IEEE TRANSACTIONS ON

SIGNAL PROCESSING

A PUBLICATION OF THE IEEE SIGNAL PROCESSING SOCIETY

IEEE Signal Processing Society

www.ieee.org/sp/index.html

APRIL 2007 VOLUME 55 NUMBER 4 ITPRED (ISSN 1053-587X)

REGULAR PAPERS

(Contents Continued on Back Cover)

IEEE

(Contents Continued from Front Cover)

Simple and Efficient Nonparametric Method for Estimating the Number of Signals Without Eigendecomposition

Jingmin Xin, *Senior Member, IEEE*, Nanning Zheng, *Fellow, IEEE*, and Akira Sano, *Member, IEEE*

***Abstract*—Inspired by the computational simplicity and numerical stability of QR decomposition, a nonparametric method for estimating the number of signals without eigendecomposition (MENSE) is proposed for the coherent narrowband signals impinging on a uniform linear array (ULA). By exploiting the array geometry and its shift invariance property to decorrelate the coherency of signals through subarray averaging, the number of signals is revealed in the rank of the QR upper-trapezoidal factor of the autoproduct of a combined Hankel matrix formed from the cross correlations between some sensor data. Since the infection of additive noise is defused, signal detection capability is improved. A new detection criterion is then formulated in terms of the row elements of the QR upper-triangular factor when finite array data are available, and the number of signals is determined as a value of the running index for which this ratio criterion is maximized, where the QR decomposition with column pivoting is also used to improve detection performance. The statistical analysis clarifies that the MENSE detection criterion is asymptotically consistent. Furthermore, the proposed MENSE algorithm is robust against the array uncertainties including sensor gain and phase errors and mutual coupling and against the deviations from the spatial homogeneity of noise model. The effectiveness of the MENSE is verified through numerical examples, and the simulation results show that the MENSE is superior in detecting closely spaced signals with a small number of snapshots and/or at relatively low signal-to-noise ratio (SNR).**

***Index Terms*—Array signal processing, direction-of-arrival estimation, eigendecomposition, multipath environment, QR decomposition, signal detection.**

I. Introduction

ESTIMATING the number of incident signals from noisy array data is an essential prerequisite for high-resolution direction-of-arrival estimation in array processing (e.g., [1]–[3] and references therein), where the performance of direction estimation can be adversely affected if the number of signals is inaccurately determined. When the incoming signals are incoherent in the presence of temporally and spatially white Gaussian additive noise, eigenstructure-based nonparametric detection methods have been proposed (e.g., [7], [13]–[19], and [32]), where the relation between the number of signals and the "multiplicity" of the smallest eigenvalues of the array covariance matrix is utilized. The most popular ones are the Akaike information criterion (AIC) and minimum description length (MDL) criterion [7] originally introduced to select model order [4]–[6], which are formulated in terms of eigenvalues without any subjective setting of a threshold (confidence levels) required in conventional hypothesis testing, and their performance have been widely studied [8]–[12]. Although nonparametric methods have been attracted considerable attention because of their relatively computational simplicity without the need to estimate direction parameters, they suffer from serious degradation, when the incident signals are coherent (i.e., fully correlated) such as in multipath propagation environments, where the rank of the signal covariance matrix is smaller than the number of signals. Furthermore, the AIC and MDL criterion are very sensitive to the array uncertainties and the deviations from the spatial homogenous assumption of the additive noise in some practical situations [17], [52]–[54].

Generally, parametric detection methods are the optimal approaches for the coherent signals by combining the problems of detecting the number of signals and estimating their directions as a whole [20]–[26], [50]. However, they are computationally unattractive, because maximum-likelihood (ML) direction estimation is involved [20]–[22], [25], [26], [50], which usually requires a nonlinear and multidimensional optimization procedure, and the choice of a judiciously chosen threshold is not easy without any *a priori* knowledge in the case of a finite number of snapshots [23], [24]. Another solution of detecting coherent signals is nonparametric methods [12], [28]–[30] with decorrelation preprocessing such as spatial smoothing (SS) [27]; unfortunately, their performance is usually susceptible to the accuracy of the estimated eigenvalues. In particular, these nonparametric methods perform poorly in difficult scenarios with closely spaced signals, low signal-to-noise ratio (SNR), and a small number of snapshots, because the bias in the estimated eigenvalues becomes quite significant and the population of eigenvalues is not well separated (e.g., [32]). In addition, the regularization-based method [3] has the drawback of slightly increased computational complexity. Furthermore, most nonparametric and parametric methods require the eigendecomposition such as eigenvalue decomposition (EVD) or singular value decomposition (SVD) of the (smoothed) correlation matrix, which is computationally demanding [33]–[36]. Therefore, a considerable computation amount of EVD/SVD

Manuscript received September 14, 2005; revised June 1, 2006. The associate editor coordinating the review of this paper and approving it for publication was Prof. R. Leyman. This work was also supported partially by the National Natural Science Foundation of China under Grant 60021302. A portion of this work was presented at the IEEE 13th Workshop on Statistical Signal Processing (SSP'05), Bordeaux, France, July 2005.

