{N}}^{(k)} - \Sigma_{\mathcal{N}}^{(k)}\right)^H\right]\right]. \qquad (7)
\end{aligned}
$$

Setting the derivative of $g(\alpha^{(k)})$ to zero and recalling $\|\mathbf{A}\|^2 = \mathrm{tr}[\mathbf{A}\mathbf{A}^H]$, we obtain the oracle estimate of $\alpha^{(k)}$:

$$
\begin{aligned}
\alpha_O^{(k)} &= \frac{\mathbb{E}\left[\mathrm{tr}\left[\left(\mu_k\mathbf{I}_{m-k} - \mathbf{S}_{\mathcal{N}}^{(k)}\right)\left(\Sigma_{\mathcal{N}}^{(k)} - \mathbf{S}_{\mathcal{N}}^{(k)}\right)^H\right]\right]}{\mathbb{E}\left[\left\|\mathbf{S}_{\mathcal{N}}^{(k)} - \mu_k\mathbf{I}_{m-k}\right\|^2\right]} \\
&= \frac{\mu_k\mathrm{tr}\left(\Sigma_{\mathcal{N}}^{(\mathrm{k})}\right) - \mathbb{E}\left[\mathrm{tr}\left(\mathbf{S}_{\mathcal{N}}^{(\mathrm{k})}\Sigma_{\mathcal{N}}^{(\mathrm{k}),\mathrm{H}}\right)\right] - \mu_{\mathrm{k}}\mathbb{E}\left[\mathrm{tr}\left(\mathbf{S}_{\mathcal{N}}^{(\mathrm{k})}\right)\right] + \mathbb{E}\left[\mathrm{tr}\left(\mathbf{S}_{\mathcal{N}}^{(\mathrm{k})}\mathbf{S}_{\mathcal{N}}^{(\mathrm{k}),\mathrm{H}}\right)\right]}{(m-k)\mu_k^2 - 2\mu_k\mathbb{E}\left[\mathrm{tr}\left(\mathbf{S}_{\mathcal{N}}^{(\mathrm{k})}\right)\right] + \mathbb{E}\left[\mathrm{tr}\left(\mathbf{S}_{\mathcal{N}}^{(\mathrm{k})}\mathbf{S}_{\mathcal{N}}^{(\mathrm{k}),\mathrm{H}}\right)\right]} \\
&= \frac{\mathbb{E}\left[\mathrm{tr}\left(\mathbf{S}_{\mathcal{N}}^{(\mathrm{k})}\mathbf{S}_{\mathcal{N}}^{(\mathrm{k}),\mathrm{H}}\right)\right] - \mathrm{tr}\left(\Sigma_{\mathcal{N}}^{(\mathrm{k})}\Sigma_{\mathcal{N}}^{(\mathrm{k}),\mathrm{H}}\right)}{\mathbb{E}\left[\mathrm{tr}\left(\mathbf{S}_{\mathcal{N}}^{(\mathrm{k})}\mathbf{S}_{\mathcal{N}}^{(\mathrm{k}),\mathrm{H}}\right)\right] - (\mathrm{m}-\mathrm{k})\mu_{\mathrm{k}}^2} \\
&= \frac{\mathbb{E}\left[\frac{1}{m-k}\sum_{i=k+1}^{m}\ell_i^2\right] - \frac{1}{m-k}\sum_{i=k+1}^{m}\lambda_i^2}{\mathbb{E}\left[\frac{1}{m-k}\sum_{i=k+1}^{m}\ell_i^2\right] - \left(\frac{\sum_{i=k+1}^{m}\lambda_i}{m-k}\right)^2}. \qquad (8)
\end{aligned}
$$

For the Gaussian observations, it follows from [25] that

$$
\frac{1}{m-k}\sum_{i=k+1}^{m}\lambda_i = \mathbb{E}\left[\frac{1}{m-k}\sum_{i=k+1}^{m}\ell_i\right] \qquad (9)
$$

$$
\begin{aligned}
\frac{1}{m-k}\sum_{i=k+1}^{m}\lambda_i^2 = \frac{1}{n+1}\Bigg(&n\mathbb{E}\left[\frac{1}{m-k}\sum_{i=k+1}^{m}\ell_i^2\right] \\
&- (m-k)\left(\mathbb{E}\left[\frac{1}{m-k}\sum_{i=k+1}^{m}\ell_i\right]\right)^2\Bigg). \qquad (10)
\end{aligned}
$$

Substituting (9) and (10) into (8) yields

$$
\alpha_O^{(k)} = \frac{\mathbb{E}\left[\frac{1}{m-k}\sum_{i=k+1}^{m}\ell_i^2\right] + (m-k)\left(\mathbb{E}\left[\frac{1}{m-k}\sum_{i=k+1}^{m}\ell_i\right]\right)^2}{(n+1)\left(\mathbb{E}\left[\frac{1}{m-k}\sum_{i=k+1}^{m}\ell_i^2\right] - \left(\mathbb{E}\left[\frac{\sum_{i=k+1}^{m}\ell_i}{m-k}\right]\right)^2\right)}. \qquad (11)
$$

An important observation herein is that when the presumed source number is equal to or larger than the true source number, i.e., $k \geq d$, all presumed noise sample eigenvalues $\{\ell_i\}_{i=k+1}^{m}$ are the true noise sample eigenvalues. As $m,n \to \infty$ and $m/n \to c$, the first- and second-order moments of the true noise sample eigenvalues have the following convergence:

$$
\frac{1}{m-k}\sum_{i=k+1}^{m}\ell_i \xrightarrow{\text{m.s}} \mathbb{E}\left[\frac{1}{m-k}\sum_{i=k+1}^{m}\ell_i\right] \qquad (12)
$$

$$
\frac{1}{m-k}\sum_{i=k+1}^{m}\ell_i^2 \xrightarrow{\text{m.s.}} \mathbb{E}\left[\frac{1}{m-k}\sum_{i=k+1}^{m}\ell_i^2\right] \qquad (13)
$$

where $\xrightarrow{\text{m.s.}}$ means convergence in the mean square. The proofs of (12) and (13) are provided in the appendix. Substituting (12) and (13) into (11) yields the consistent estimate of $\hat{\alpha}_O^{(k)}$:

$$
\hat{\alpha}_c^{(k)} = \frac{\frac{1}{m-k}\sum_{i=k+1}^{m}\ell_i^2 + (m-k)\left(\frac{1}{m-k}\sum_{i=k+1}^{m}\ell_i\right)^2}{(n+1)\left(\frac{1}{m-k}\sum_{i=k+1}^{m}\ell_i^2 - \left(\frac{\sum_{i=k+1}^{m}\ell_i}{m-k}\right)^2\right)}. \qquad (14)
$$

Because $\hat{\alpha}_c^{(k)}$ can be larger than one, we use $\hat{\alpha}^{(k)} = \min(\hat{\alpha}_c^{(k)}, 1)$ [20] rather than $\hat{\alpha}_c^{(k)}$ as the estimated shrinkage coefficient. The shrinkage coefficient has a different behavior from that of $k < d$ because the signal sample eigenvalues, i.e., $\ell_k, \cdots, \ell_d$, do not have limiting values in (12) and (13), thereby providing a good indicator for source enumeration. Compared with the LS-MDL

approach [9], which employs LS to improve the estimate of the eigenvalues, the methods devised in this paper rely on the distribution of the shrinkage coefficient and the difference between the signal and noise shrinkage coefficients. Because this distribution can be accurately determined and the gap between the signal and noise shrinkage coefficients is usually quite large, the proposed source enumerators are able to provide the superior performance.

C. Threshold Testing for Source Enumeration

Let $\mathcal{H}_0$ and $\mathcal{H}_1$ be the hypotheses of at most k sources and at least $k+1$ sources, respectively. For hypothesis $\mathcal{H}_0$, $\ell_{k+1}, \cdots, \ell_m$, are the noise sample eigenvalues. Moreover, it follows from (14) that $\hat{\alpha}^{(k)}(k = d, \cdots, m-1)$ corresponds to the noise shrinkage coefficient because it is determined by the true noise eigenvalues, whereas $\hat{\alpha}^{(k)}(k = 0, \cdots, d-1)$ corresponds to the signal shrinkage coefficient because it is dominated by the signal eigenvalues. By determining the fluctuation of $\hat{\alpha}^{(k)}$ under $\mathcal{H}_0$, we can calculate the decision threshold according to a prescribed probability of false alarm. Moreover, $\hat{\alpha}^{(k)}$ under $\mathcal{H}_1$ is smaller than that under $\mathcal{H}_0$. Consequently, source enumeration amounts to making the following decision:

$$\hat{\alpha}^{(k)} \underset{\mathcal{H}_1}{\overset{\mathcal{H}_0}{\gtrless}} \eta_k, k = 0, \cdots, m-1 \tag{15}$$

where η_k is the prescribed threshold, which needs to be determined and must be independent of the noise variance so that the CFAR property can be achieved. To correctly calculate the threshold for source number detection, it is important to determine the distribution of the shrinkage coefficients for the situation of $k \geq d$, which corresponds to the null hypothesis $\mathcal{H}_0$. To this end, we need the following results.

PROPOSITION 1 The shrinkage coefficient $\hat{\alpha}^{(k)}$ with $k \geq d$ asymptotically obeys the Gaussian distribution as $m,n \to \infty$ and $m/n \to c$. That is,

$$\hat{\alpha}^{(k)} \xrightarrow{\mathcal{D}} \mathcal{N}(\nu_k, \sigma_k^2) \tag{16}$$

where $\xrightarrow{\mathcal{D}}$ denotes convergence in distribution and

$$\nu_k = \frac{c + (m-k+1)}{c(n+1)} \tag{17}$$

$$\sigma_k^2 = \frac{2(m-k+1)^2}{c^2(n+1)^2(m-k)^2}. \tag{18}$$

PROOF Setting $x = \frac{1}{m-k}\sum_{i=k+1}^{m} \ell_i$, $y = \frac{1}{m-k}\sum_{i=k+1}^{m} \ell_i^2$, and $z_k = f(x, y) \triangleq \hat{\alpha}^{(k)}$, it follows from (14) that

$$f(x, y) = \frac{y + (m-k)x^2}{(n+1)\left(y - x^2\right)}. \tag{19}$$

To determine the mean and variance of z_k, we resort to the delta method. In particular, by means of the Taylor series expansion of $f(x, y)$ around $f(x_0, y_0)$, with $x_0 = \mathbb{E}[x]$ and $y_0 = \mathbb{E}[y]$, we obtain

$$\begin{aligned} f(x, y) = {} & f(x_0, y_0) + \partial f_x(x_0, y_0)(x - x_0) \\ & + \partial f_y(x_0, y_0)(y - y_0) + \cdots \end{aligned} \tag{20}$$

where $\partial f_x(x_0, y_0)$ denotes the partial derivative with respect to x. It follows from Lemma 1 in the appendix that $x_0 = \tau$ and $y_0 = \tau^2(1+c)$. We then easily obtain $\mathbf{v} \triangleq [x - x_0, y - y_0]^T = [x - \tau, y - \tau^2(1+c)]^T$ and approximate (20) as

$$f(x, y) \approx \nu_k + \boldsymbol{\nabla}_k^T \mathbf{v} \tag{21}$$

where

$$\nu_k \triangleq f(x_0, y_0) = \frac{y_0 + (m-k)x_0^2}{(n+1)\left(y_0 - x_0^2\right)} = \frac{c + (m-k+1)}{c(n+1)} \tag{22}$$

$$\begin{aligned} \boldsymbol{\nabla}_k &\triangleq \left[\partial f_x(x_0, y_0), \partial f_y(x_0, y_0)\right]^T \\ &= \frac{m-k+1}{(n+1)c^2}\left[\frac{2(1+c)}{\tau}, -\frac{1}{\tau^2}\right]^T. \end{aligned} \tag{23}$$

As $m,n \to \infty$ and $m/n \to c$, substituting $z_k = f(x, y)$ and $\nu_k = f(x_0, y_0)$ into (21) and recalling that z_k has the Gaussian distribution in (16), we have

$$z_k - f(x_0, y_0) \xrightarrow{\mathcal{D}} \mathcal{N}\left(0, \sigma_k^2\right) \tag{24}$$

where

$$\sigma_k^2 = \frac{\boldsymbol{\nabla}_k^T \mathbf{D} \boldsymbol{\nabla}_k}{(m-k)^2} \tag{25}$$

and

$$\mathbf{D} = \begin{bmatrix} \tau^2 c & 2\tau^3 c(1+c) \\ 2\tau^3 c(1+c) & 2\tau^4 c(2c^2+5c+2) \end{bmatrix}. \tag{26}$$

Substituting (23) and (26) into (25) leads to (18). This completes the proof of Proposition 1.

Given a false-alarm level, say, ε, it follows from Proposition 1 that

$$\begin{aligned} \text{Prob}\left(\mathcal{H}_1 | \mathcal{H}_0\right) &= \text{Prob}\left(\frac{\hat{\alpha}^{(k)} - \nu_k}{\sigma_k} < \tilde{\eta}_k \middle| \mathcal{H}_0\right) \\ &= 1 - Q(\tilde{\eta}_k) \triangleq \varepsilon \end{aligned} \tag{27}$$

where $\tilde{\eta}_k$ represents the threshold for the Gaussian distribution and $Q(\tilde{\eta}_k) = \int_{\tilde{\eta}_k}^{\infty} 1/\sqrt{2\pi}\exp(-t^2/2)dt$. As a result, the threshold for the shrinkage coefficient turns out to be

$$\begin{aligned} \eta_k &= Q^{-1}(1-\varepsilon)\sigma_k + \nu_k \\ &= \frac{\sqrt{2}(m-k+1)}{c(n+1)(m-k)}Q^{-1}(1-\varepsilon) + \frac{c + (m-k+1)}{c(n+1)}. \end{aligned} \tag{28}$$

It is shown in [7, 26] that the limiting parameter $c = \lim_{m,n \to \infty} m/n$ can be replaced by $c_m \triangleq m/n$ in

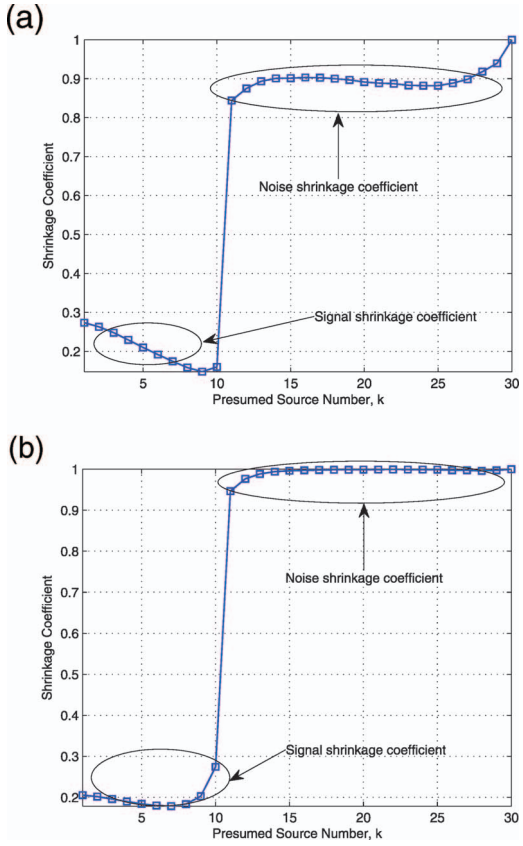


Fig. 2. Shrinkage coefficient versus presumed source number, showing 10 sources with DOAs: [0,6.67, 13.33, 20.00, 26.67, 33.33, 40.00, 46.67, 53.33, 60.00]°. (a) $m = 30$, $n = 50$, $d = 10$, and SNR $= 0$ dB. (b) $m = 30$, $n = 100$, $d = 10$, and SNR $= -5$ dB.