J. Xin was with the Network Systems Laboratories, Fujitsu Laboratories, Limited, Yokosuka 239-0847, Japan. He is now with the Institute of Artificial Intelligence and Robotics, Xi'an Jiaotong University, Xi'an 710049, China (e-mail: jxin@mail.xjtu.edu.cn).

N. Zheng is with the Institute of Artificial Intelligence and Robotics, Xi'an Jiaotong University, Xi'an 710049, China (e-mail: nnzheng@mail.xjtu.edu.cn).

A. Sano is with the Department of System Design Engineering, Keio University, Yokohama 223-8522, Japan (e-mail: sano@sd.keio.ac.jp).

Digital Object Identifier 10.1109/TSP.2006.889982

turns out to be a major obstacle to real-time implementation of most detection methods, especially when the number of sensors is large and/or online detection is required.

To eliminate the need for the costly computation of eigenvalues, some detection methods without EVD/SVD have emerged [35], [37]–[40]. Because the QR decomposition, which factors a matrix into the product of a unitary matrix and an upper-triangular matrix (customarily represented as $\boldsymbol{Q}$ and $\boldsymbol{R}$) requires much lesser computational effort than the complete eigendecomposition, it is a useful alternative to the EVD/SVD and is more amenable to real-time implementation (e.g., [33], [34], [41], and [42]). However, the first QR-based detection method with SS preprocessing needs *a priori* knowledge of true noise variance and subjective assessment [35], [36], and its performance generally degrades in difficult scenarios. Other QR-based detectors for the incoherent signals are essentially the modified/approximate MDL method [37], [38], where the ordinary eigenvalues are replaced by the diagonal elements of the QR upper-triangular factor of the array covariance matrix or their lower/upper bounds provided by the rank revealing QR factorization (RRQR) [43]. However, the former [37] is poor in difficult scenarios, while the latter [38] is complicated because of more computations of the involved RRQR. The supervised training approach [39] is restricted for no more than two incoherent/coherent signals and is not computationally efficient enough due to its implementation with neural network. Though the maximum-likelihood update (MUD) detection scheme is applicable to uncorrelated and correlated signals [40], which is a modification to the parametric method of weight subspace fitting (WSF) detection [23], [24] without EVD, it is still computationally more complicated.

In this paper, we investigate the detection of coherent narrowband signals impinging on the a uniform linear array (ULA) and propose a QR-based nonparametric method for estimating the number of signals without eigendecomposition (MENSE). By exploiting the array geometry and its shift invariance property to decorrelate the coherency of signals through subarray averaging, the number of signals is equal to the rank of a combined Hankel matrix formed from the cross correlations between some sensor data and can be revealed in the rank of the QR upper-trapezoidal factor of the autoproduct of this combined matrix. Since the influence of additive noise is eliminated, the robustness to noise can be improved. A new detection criterion is then formulated in terms of the elements of the corresponding QR upper-triangular matrix when finite array data are available, and hence the number of signals is determined as the value of the running index for which this criterion is maximized, where the QR decomposition with column pivoting is also used to improve the detection performance. The MENSE criterion is asymptotically consistent, and its detection performance can be predicted by examining the QR decomposition of the asymptotical autoproduct matrix with different permutation matrices. The choices of the subarray size and the predetermined permutation matrix are also considered. The proposed MENSE is computationally efficient and suitable for real-time implementation having remarkable insensitivity to the correlation of incident signals, and it can be extended to the spatially correlated noise by appropriately choosing the used subarrays (i.e., cross correlations of array data). Furthermore, the MENSE algorithm is robust to the array uncertainties including sensor gain and phase errors and mutual coupling and to the statistical modeling errors corresponding to deviations from the Gaussian assumption of signals and from the spatial homogenous assumption of additive noise. It is a companion to the subspace-based method without eigendecomposition (SUMWE) for direction estimation [44] and its online implementation [45]. The simulation results demonstrate that the MENSE is superior to the conventional QR-based method and can effectively detect the closely spaced signals with insignificantly degraded performances as compared to the MDL and AIC methods, when the number of snapshots is small and/or the SNR is low.