practical applications. As a result, the SCD_{thre} approach can be used to obtain the source number estimate, i.e.,

$$\hat{d} = \max_{k=0,\cdots,\bar{m}-1} \{i_k\} \tag{29}$$

where $\bar{m} = \min(m, n)$ and

$$i_k = \begin{cases} k & \hat{\alpha}^{(k)} < \eta_k \\ 0 & \hat{\alpha}^{(k)} \geq \eta_k. \end{cases} \tag{30}$$

D. Heuristic Approach to Source Enumeration

Although the SCD_{thre} method is able to accurately determine the theoretical threshold for source number detection, it cannot efficiently employ the information inherent in the distance between the signal and noise shrinkage coefficients. The noise shrinkage coefficients are usually separated by a large gap from the signal shrinkage coefficients, as illustrated in Fig. 2. An intuition can be attained from (6), which agrees well with the results in Fig. 2. That is, the signal shrinkage coefficient is a small number close to zero, while the noise shrinkage coefficient is close to one. In particular, for $k < d$, the unknown true covariance matrix $\mathbf{\Sigma}_{\mathcal{N}}^{(k)}$ is diagonal but not identical. To minimize the MSE in (6), we need to assign a weight for the identity matrix that is as small as possible, thereby indicating that $\alpha^{(k)}$ should be a small number. In contrast, for $k \geq d$, $\mathbf{\Sigma}_{\mathcal{N}}^{(k)}$ has the identity structure, calling for a large weight for $\mu_k \mathbf{I}_{m-k}$ in (6). This in turn implies that $\alpha^{(k)}$ approaches one. Consequently, using this gap between the signal and noise shrinkage coefficients, we can significantly improve the detection performance.

PROPOSITION 2 As $m,n \to \infty$ and $m/n \to c$, the noise shrinkage coefficients $\hat{\alpha}^{(k)}(k = d, \cdots, m-1)$ tend to one almost surely:

$$\hat{\alpha}^{(k)} = 1, k = d, \cdots, m-1 \tag{31}$$

whereas the signal shrinkage coefficients $\hat{\alpha}^{(k)}(k = 0, \cdots, d-1)$ are almost surely smaller than one in large SNR region $(\ell_i/\tau \gg 1, i = 1, \cdots, d)$ and strictly lower than one in small SNR region $(\ell_i/\tau \ll m, i = 1, \cdots, d)$, i.e.,

$$\hat{\alpha}^{(k)} \begin{cases} \ll 1, \ell_i/\tau \gg 1, i = 1, \cdots, d \\ < 1, \ell_i/\tau \ll m, i = 1, \cdots, d. \end{cases} \tag{32}$$

PROOF It follows from (14) that the shrinkage coefficient can be rewritten as

$$\begin{aligned} \hat{\alpha}^{(k)} &= \frac{\frac{\frac{1}{m-k}\sum_{i=k+1}^{m} \ell_i^2}{\left(\frac{1}{m-k}\sum_{i=k+1}^{m} \ell_i\right)^2} + (m-k)}{(n+1)\left(\frac{\frac{1}{m-k}\sum_{i=k+1}^{m} \ell_i^2}{\left(\frac{1}{m-k}\sum_{i=k+1}^{m} \ell_i\right)^2} - 1\right)} \\ &\triangleq \frac{\gamma_k + (m-k)}{(n+1)(\gamma_k - 1)} \end{aligned} \tag{33}$$

where

$$\gamma_k = \frac{\frac{1}{m-k}\sum_{i=k+1}^{m} \ell_i^2}{\left(\frac{1}{m-k}\sum_{i=k+1}^{m} \ell_i\right)^2}. \tag{34}$$

For $k \geq d$ and $m,n \to \infty$ with $m/n \to c$, it readily follows from (34) that

$$\gamma_k = \frac{\tau^2(1+c)}{\tau^2} = 1 + c \tag{35}$$

which, when substituted into (33), leads to

$$\hat{\alpha}^{(k)} = \frac{1 + c + (m-k)}{(n+1)c} \to 1 \tag{36}$$

for $k = d, \ldots, m-1$, along with $m,n \to \infty$ and $m/n \to c$.

Now consider the signal shrinkage coefficient that corresponds to the situation of $k < d$. As $m,n \to \infty$ and $m/n \to c$, it follows from (33) that

$$\hat{\alpha}^{(k)} = \frac{1}{n+1} + \frac{\frac{m-k+1}{n+1}}{\gamma_k - 1} \xrightarrow{\text{a.s.}} \frac{c}{\gamma_k - 1}. \tag{37}$$

where $\xrightarrow{\text{a.s.}}$ means almost surely convergence. However, as $m,n \to \infty$ with $m/n \to c$, it is indicated in [7]

that for $i = 1, \ldots, d$,

$$\ell_i \xrightarrow{\text{a.s}} \begin{cases} \lambda_i(1+\tau c/(\lambda_i-\tau)) \triangleq \tilde{\lambda}_i, & \text{if } \lambda_i > \tau(1+\sqrt{c}) \\ \tau(1+\sqrt{c})^2 \triangleq \lambda_{\text{DET}}, & \text{if } \lambda_i \le \tau(1+\sqrt{c}). \end{cases} \tag{38}$$

At high SNRs, it is clear that $\lambda_i > \tau(1+\sqrt{c})$, $i=1, \cdots, d$. For $m,n \to \infty$ and $m/n \to c$, it is easy to obtain

$$\begin{aligned} \gamma_k &= (m-k)\frac{\sum_{i=k+1}^{d} \ell_i^2 + \sum_{i=d+1}^{m} \ell_i^2}{\left(\sum_{i=k+1}^{d} \ell_i + \sum_{i=d+1}^{m} \ell_i\right)^2} \\ &\xrightarrow{\text{a.s.}} m\frac{\sum_{i=k+1}^{d} \tilde{\lambda}_i^2 + m\tau^2(1+c)}{\left(\sum_{i=k+1}^{d} \tilde{\lambda}_i + m\tau\right)^2} \end{aligned} \tag{39a}$$

$$= m\frac{\sum_{i=k+1}^{d} \rho_i^2 + m(1+c)}{\left(\sum_{i=k+1}^{d} \rho_i + m\right)^2} \tag{39b}$$

where $\rho_i = \tilde{\lambda}_i/\tau$ corresponds to the SNR. For the high SNR case, i.e., $\ell_i/\tau \to \rho_i \gg 1$, $\sum_{i=k+1}^{d} \rho_i$ cannot be ignored when compared with m. This indicates that $\sum_{i=k+1}^{d} \rho_i = \mathcal{O}(m)$ and $\sum_{i=k+1}^{d} \rho_i^2 = \mathcal{O}(m^2)$, which, when inserted into (39b), lead to

$$\gamma_k \xrightarrow{\text{a.s.}} m\frac{c_1 m^2 + m(1+c)}{(c_2 m + m)^2} \to \frac{c_1}{(c_2+1)^2} m \gg 1 \tag{40}$$

with c_1 and c_2 as constant numbers. Substituting (40) into (37) yields

$$\hat{\alpha}^{(k)} \ll 1. \tag{41}$$

For the low SNR case, we have $\ell_{k+1}/\tau \ll m$ and $\ell_d \xrightarrow{\text{a.s.}} \lambda_{\text{DET}}$. Recalling that $\ell_{k+1} \ge \cdots \ge \ell_d$, it follows from (38) and (39a) that

$$\begin{aligned} \gamma_k &\xrightarrow{\text{a.s.}} m\frac{\sum_{i=k+1}^{d} \ell_i^2 + m\tau^2(1+c)}{\left(\sum_{i=k+1}^{d} \ell_i + m\tau\right)^2} \\ &> m\frac{(d-k)\ell_d^2 + m\tau^2(1+c)}{((d-k)\ell_{k+1} + m\tau)^2} \\ &\xrightarrow{\text{a.s.}} 1+c. \end{aligned} \tag{42}$$

Substituting (42) into (37) leads to

$$\hat{\alpha}^{(k)} < 1, \text{ (a.s.)}. \tag{43}$$

This completes the proof of Proposition 2.