II. Data Model and Assumptions

We consider a ULA of M identical and omnidirectional sensors (i.e., unity gain and zero phase shift in all directions) with spacing d and suppose that p narrowband signals $\{s_k(n)\}$ with the carrier frequency f_0 are in the far field of the array and arrive from distinct directions $\{\theta_k\}$, where $M > 2p$, and d is less than or equal to half a wavelength of the carrier. The received signal $y_i(n)$ at the ith sensor can be expressed as

$$y_i(n) = \sum_{k=1}^{p} s_k(n) e^{j\omega_0(i-1)\tau(\theta_k)} + w_i(n) \tag{1}$$

where $w_i(n)$ is the additive noise, $\omega_0 \triangleq 2\pi f_0$, $\tau(\theta_k) \triangleq (d/c)\sin\theta_k$, and c is the propagation speed. Then, the received signals can be rewritten more compactly as

$$\boldsymbol{y}(n) = \boldsymbol{A}\boldsymbol{s}(n) + \boldsymbol{w}(n) \tag{2}$$

where $\boldsymbol{y}(n)$, $\boldsymbol{s}(n)$, and $\boldsymbol{w}(n)$ are the vectors of the received signals, the incident signals, and the additive noise given by $\boldsymbol{y}(n) \triangleq [y_1(n), y_2(n), \cdots, y_M(n)]^T$, $\boldsymbol{s}(n) \triangleq [s_1(n), s_2(n), \cdots, s_p(n)]^T$, $\boldsymbol{w}(n) \triangleq [w_1(n), w_2(n), \cdots, w_M(n)]^T$, $\boldsymbol{A}$ is the array response matrix given by $\boldsymbol{A} \triangleq [\boldsymbol{a}(\theta_1), \boldsymbol{a}(\theta_2), \cdots, \boldsymbol{a}(\theta_p)]$ with $\boldsymbol{a}(\theta_k) \triangleq [1, e^{j\omega_0\tau(\theta_k)}, \cdots, e^{j\omega_0(M-1)\tau(\theta_k)}]^T$, and $(\cdot)^T$ denotes transposition.

Herein, we make the following basic assumptions on the data model, which are similar to those in the SUMWE [44].

A1) The signals $\{s_k(n)\}$ are coherent under the flat-fading multipath propagation and given by

$$s_k(n) = \beta_k s_1(n), \quad \text{for} \quad k = 1, 2, \cdots, p \tag{3}$$

where β_k is the complex attenuation coefficient with $\beta_k \neq 0$ and $\beta_1 = 1$.

A2) For the simplicity of theoretical analysis, the incident signal $s_1(n)$ is a temporally complex white Gaussian random process with zero-mean and the variance given by

$$\begin{aligned} E\{s_1(n)s_1^*(t)\} &= c_s\delta_{n,t} \\ E\{s_1(n)s_1(t)\} &= 0 \end{aligned} \tag{4}$$

$\forall n, t$, where $E\{\cdot\}$, $(\cdot)^*$, and $\delta_{n,t}$ denote the expectation, the complex conjugate, and Kronecker delta.

A3) The additive noise $\{w_i(n)\}$ is a temporally and spatially complex white Gaussian random process with zero-mean and the following variance:

$$E\{w_i(n)w_k^*(t)\} = \sigma^2\delta_{i,k}\delta_{n,t}$$
$$E\{w_i(n)w_k(t)\} = 0 \tag{5}$$

$\forall i,k$ and $\forall n,t$. And the additive noise is uncorrelated with the incident signals.

A4) The number of incident signals p satisfies the inequality that $p < M/2$ for an array of M sensors.

Remark 1: The proposed method can be extended to uncorrelated or correlated signals and to partly correlated or coherent signals, and it can accommodate a more general noise model for the spatially correlated noise if we choose appropriate subarrays (cf. Remark 3 and [44]). The detectability condition is that $M > 2p$ (cf. Remark 4), which is the same as the identifiability condition for the SUMWE [44] and less strict than the necessary condition $M \geq 3p/2$ with probability one for estimation problem [46]. □

Remark 2: Although the ideal omnidirectional array and the normal signal and noise models are assumed, the proposed method is robust against the array uncertainties including sensor gain and phase errors and mutual coupling and against the deviations from the complex white Gaussian assumption of incident signals and from the spatial homogenous assumption of additive noise (cf. Section V). □