It follows from Proposition 2 that when k varies from 0 to $m-1$, the noise shrinkage coefficients asymptotically converge to one while the signal shrinkage coefficients are asymptotically smaller than one at large SNRs and strictly smaller than one at small SNRs. As a result, similar to [27, 28], the source number can be determined by SCD_{heur}:

$$\hat{d} = \max_{k=0,\cdots,\bar{m}-1} \{j_k\} \tag{44}$$

where

$$j_k = \begin{cases} k & \hat{\alpha}^{(k)} < \dfrac{D_n}{\bar{m}} \sum_{i=1}^{\bar{m}} \hat{\alpha}^{(i)} \\ 0 & \hat{\alpha}^{(k)} \ge \dfrac{D_n}{\bar{m}} \sum_{i=1}^{\bar{m}} \hat{\alpha}^{(i)} \end{cases} \tag{45}$$

and $D_n = \kappa/\log(n)$. Here, κ is a constant number close to $\log(n)$, which is the natural logarithm of n.

REMARK Both the SCD_{heur} and SCD_{thre} algorithms involve the computations of the SCM and EVD, which require around $\mathcal{O}(m^2n + m^3)$ flops. Moreover, it is indicated in (30) and (45) that the constructions of the source enumerators for SCD_{heur} and SCD_{thre} only rely on the shrinkage coefficient, which turns out to be a simple function of the sample eigenvalues, as shown in (14). Hence, the computational complexity in source enumerator construction can be ignored and SCD_{heur} or SCD_{thre} needs merely about $\mathcal{O}(m^2n + m^3)$ flops.

IV. SIMULATION RESULTS

A. Accuracy of the Theoretical Threshold for SCD_{thre}

In this simulation, we evaluate the accuracy of the theoretical threshold for SCD_{thre}. For the purpose of comparison, the empirical results for RMT-TT [19] are presented as well. To numerically determine the true threshold used as the performance benchmark, we repeatedly calculate $\hat{\alpha}^{(d)}$ by assuming that the true source number d is known based on a large number of independent trials. The true threshold is the one corresponding to a desired probability of false alarm: $P_{\text{fa}} = \varepsilon$. In this experiment, 50 000 independent runs have been performed, yielding 50 000 $\{\hat{\alpha}^{(d)}\}$. After arranging them in increasing order, the $(P_{\text{fa}} \cdot 50\,000)$th coefficient is selected as the true threshold. The theoretical threshold for SCD_{thre} is obtained by using (28), while that for RMT-TT is yielded by employing the right-hand side within the braces of (19), along with (22) in [19]. The relative error is defined as Error $\triangleq (|\eta_{\text{theo}} - \eta_{\text{simu}}|)/\eta_{\text{simu}} \times 100\%$, where η_{theo} and η_{simu} are the theoretical and simulated thresholds, respectively. The noise is the IID Gaussian process with zero mean and unknown variance τ.

The empirical results for the threshold comparison are given in Table I, where the number of antennas is 30, the number of snapshots is 40, and the DOAs of five sources are $[1.2, 5.8, 11.7, 17, 20.3]^\circ$. It is seen that SCD_{thre} is superior to RMT-TT in terms of accuracy in the theoretical threshold determination. When the number of snapshots increases to 80 while other parameters remain unchanged, the SCD_{thre} algorithm is more accurate than RMT-TT, as depicted in Table II. That is, SCD_{thre} yields a relative error as small as about 2%, while RMT-TT has an error of around 10% at $P_{\text{fa}} = 0.1$. When the numbers of antennas

TABLE I
Thresholds for SCD_{thre} and RMT-TT at $m = 30$, $n = 40$, $c = 0.75$, SNR $= 0$ dB, and DOAs of [1.2, 5.8, 11.7, 17, 20.3]°

Method	SCD_{thre}				RMT-TT			
P_{fa}	0.1	0.05	0.01	0.001	0.1	0.05	0.01	0.001
η_{theo}	0.8086	0.7912	0.7586	0.7221	3.3788	3.4283	3.5324	3.6646
η_{simu}	0.8540	0.8369	0.8066	0.7712	2.8454	2.9131	3.0500	3.2054
Error (%)	5.3102	5.4548	5.9439	6.3672	18.7452	17.6827	15.8158	14.3229

TABLE II
Thresholds for SCD_{thre} and RMT-TT at $m = 60$, $n = 80$, $c = 0.75$, SNR $= 0$ dB, and DOAs of [1.2, 5.8, 11.7, 17, 20.3]°

Method	SCD_{thre}				RMT-TT			
P_{fa}	0.1	0.05	0.01	0.001	0.1	0.05	0.01	0.001
η_{theo}	0.9038	0.8952	0.8790	0.8609	3.4746	3.5067	3.5743	3.6601
η_{simu}	0.9256	0.9170	0.9007	0.8833	3.1495	3.1954	3.2898	3.3964
Error (%)	2.3557	2.3781	2.4111	2.5401	10.3207	9.7411	8.6501	7.7652

TABLE III
Thresholds for SCD_{thre} and RMT-TT at $m = 90$, $n = 120$, $c = 0.75$, SNR $= 0$ dB, and DOAs of [1.2, 5.8, 11.7, 17, 20.3]°

Method	SCD_{thre}				RMT-TT			
P_{fa}	0.1	0.05	0.01	0.001	0.1	0.05	0.01	0.001
η_{theo}	0.9357	0.9300	0.9192	0.9072	3.4733	3.4979	3.5497	3.6153
η_{simu}	0.9498	0.9441	0.9336	0.9224	3.2526	3.2878	3.3588	3.4369
Error (%)	1.4777	1.4957	1.5373	1.6509	6.7865	6.3909	5.6811	5.1910

and snapshots become sufficiently large but at the same rate $c = 0.75$, i.e., $m = 90$ and $n = 120$, the error of SCD_{thre} is less than that of RMT-TT, as illustrated in Table III. This means that SCD_{thre} significantly surpasses RMT-TT in the theoretical threshold computation. RMT-TT is able to attain asymptotic limit of detection, i.e., in equation (11) in [(19)], only when the noise variance τ is known *a priori.* When τ is unknown and needs to be estimated, however, RMT-TT cannot attain this asymptotic limit of detection, thereby indicating that it is no longer optimal in this sense. This is why the RMT-TT algorithm is inferior to the SCD_{thre} method.

B. Effect of κ on SCD_{heur}

Although SCD_{heur} is able to employ the information inherent in the distance between the signal and noise shrinkage coefficients, it highly relies on the user-defined parameter κ. As a result, it is necessary to investigate the effect of κ on the behavior of SCD_{heur}. To this end, we plot the empirical probability of correct detection for SCD_{heur} versus κ under different parameter settings in Fig. 3. All empirical results are obtained from 1000 independent trials. It is observed from Figs. 3a–3c that when the number of sources is $d = 5$, $\log(n)$ is close to the optimal κ no matter whether c_m is larger or smaller than one. The same results can be seen in Fig. 3d–3f, where the number of sources is $d = 10$. Therefore, the user-defined parameter κ can be set as $\log(n)$, although it is not optimal. Indeed, as demonstrated in the following simulation results, SCD_{heur} with $\kappa = \log(n)$ is considerably superior to other methods in detection performance because it efficiently uses the gap information between the signal and noise shrinkage coefficients, which is not shared with the other source enumerators.

C. Detection Performance

Now let us evaluate the detection performance of the SCDs. For comparison, the empirical results of the state-of-the-art algorithms for source enumeration, i.e., RMT-TT [19], Bayesian information criterion (BIC) [8], LS-MDL [9], and MDL [12], are presented. In what follows, the simulated results are based on 1000 independent trials, the probability of false alarm is set as 0.001 for RMT-TT and SCD_{thre}, and $\kappa = \log(n)$ for SCD_{heur}.

The empirical probabilities of correct detection versus SNR are plotted in Fig. 4, where the number of antennas is 30 and the DOAs of five sources are [1.2, 5.8, 11.7, 17, 20.3]°. It is seen in Fig. 4a—where the number of snapshots is 15, which is less than the number of antennas—that the classical MDL and recently devised BIC schemes fail because both of them are developed for the classical asymptotic case; i.e., the number of snapshots is larger than the number of antennas. The SCD_{thre} has similar behavior to that of RMT-TT, and both of them are inferior to the LS-MDL. Because the SCD_{heur} is able to efficiently use the distance information between the signal and noise shrinkage coefficients, it considerably outperforms other enumerators. When the number of snapshots becomes larger than the number of antennas,

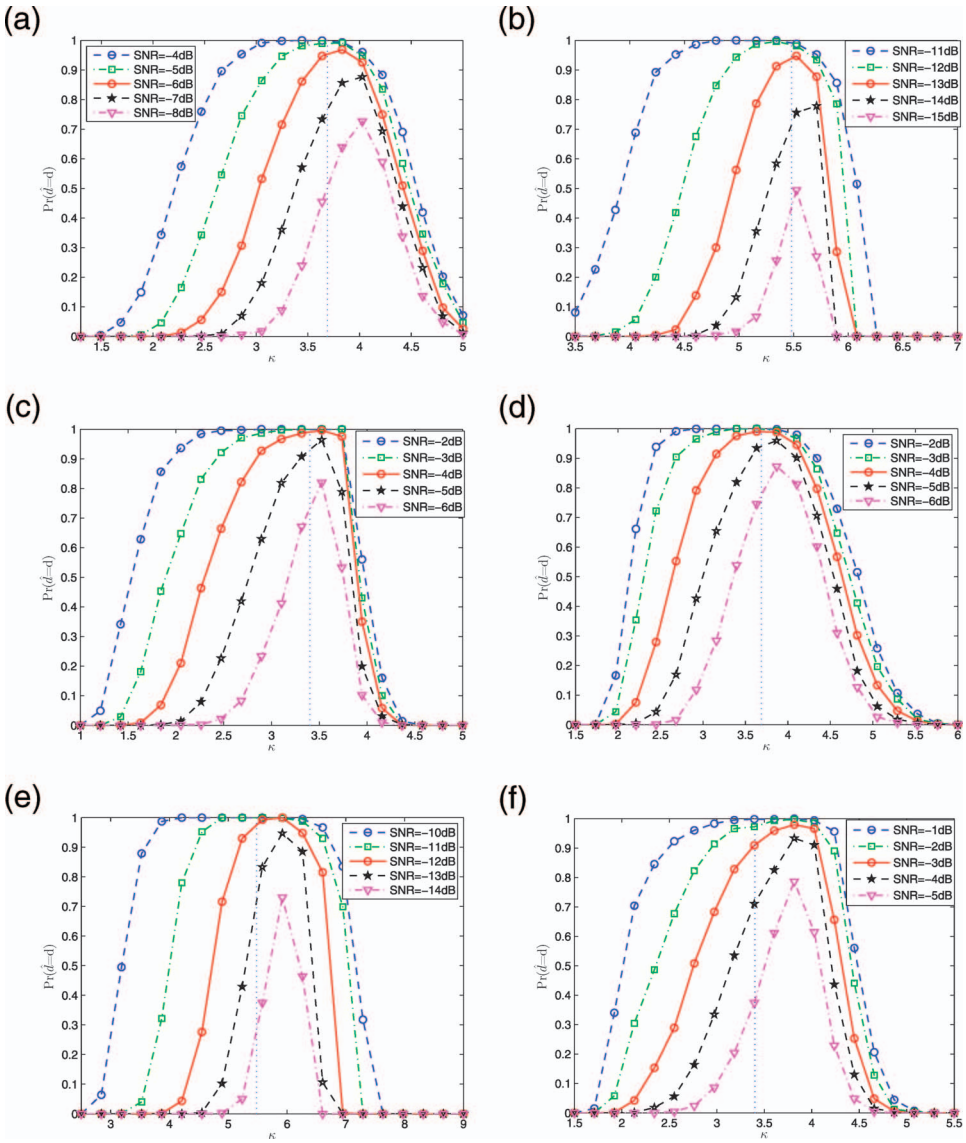


Fig. 3. Probability of correction detection for SCD_{heur} versus κ. Blue dashed line denotes location of $\log(n)$, DOAs are $[1.2, 5.8, 11.7, 17, 20.3]^\circ$ for $d = 5$, and DOAs are $[0, 4.44, 8.89, 13.33, 17.78, 22.22, 26.67, 31.11, 35.56, 40]^\circ$ for $d = 10$. (a) $m = 30$, $n = 40$, and $d = 5$. (b) $m = 30$, $n = 240$, and $d = 5$. (c) $m = 40$, $n = 30$, and $d = 5$. (d) $m = 30$, $n = 40$, and $d = 10$. (e) $m = 30$, $n = 240$, and $d = 10$. (f) $m = 40$, $n = 30$, and $d = 10$.

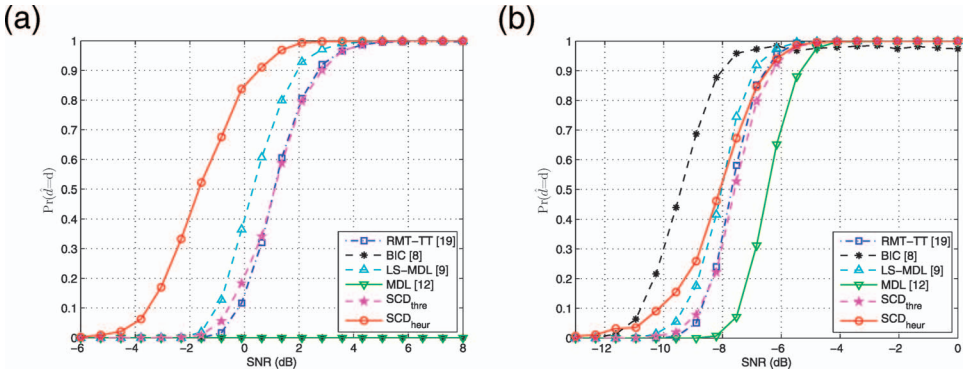


Fig. 4. Probability of correct detection versus SNR. $\kappa = \log(n)$ and DOAs are $[1.2, 5.8, 11.7, 17, 20.3]^\circ$ for $d = 5$. (a) $m = 30$ and $n = 15$. (b) $m = 30$ and $n = 40$.