III. Estimating the Number of Signals Without Eigendecomposition

Under the assumptions of data model, we easily get the array covariance matrix $\boldsymbol{C}$ as

$$\boldsymbol{C} \triangleq E\{\boldsymbol{y}(n)\boldsymbol{y}^H(n)\} = \boldsymbol{A}\boldsymbol{C}_s\boldsymbol{A}^H + \sigma^2\boldsymbol{I}_M \tag{6}$$

where $(\cdot)^H$ and $\boldsymbol{I}_m$ indicate the Hermitian transposition and $m \times m$ identity matrix, and $\boldsymbol{C}_s$ is the signal covariance matrix given by

$$\boldsymbol{C}_s \triangleq E\{\boldsymbol{s}(n)\boldsymbol{s}^H(n)\} = c_s\boldsymbol{\beta}\boldsymbol{\beta}^H \tag{7}$$

with $\boldsymbol{\beta} \triangleq [\beta_1, \beta_2, \cdots, \beta_p]^T$. By defining the correlation c_{ik} between the noisy signals $y_i(n)$ and $y_k(n)$ as $c_{ik} \triangleq E\{y_i(n)y_k^*(n)\}$, where $c_{ik} = c_{ki}^*$, clearly the diagonal elements $\{c_{ii}\}$ of $\boldsymbol{C}$ in (6) are affected by the noise variance. In addition, since the incident signals are coherent, the matrix $\boldsymbol{C}_s$ in (7) is singular, and the dimension of the signal subspace of $\boldsymbol{C}$ is smaller than the number of incident signals (i.e., $\dim(\mathcal{R}(\boldsymbol{C})) = \text{rank}(\boldsymbol{C}_s) < p$, when $p > 1$). Hence, the number of signals cannot be estimated directly from the multiplicity of the eigenvalues of the matrix $\boldsymbol{C}$ in (6).

A. Decorrelation of Signal Coherency and Insensitivity to Additive Noise

First, for the sake of tractability in theoretical analysis, the noisy received signal $y_i(n)$ in (1) can be reexpressed by

$$y_i(n) = \boldsymbol{b}_i^T(\theta)\boldsymbol{s}(n) + w_i(n) \tag{8}$$

where $\boldsymbol{b}_i(\theta) \triangleq [e^{j\omega_0(i-1)\tau(\theta_1)}, e^{j\omega_0(i-1)\tau(\theta_2)}, \cdots, e^{j\omega_0(i-1)\tau(\theta_p)}]^T$. Now, we can divide the full array into L overlapping subarrays with $\bar{p}$ sensors in the forward and backward directions [31], and the lth forward or backward subarray comprises $\{l, l+1, \cdots, l+\bar{p}-1\}$ or $\{M-l+1, M-l, \cdots, L-l+1\}$ sensors for $l = 1, 2, \cdots, L$, where $L = M - \bar{p} + 1$ and $\bar{p} \geq p$ (see Remark 4 for the choice of subarray size $\bar{p}$). The signal vectors of the lth forward and backward subarrays are given by

$$\begin{aligned}\boldsymbol{y}_{fl}(n) &\triangleq [y_l(n), y_{l+1}(n), \cdots, y_{l+\bar{p}-1}(n)]^T \\ &= \bar{\boldsymbol{A}}_1\boldsymbol{D}^{l-1}\boldsymbol{s}(n) + \boldsymbol{w}_{fl}(n)\end{aligned} \tag{9}$$

$$\begin{aligned}\boldsymbol{y}_{bl}(n) &\triangleq [y_{M-l+1}(n), y_{M-l}(n), \cdots, y_{L-l+1}(n)]^H \\ &= \bar{\boldsymbol{A}}_1\boldsymbol{D}^{-(M-l)}\boldsymbol{s}^*(n) + \boldsymbol{w}_{bl}(n)\end{aligned} \tag{10}$$

where $\boldsymbol{w}_{fl}(n) \triangleq [w_l(n), w_{l+1}(n), \cdots, w_{l+\bar{p}-1}(n)]^T$, $\boldsymbol{w}_{bl}(n) \triangleq [w_{M-l+1}(n), w_{M-l}(n), \cdots, w_{L-l+1}(n)]^H$, $\bar{\boldsymbol{A}}_1$ is the submatrix of $\boldsymbol{A}$ in (2) consisting of the first $\bar{p}$ rows with the column $\bar{\boldsymbol{a}}_1(\theta_k) \triangleq [1, e^{j\omega_0\tau(\theta_k)}, \cdots, e^{j\omega_0(\bar{p}-1)\tau(\theta_k)}]^T$, and $\boldsymbol{D} \triangleq \text{diag}(e^{j\omega_0\tau(\theta_1)}, e^{j\omega_0\tau(\theta_2)}, \cdots, e^{j\omega_0\tau(\theta_p)})$. Then by defining the correlations $\boldsymbol{\varphi}_{fl}$, $\bar{\boldsymbol{\varphi}}_{fl}$, $\boldsymbol{\varphi}_{bl}$, and $\bar{\boldsymbol{\varphi}}_{bl}$ between the signal vectors $\boldsymbol{y}_{fl}(n)$ and $\boldsymbol{y}_{bl}(n)$ in (9) and (10) and the signals $y_1(n)$ and $y_M(n)$ in (1) as $\boldsymbol{\varphi}_{fl} \triangleq E\{\boldsymbol{y}_{fl}(n)y_M^*(n)\}$, $\bar{\boldsymbol{\varphi}}_{fl} \triangleq E\{\boldsymbol{y}_{fl}(n)y_1^*(n)\}$, $\boldsymbol{\varphi}_{bl} \triangleq E\{y_1(n)\boldsymbol{y}_{bl}(n)\}$, and $\bar{\boldsymbol{\varphi}}_{bl} \triangleq E\{y_M(n)\boldsymbol{y}_{bl}(n)\}$, after some algebraic manipulations, we can obtain four Hankel correlation matrices (cf. [44])