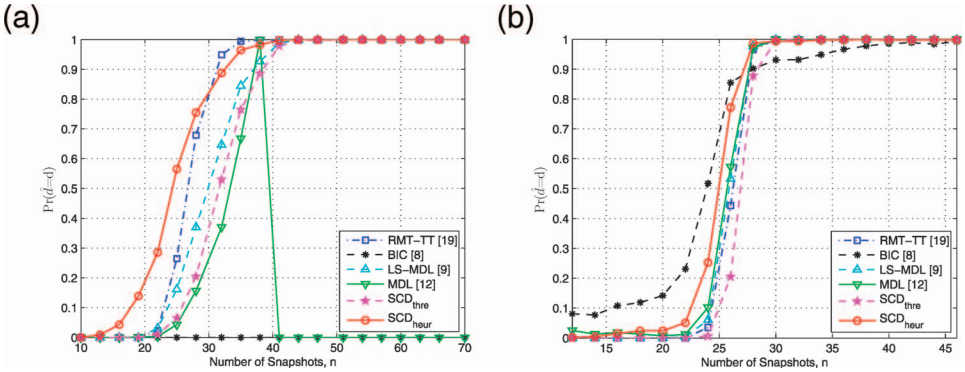


Fig. 5. Probability of correct detection versus number of snapshots. $\kappa = \log(n)$ and DOAs are $[1.2, 5.8, 11.7, 17, 20.3]^\circ$ for $d = 5$. (a) $m/n = 1.6$ and SNR $= -5$ dB. (b) $m/n = 0.6$ and SNR $= 5$ dB.

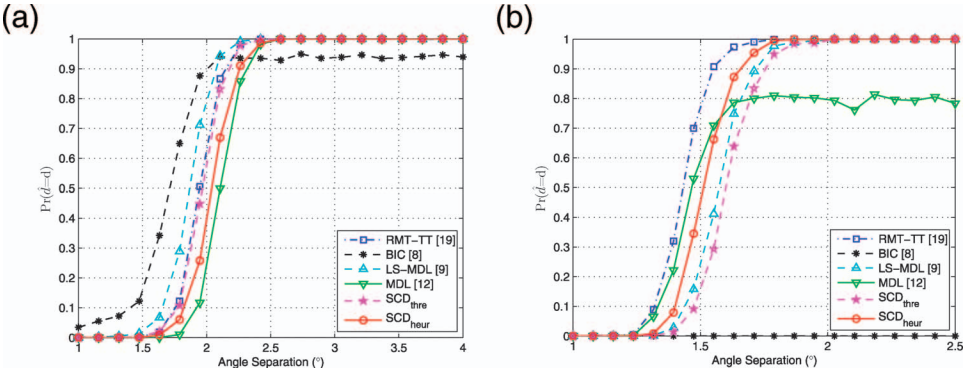


Fig. 6. Probability of correct detection versus angle separation Δ. $[\theta_1, \theta_2, \theta_3] = [0, \Delta, 2\Delta]$ and $\kappa = \log(n)$. (a) $m = 25$, $n = 30$, and SNR $= 3$ dB. (b) $m = 50$, $n = 30$, and SNR $= -3$ dB.

i.e., $m/n = 3/4$, all the investigated methods can properly work. However, although the BIC algorithm can provide accurate detection at small SNRs, it cannot attain the probability of correction detection of one even when the SNR is large enough. As a threshold-like detector, the SCD_{thre} yields similar detection accuracy to that of the RMT-TT approach. The detection performance of SCD_{heur} is comparable with the LS-MDL scheme but superior to the threshold-like and MDL schemes. It is easy to interpret the superiority of SCD_{heur} over LS-MDL, especially when the number of antennas is larger than the number of snapshots. Although the LS-MDL approach is able to use the LS technique to enhance the estimate of the covariance matrix of the noise subspace components, leading to an efficient MDL variant for the general asymptotic regime, it ignores the significant difference between the signal and noise shrinkage coefficients. Unlike the LS-MDL method, the SCD_{heur} scheme is capable of exploiting this gap information for source enumeration, thereby considerably improving the detection accuracy, particularly when the number of antennas is larger than the number of snapshots.

Fig. 5 displays the empirical probability of correct detection versus the number of snapshots. We observe from Fig. 5a that when the number of snapshots is less than the number of antennas, the MDL and BIC methods cannot correctly detect the source number. In this case, the RMT-TT scheme surpasses the SCD_{thre} and LS-MDL approaches but still is not as accurate as the SCD_{heur}. For the situation of $m/n = 3/5$, however, all studied methods are able to properly enumerate the sources. Moreover, except for the BIC scheme, all methods provide similar detection accuracy. The SCD_{heur} is slightly superior to the other methods, while the SCD_{thre} is somewhat inferior to the other schemes. However, the probability of correct detection of the BIC algorithm converges to one more slowly than it does with the other methods, while it has relatively higher accuracy at small numbers of snapshots and antennas.

The correct detection probabilities versus the angle separation are depicted in Fig. 6, where the DOAs of three sources are set as $[0, \Delta, 2\Delta]^\circ$. It is seen in Fig. 6a that when the number of snapshots is larger than the number of antennas, all methods can correctly detect the source number. However, the BIC algorithm fails to attain the probability of correct detection of one even when the angle separation is large enough. The LS-MDL performs the best, while MDL is the worst. The detection accuracy of SCD_{thre} is close to that of RMT-TT, while the SCD_{heur} is slightly inferior to them. When the number of snapshots is less than the number of antennas, nevertheless, the MDL method cannot work properly; neither does the BIC scheme, as demonstrated in Fig. 6b. In this setting, RMT-TT surpasses all other schemes.

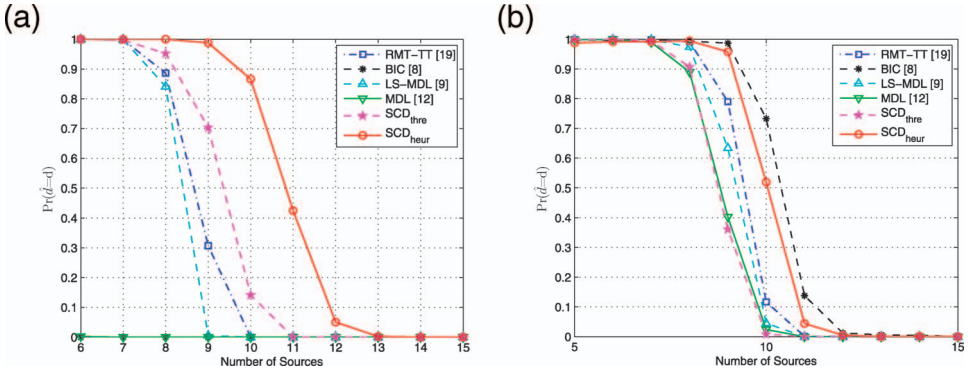


Fig. 7. Probability of correct detection versus number of sources. $\kappa = \log(n)$ and DOAs are $[\theta_1, \ldots, \theta_d] = \boldsymbol{\theta}(1{:}\, d)$. (a) $m = 30$, $n = 20$, and SNR $= 3$ dB. (b) $m = 30$, $n = 60$, and SNR $= -5$ dB.