$$\boldsymbol{\Phi}_f \triangleq [\boldsymbol{\varphi}_{f1}, \boldsymbol{\varphi}_{f2}, \cdots, \boldsymbol{\varphi}_{fL-1}]^T = c_s\bar{\rho}_M\bar{\boldsymbol{A}}\boldsymbol{B}\bar{\boldsymbol{A}}_1^T \tag{11}$$

$$\bar{\boldsymbol{\Phi}}_f \triangleq [\bar{\boldsymbol{\varphi}}_{f2}, \bar{\boldsymbol{\varphi}}_{f3}, \cdots, \bar{\boldsymbol{\varphi}}_{fL}]^T = c_s\bar{\rho}_1\bar{\boldsymbol{A}}\boldsymbol{B}\boldsymbol{D}\bar{\boldsymbol{A}}_1^T \tag{12}$$

$$\boldsymbol{\Phi}_b \triangleq [\boldsymbol{\varphi}_{b1}, \boldsymbol{\varphi}_{b2}, \cdots, \boldsymbol{\varphi}_{bL-1}]^T = c_s\bar{\rho}_1^*\bar{\boldsymbol{A}}\boldsymbol{B}^*\boldsymbol{D}^{-(M-1)}\bar{\boldsymbol{A}}_1^T \tag{13}$$

$$\bar{\boldsymbol{\Phi}}_b \triangleq [\bar{\boldsymbol{\varphi}}_{b2}, \bar{\boldsymbol{\varphi}}_{b3}, \cdots, \bar{\boldsymbol{\varphi}}_{bL}]^T = c_s\bar{\rho}_M^*\bar{\boldsymbol{A}}\boldsymbol{B}^*\boldsymbol{D}^{-(M-2)}\bar{\boldsymbol{A}}_1^T \tag{14}$$

where $\bar{\boldsymbol{A}}$ is the submatrix of $\boldsymbol{A}$ in (2) consisting of its first $M - \bar{p}$ rows with the column $\bar{\boldsymbol{a}}(\theta_k) = [1, e^{j\omega_0\tau(\theta_k)}, \cdots, e^{j\omega_0(L-2)\tau(\theta_k)}]^T$, $\boldsymbol{B} \triangleq \text{diag}(\beta_1, \beta_2, \cdots, \beta_p)$, and $\bar{\rho}_i \triangleq \boldsymbol{\beta}^H\boldsymbol{b}_i^*(\theta)$.

Obviously, the Hankel correlation matrices $\boldsymbol{\Phi}_f$, $\bar{\boldsymbol{\Phi}}_f$, $\boldsymbol{\Phi}_b$, and $\bar{\boldsymbol{\Phi}}_b$ with the relations $\boldsymbol{\Phi}_b = \boldsymbol{J}_{M-\bar{p}}\bar{\boldsymbol{\Phi}}_f^*\boldsymbol{J}_{\bar{p}}$ and $\bar{\boldsymbol{\Phi}}_b = \boldsymbol{J}_{M-\bar{p}}\boldsymbol{\Phi}_f^*\boldsymbol{J}_{\bar{p}}$ are not affected by the additive noise and can be simply formed from the elements $\{c_{iM}\}$ and $\{c_{i1}\}$ in the Mth and first columns $\boldsymbol{\varphi}$ and $\bar{\boldsymbol{\varphi}}$ of $\boldsymbol{C}$ in (6) except for the autocorrelations c_{11} and c_{MM}, which contain the noise variance σ^2, where $\boldsymbol{J}_m$ is an $m \times m$ counteridentity matrix, $\boldsymbol{\varphi} \triangleq E\{\boldsymbol{y}(n)y_M^*(n)\}$, and $\bar{\boldsymbol{\varphi}} \triangleq E\{\boldsymbol{y}(n)y_1^*(n)\}$. The ranks of these matrices equal the number of signals p iff $M - \bar{p} \geq p$ and $\bar{p} \geq p$ (cf. [3], [44], and [47]),

i.e., the dimension of the signal subspace of these matrices is restored to the number of signals without the ordinary SS preprocessing [27], in which the spatial averaging of subarray covariance matrices is required. The rank property of these matrices and their insensitivity to the additive noise lay the basis for the proposed method.

Remark 3: When the incoming signals are uncorrelated, four Hankel correlation matrices are given by (cf. [44])

$$\boldsymbol{\Phi}_f = \boldsymbol{\Phi}_b = \bar{\boldsymbol{A}}\boldsymbol{D}^{-(M-1)}\boldsymbol{C}_s\bar{\boldsymbol{A}}_1^T \tag{15}$$

$$\bar{\boldsymbol{\Phi}}_f = \bar{\boldsymbol{\Phi}}_b = \bar{\boldsymbol{A}}\boldsymbol{D}\boldsymbol{C}_s\bar{\boldsymbol{A}}_1^T \tag{16}$$

where $\boldsymbol{C}_s \triangleq E\{\boldsymbol{s}(n)\boldsymbol{s}^H(n)\} = \mathrm{diag}(c_{s_1}, c_{s_2}, \cdots, c_{s_p})$, and $c_{s_k} \triangleq E\{s_k(n)s_k^*(n)\}$. Clearly, the ranks of these matrices still equal the number of incident signals p. □

B. QR-Based Approach to Estimating the Number of Signals

From (11)–(14), we define an $(M-\bar{p}) \times 4\bar{p}$ combined correlation matrix $\bar{\boldsymbol{\Phi}}$ as

$$\boldsymbol{\Phi} \triangleq [\boldsymbol{\Phi}_f, \bar{\boldsymbol{\Phi}}_f, \boldsymbol{\Phi}_b, \bar{\boldsymbol{\Phi}}_b] = c_s\bar{\rho}_M\bar{\boldsymbol{A}}\boldsymbol{B}\boldsymbol{F} \tag{17}$$

where $\boldsymbol{F} \triangleq [\bar{\boldsymbol{A}}_1^T, (\bar{\rho}_1/\bar{\rho}_M)\boldsymbol{D}\bar{\boldsymbol{A}}_1^T, (\bar{\rho}_1^*/\bar{\rho}_M)\bar{\boldsymbol{B}}\boldsymbol{D}^{-(M-1)}\bar{\boldsymbol{A}}_1^T, (\bar{\rho}_M^*/\bar{\rho}_M)\bar{\boldsymbol{B}}\boldsymbol{D}^{-(M-2)}\bar{\boldsymbol{A}}_1^T]$, and $\bar{\boldsymbol{B}} \triangleq \boldsymbol{B}^{-1}\boldsymbol{B}^* = \mathrm{diag}(\beta_1^*/\beta_1, \beta_2^*/\beta_2, \cdots, \beta_p^*/\beta_p)$. Then, we obtain an $(M-\bar{p}) \times (M-\bar{p})$ autoproduct $\boldsymbol{\Psi}$ of matrix $\boldsymbol{\Phi}$ as

$$\boldsymbol{\Psi} \triangleq \boldsymbol{\Phi}\boldsymbol{\Phi}^H = c_s^2|\bar{\rho}_M|^2\bar{\boldsymbol{A}}\boldsymbol{B}\boldsymbol{F}\boldsymbol{F}^H\boldsymbol{B}^H\bar{\boldsymbol{A}}^H \tag{18}$$

which is a special centrosymmetric matrix with double symmetry (Hermitian and persymmetric about the principle and cross diagonals, respectively), i.e.,

$$\boldsymbol{\Psi} = \boldsymbol{\Psi}^H = \boldsymbol{J}_{M-\bar{p}}\boldsymbol{\Psi}^T\boldsymbol{J}_{M-\bar{p}} = \boldsymbol{J}_{M-\bar{p}}\boldsymbol{\Psi}^*\boldsymbol{J}_{M-\bar{p}}. \tag{19}$$

Then, we have the following relation between the number of incoming signals and the rank of the QR upper-trapezoidal factor of $\boldsymbol{\Psi}$ in (18).