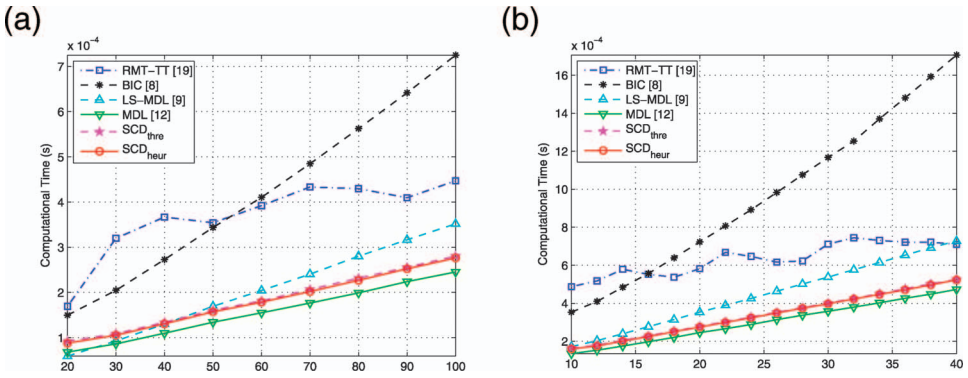


Fig. 8. Computational cost versus number of snapshots. $\kappa = \log(n)$ and DOAs are $[1.2, 5.8, 11.7, 17, 20.3]^\circ$ for $d = 5$. (a) $m/n = 0.1$ and SNR $= 0$ dB. (b) $m/n = 0.5$ and SNR $= 5$ dB.

To examine the identifiability of the studied methods, the empirical probabilities of correct detection versus the source number are plotted in Fig. 7. Here, the DOAs of $d \in [0{,}15]$ sources are set as $[\theta_1, \cdots, \theta_d] = \boldsymbol{\theta}(1 : d)$ with

$$\boldsymbol{\theta} = [\underbrace{0, 3.57, 7.14, 10.71, \cdots, 50}_{15}]^\circ. \tag{46}$$

It is indicated in Fig. 7a that both the SCD_{heur} and SCD_{thre} outperform the other methods in identifiability for $m = 30$ and $n = 20$. When $m = 30$ and $n = 60$, all methods are capable of correctly detecting the source number. The identifiability of SCD_{heur} is still superior to the other schemes but slightly inferior to the BIC algorithm.

The computational costs of various algorithms versus the number of snapshots are depicted in Fig. 8. We observe from Fig. 8 that the proposed SCDs are computationally simpler than the RMT, MDL, LS-MDL, and BIC schemes, especially when both m and n become large. In addition, they require similar a computational cost to that of the MDL approach, because the computational complexity of the MDL scheme also mainly results from the calculations of SCM and EVD. The simulation results agree well with the complexity analysis for the SCD algorithms in Section III.D. The RMT-TT depends on an iterative procedure to estimate the noise variance, and the iterations cannot be terminated until a stopping criterion is reached. Moreover, the BIC scheme is more computationally intensive than the MDL, LS-MDL, SCD_{heur}, and SCD_{thre} because it needs to compute an additional likelihood term for the eigenvalues. As a result, both the RMT-TT and the BIC methods are more computationally demanding than the SCD_{heur}, SCD_{thre}, MDL, and LS-MDL, as indicated in Fig. 8.

V. CONCLUSION

By means of the properties of the LS coefficients, two source enumerators have been devised for the general asymptotic regime, where the number of snapshots tends to infinity at the same rate as the number of antennas. Because the SCD_{thre} does not require the estimated noise variance in determining its theoretical threshold, it is superior to the RMT-TT in robustness against noise uncertainty. Compared with the SCD_{thre}, the SCD_{heur} is able to efficiently exploit the distance information between the signal and noise shrinkage coefficients. As a result, it is more accurate than the other source enumerators, particularly when the number of antennas is larger than the number of snapshots. Furthermore, because of the computational simplicity of the LS coefficients, the SCDs are comparable with the classical MDL scheme and superior to the BIC, LS-MDL, and RMT-TT methods in terms of computational complexity.

APPENDIX. PROOF OF (12) AND (13)

To prove (12) and (13), we need the following results.

Lemma 1: Let $\ell_{k+1} \geq \cdots \geq \ell_m$ be the sample eigenvalues associated with the SCM of the $(m-k) \times n$ IID Gaussian observations with mean zero and variance τ. As $m,n \to \infty$ and $m/n \to c \in (0, \infty)$, we have

$$(m-k)\left(\begin{bmatrix} \frac{1}{m-k}\sum_{i=k+1}^{m} \ell_i \\ \frac{1}{m-k}\sum_{i=k+1}^{m} \ell_i^2 \end{bmatrix} - \begin{bmatrix} \tau \\ \tau^2(1+c) \end{bmatrix}\right) \xrightarrow{\mathcal{D}} \mathcal{N}(\mathbf{0}_2, \mathbf{D}) \tag{47}$$

where $\xrightarrow{\mathcal{D}}$ denotes convergence in distribution and $\mathbf{D}$ is defined in (26).

PROOF: Noting that $m-k \to m$ as $m \to \infty$, Lemma 1 turns out to be Proposition 3.2 in [7], which was proved in [26, 29].

It is indicated in Lemma 1 that $\sum_{i=k+1}^{m} \ell_i/(m-k)$ and $\sum_{i=k+1}^{m} \ell_i^2/(m-k)$ are the unbiased estimates of τ and $\tau^2(1+c)$, respectively. Meanwhile, their variances converge to zero as $m,n \to \infty$ and $m/n \to c \in (0, \infty)$, i.e.,

$$\operatorname{var}\left(\frac{1}{m-k}\sum_{i=k+1}^{m} \ell_i - \tau\right) = \frac{\tau^2 c}{(m-k)^2} \to 0 \tag{48}$$

$$\operatorname{var}\left(\frac{1}{m-k}\sum_{i=k+1}^{m} \ell_i^2 - \tau^2(1+c)\right) = \frac{2\tau^4 c(2c^2+5c+2)}{(m-k)^2} \to 0. \tag{49}$$

This eventually leads to (12) and (13).

REFERENCES

[1] Schmidt, R. O.
Multiple emitter location and signal parameter estimation.
IEEE Transactions on Antennas and Propagation, **34** (1986), 276–280.

[2] Roy, R., Paulraj, A., and Kailath, T.
ESPRIT: A subspace rotation approach to estimation of parameters of cisoids in noise.
IEEE Transactions on Acoustics, Speech and Signal Processing, **34**, 5 (Oct. 1986), 1340–1342.

[3] Stoica, P., and Sharman, K. C.
Novel eigenanalysis method for direction estimation.
Proceedings of the IEEE, **137**, 1 (Feb. 1990), 19–26.

[4] Sun, W., and So, H. C.
Accurate and computationally efficient tensor-based subspace approach for multi-dimensional harmonic retrieval.
IEEE Transactions on Signal Processing, **60**, 10 (Oct. 2012), 5077–5088.

[5] Sun, W., So, H. C., Chan, F. K. W., and Huang, L.
Tensor approach for eigenvector-based multi-dimensional harmonic retrieval.
IEEE Transactions on Signal Processing, **61**, 13 (July 2013), 3378–3388.

[6] Huang, L., Wu, S., and Li, X.
Reduced-rank MDL method for source enumeration in high-resolution array processing.
IEEE Transactions on Signal Processing, **55**, 12 (Dec. 2007), 5658–5667.

[7] Nadakuditi, R. R., and Edelman, A.
Sample eigenvalue based detection of high-dimensional signals in white noise using relatively few samples.
IEEE Transactions on Signal Processing, **56**, 7 (July 2008), 2625–2638.

[8] Lu, M., and Zoubir, A. M.
Generalized Bayesian information criterion for source enumeration in array processing.
IEEE Transactions on Signal Processing, **61**, 6 (Mar. 2013), 1470–1480.

[9] Huang, L., and So, H. C.
Source enumeration via MDL criterion based on linear shrinkage estimation of noise subspace covariance matrix.
IEEE Transactions on Signal Processing, **61**, 19 (July 2013), 1–16.

[10] Williams, D., and Johnson, D.
Using the sphericity test for source detection with narrow-band passive arrays.
IEEE Transactions on Acoustics, Speech and Signal Processing, **38**, 11 (Nov. 2011), 2008–2014.

[11] Akaike, H.
A new look at the statistical model identification.
IEEE Transactions on Automatic Control, **AC-19** (1974), 716–723.

[12] Wax, M., and Kailath, T.
Detection of signals by information theoretic criteria.
IEEE Transactions on Acoustics, Speech and Signal Processing, **33**, 2 (Apr. 1985), 387–392.

[13] Kay, S. M.
Exponentially embedded families: A new approach to model order estimation.
IEEE Transactions on Aerospace and Electronic Systems, **41**, 1 (Jan. 2005), 333–345.

[14] Seghouane, A.-K.
New AIC corrected variants for multivariate linear regression model selection.
IEEE Transactions on Aerospace and Electronic Systems, **47**, 2 (Apr. 2011), 1154–1165.

[15] Niu, R., Blum, R., Varshney, P., and Drozd, A.
Target localization and tracking in on coherent multiple-input multiple-output radar systems.
IEEE Transactions on Aerospace and Electronic Systems, **48**, 2 (2012), 1466–1489.

[16] Xu, L., Li, J., and Stoica, P.
Target detection and parameter estimation for MIMO radar systems.
IEEE Transactions on Aerospace and Electronic Systems, **44**, 3 (2008), 927–939.

[17] Gorji, A. A., Tharmarasa, R., Blair, W., and Kirubarajan, T.
Multiple unresolved target localization and tracking using colocated MIMO radars.
IEEE Transactions on Aerospace and Electronic Systems, **48**, 3 (2012), 2498–2517.

[18] Mestre, X., and Lagunas, M. A.
Modified subspace algorithms for DOA estimation with large arrays.
IEEE Transactions on Signal Processing, **56**, 2 (Feb. 2008), 598–613.

[19] Kritchman, S., and Nadler, B.
Non-parametric detection of the number of signals: Hypothesis testing and random matrix theory.
IEEE Transactions on Signal Processing, **57**, 10 (Oct. 2009), 3930–3941.

[20] Ledoit, O., and Wolf, M.
A well-conditioned estimator for large-dimensional covariance matrices.
Journal of Multivariate Analysis, **88** (2004), 365–411.

[21] Du, L., Li, J., and Stoica, P.
Fully automatic computation of diagonal loading levels for robust adaptive beamforming.
IEEE Transactions on Aerospace and Electronic Systems, **46**, 1 (Jan. 2010), 449–458.

[22] Stoica, P., Li, J., Zhu, X., and Guerci, J. R.
On using a priori knowledge in space–time adaptive processing.
IEEE Transactions on Signal Processing, **56**, 6 (July 2008), 2598–2602.

[23] Tracy, C. A., and Widom, H.
On orthogonal and symplectic matrix ensembles.
Communications in Mathematical Physics, **177** (1996), 727–754.

[24] Anderson, T. W.
Asymptotic theory for principal component analysis.
Annals of Mathematical Statistics, **34** (1963), 122–148.

[25] Letac, G., and Massam, H.
All invariant moments of the Wishart distribution.
Scandinavian Journal of Statistics, **31**, 2 (2004), 285–318.

[26] Bai, Z. D., and Silverstein, J. W.
CLT for linear spectral statistics of a large dimensional sample covariance matrix.
Annals of Probability, **32** (2004), 553–605.

[27] Wu, H. T., Yang, J. F., and Chen, F. K.
Source number estimators using transformed Gerschgorin radii.
IEEE Transactions on Signal Processing, **43**, 6 (June 1995), 1325C1333.

[28] Huang, L., Long, T., and Wu, S.
Source enumeration for high-resolution array processing using improved Gerschgorin radii without eigendecomposition.
IEEE Transactions on Signal Processing, **56**, 12 (Dec. 2008), 5916–5925.

[29] Jonsson, D.
Some limit theorems for the eigenvalues of a sample covariance matrix.
Journal of Multivariate Analysis, **12** (Sept. 1982), 1–38.

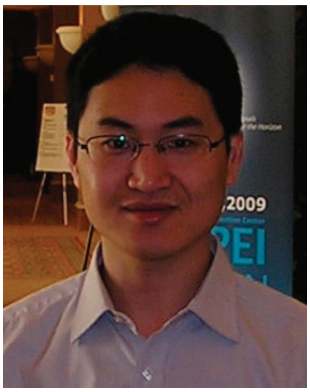
Lei Huang (M'07—SM'14) was born in Guangdong, China. He received B.Sc., M.Sc., and Ph.D. degrees in electronic engineering from Xidian University, Xian, China, in 2000, 2003, and 2005, respectively.

From 2005 to 2006, he was a research associate with the Department of Electrical and Computer Engineering, Duke University, Durham, NC. From 2009 to 2010, he was a research fellow with the Department of Electronic Engineering, City University of Hong Kong, and a research associate with the Department of Electronic Engineering, The Chinese University of Hong Kong. From 2011 to 2014, he was a Professor with the Department of Electronic and Information Engineering, Harbin Institute of Technology Shenzhen Graduate School. Since November 2014, he has joined the College of Information Engineering, Shenzhen University, where he is currently a Chair Professor. His research interests include spectral estimation, array signal processing, and statistical signal processing and their applications in radar and wireless communication systems. He is an editorial board member of *Digital Signal Processing.*

Cheng Qian was born in Zhejiang on November 27, 1988. He received a B.E. degree in communication engineering from Hangzhou Dianzi University, Hangzhou, China, in 2011, and an M.E. degree in information and communication engineering from Harbin Institute of Technology (HIT), Shenzhen, China, in 2013. He is pursuing a Ph.D. degree in the field of information and communication engineering at HIT. His research interests are in array signal processing and MIMO radar.

Hing Cheung So (S'90—M'95—SM'07—F'15) was born in Hong Kong. He received a B.Eng. degree from the City University of Hong Kong and a Ph.D. degree from The Chinese University of Hong Kong, both in electronic engineering, in 1990 and 1995, respectively. From 1990 to 1991, he was an electronic engineer with the Research and Development Division, Everex Systems Engineering Ltd., Hong Kong. During 1995–1996, he worked as a postdoctoral fellow with The Chinese University of Hong Kong. From 1996 to 1999, he was a research assistant professor with the Department of Electronic Engineering, City University of Hong Kong, where he is now an associate professor. His research interests include statistical signal processing, fast and adaptive algorithms, signal detection, parameter estimation, and source localization.

Dr. So has been on the editorial boards of *IEEE Signal Processing Magazine, IEEE Transactions on Signal Processing, Signal Processing,* and *Digital Signal Processing,* as well as a member of the Signal Processing Theory and Methods Technical Committee of the Institute of Electrical and Electronics Engineers (IEEE) Signal Processing Society.



Jun Fang (M'08) received B.S. and M.S. degrees from the Xidian University, Xian, China, in 1998 and 2001, respectively, and a Ph.D. degree from the National University of Singapore, Singapore, in 2006, all in electrical engineering.

During 2006, he was a postdoctoral research associate in the Department of Electrical and Computer Engineering, Duke University. From January 2007 to December 2010, he was a research associate with the Department of Electrical and Computer Engineering, Stevens Institute of Technology. Since January 2011, he has been with the University of Electronic of Science and Technology of China. His current research interests include statistical signal processing, sparse theory and compressed sensing, and Bayesian statistical inference.

Dr. Fang received the IEEE Vehicular Technology Society Jack Neubauer Memorial Award in 2013 and the IEEE Africon Outstanding Paper Award in 2011. He is an associate technical editor for *IEEE Communications Magazine* and an associate editor for *IEEE Signal Processing Letters.*