Theorem 1: The number of incident signals equals the rank of the QR upper-trapezoidal factor $\boldsymbol{R}$ of the matrix $\boldsymbol{\Psi}$ in (18), when the detectability condition that $p \le \bar{p} < M-p$ is satisfied, i.e., $\mathrm{rank}(\boldsymbol{R}) = p$, where the Householder QR decomposition of the matrix $\boldsymbol{\Psi}$ is given by

$$\boldsymbol{\Psi} = \boldsymbol{Q}\boldsymbol{R} = \boldsymbol{Q}\begin{bmatrix} \boldsymbol{R}_{11}, & \boldsymbol{R}_{12} \\ \boldsymbol{O}_{(M-\bar{p}-p)\times(M-\bar{p})} & \end{bmatrix}\begin{matrix}\}p \\ \}M-\bar{p}-p\end{matrix} \tag{20}$$

where $\boldsymbol{Q}$ is the $(M-\bar{p}) \times (M-\bar{p})$ unitary matrix, $\boldsymbol{R}_{11}$ is the $p \times p$ upper-triangular and nonsingular matrix, $\boldsymbol{R}_{12}$ is the $p \times (M-\bar{p}-p)$ matrix with nonzero elements, and $\boldsymbol{O}_{m\times q}$ denotes the $m \times q$ null matrix.

Proof: See Appendix A. ■

Thus, the rank property of the QR factor $\boldsymbol{R}$ in Theorem 1 has the potential for detecting the number of signals due to the avoidance of EVD and the robustness to additive noise.

Remark 4: As the information on the number of signals is unavailable, the subarray size $\bar{p}$ should be chosen appropriately. From the detectability condition $p \le \bar{p} < M-p$, where $p \ge 1$, we can find that the maximum detectable number of signals is $p < M/2$, i.e., $p_{\max} = \lceil M/2 \rceil - 1$, where $\lceil x \rceil$ denotes the smallest integer not less than x. Therefore, we can choose a conservative value of $\bar{p} = \lfloor M/2 \rfloor$, which satisfies the inequality condition that $p_{\max} \le \bar{p} = \lfloor M/2 \rfloor < M - p_{\max}$, where $\lfloor x \rfloor$ denotes the largest integer not greater than x. □

C. Detecting Absence of Incident Signal

In many fields such as radar and sonar (e.g., [52] and [55]), we sometimes encounter the absence of incident signal(s). Hence, we should firstly infer whether the measurements consist of noise only or of a number of incident signals corrupted by noise and then have to determine the exact number of signals if they are present. From Theorem 1, we can easily get the following lemma for the case of one signal, which is useful for setting a detection threshold automatically without the need for "manual" adjustment to detect the absence or presence of incident signal(s).

Lemma: For the special case of a single signal, the correlation matrix $\boldsymbol{\Phi}$ in (17) is given by

$$\begin{aligned}\boldsymbol{\Phi} &= [\boldsymbol{\Phi}_f, \bar{\boldsymbol{\Phi}}_f, \boldsymbol{\Phi}_b, \bar{\boldsymbol{\Phi}}_b] \\ &= c_s e^{-j\omega_0(M-1)\tau(\theta_1)}\bar{\boldsymbol{a}}(\theta_1) \\ &\quad \times \left[\bar{\boldsymbol{a}}_1^T(\theta_1), e^{j\omega_0 M\tau(\theta_1)}\bar{\boldsymbol{a}}_1^T(\theta_1), \bar{\boldsymbol{a}}_1^T(\theta_1), e^{j\omega_0 M\tau(\theta_1)}\bar{\boldsymbol{a}}_1^T(\theta_1)\right]\end{aligned} \tag{21}$$

and the analytical expressions of the QR factors of the autoproduct $\boldsymbol{\Psi}$ of $\boldsymbol{\Phi}$ in (21) can be explicitly given by

$$\begin{aligned}\boldsymbol{Q} &= \boldsymbol{I}_{M-\bar{p}} - \frac{1}{(M-\bar{p})^{1/2} + M - \bar{p}} \\ &\quad \times \left(\bar{\boldsymbol{a}}(\theta_1)\bar{\boldsymbol{a}}^H(\theta_1) + (M-\bar{p})^{1/2}\left(\boldsymbol{e}_1\bar{\boldsymbol{a}}^H(\theta_1) + \bar{\boldsymbol{a}}(\theta_1)\boldsymbol{e}_1^H\right)\right. \\ &\quad \left. + (M-\bar{p})\boldsymbol{e}_1\boldsymbol{e}_1^H\right)\end{aligned} \tag{22}$$

$$\boldsymbol{R} = -4c_s^2\bar{p}\,\|\bar{\boldsymbol{a}}(\theta_1)\|\,\boldsymbol{e}_1\bar{\boldsymbol{a}}^H(\theta_1) \tag{23}$$

where $\boldsymbol{e}_i$ is an $(M-\bar{p}) \times 1$ unit vector with a unity element at the ith location and zeros elsewhere, and $\mathrm{rank}(\boldsymbol{R}) = p = 1$.

Proof: This Lemma can be established from Theorem 1 and Appendix A immediately. ■

When there is only one impinging signal, from (6), the autocorrelation c_{ii} of the received signal $y_i(n)$ is given by

$$c_{ii} = c_s + \sigma^2, \quad \text{for} \quad i = 1, 2, \cdots, M. \tag{24}$$

By comparing this correlation c_{11} with the element r_{11} of the QR factor matrix $\boldsymbol{R}$ in (23), we have

$$\frac{|r_{11}|}{c_{11}} = \frac{4c_s^2\bar{p}(M-\bar{p})^{1/2}}{c_s + \sigma^2} = 4\bar{p}(M-\bar{p})^{1/2}c_s\frac{\mathrm{SNR}}{\mathrm{SNR}+1} \tag{25}$$

where we set $\mathrm{SNR} \triangleq c_s/\sigma^2$ and assume that $\sigma^2 \neq 0$. If the SNR and the power c_s of signal $s_1(n)$ satisfy the condition

$$c_s \ge \frac{1}{4\bar{p}(M-\bar{p})^{1/2}}\left(1 + \frac{1}{\mathrm{SNR}}\right) \triangleq \alpha_{c_s} \tag{26}$$

we find that $|r_{11}| \geq c_{11}$. Hence, if the subarray size is chosen as $\bar{p} = \lfloor M/2 \rfloor$ (cf. Remark 4), when SNR ≥ 0.1 (i.e., SNR ≥ -10 dB), the low threshold α_{c_s} can be approximated as

$$\alpha_{c_s} \leq \frac{11}{4\bar{p}(M-\bar{p})^{1/2}} \approx 5.5\sqrt{2}M^{-3/2} \triangleq \bar{\alpha}_{c_s}. \tag{27}$$

Consequently, as long as the signal power c_s is not less than $\bar{\alpha}_{c_s}$, we can find that $|r_{11}| \geq c_{11}$; whereas when there is no signal, evidently we can obtain that $|r_{11}| = 0 < c_{11} = \sigma^2$.

Thus, for the case of SNR ≥ -10 dB and $\bar{p} = \lfloor M/2 \rfloor$, when $c_s \geq \bar{\alpha}_{c_s}$, which is a reasonable condition in practice where the number of sensors M is usually not too small and hence results in small $\bar{\alpha}_{c_s}$, by letting the binary hypotheses be

$$\begin{aligned} \mathcal{H}_0: &\quad p = 0 \quad \text{(absence)} \\ \mathcal{H}_1: &\quad p > 0 \quad \text{(presence)} \end{aligned} \tag{28}$$

we can accept $\mathcal{H}_0$ or $\mathcal{H}_1$ by quantitatively comparing the element r_{11} of the QR factor $\boldsymbol{R}$ and the autocorrelation c_{11} of the signal $y_1(n)$, i.e.,

$$|r_{11}| \underset{\mathcal{H}_1}{\overset{\mathcal{H}_0}{\lessgtr}} c_{11}. \tag{29}$$

If $\mathcal{H}_0$ is accepted, we can see that there is no incident signal; otherwise $\mathcal{H}_1$ is accepted, then the number of incident signals should further be estimated by the rank property in Theorem 1.

D. Implementation of MENSE With Accessible Noisy Data

When only finite and noisy array data are available, the Hankel correlation matrices in (11)–(14) (and hence $\bar{\boldsymbol{\Phi}}$) should be replaced with their estimates $\hat{\bar{\boldsymbol{\Phi}}}_f$, $\hat{\bar{\bar{\boldsymbol{\Phi}}}}_f$, $\hat{\bar{\boldsymbol{\Phi}}}_b$, and $\hat{\bar{\bar{\boldsymbol{\Phi}}}}_b$ (and hence $\hat{\bar{\boldsymbol{\Phi}}}$), and then the sample estimate of autoproduct $\hat{\bar{\boldsymbol{\Psi}}}$ of $\hat{\bar{\boldsymbol{\Phi}}}$ is given by

$$\hat{\bar{\boldsymbol{\Psi}}} = \hat{\bar{\boldsymbol{\Phi}}}\hat{\bar{\boldsymbol{\Phi}}}^H = \hat{\bar{\boldsymbol{\Phi}}}_f\hat{\bar{\boldsymbol{\Phi}}}_f^H + \hat{\bar{\bar{\boldsymbol{\Phi}}}}_f\hat{\bar{\bar{\boldsymbol{\Phi}}}}_f^H + \hat{\bar{\boldsymbol{\Phi}}}_b\hat{\bar{\boldsymbol{\Phi}}}_b^H + \hat{\bar{\bar{\boldsymbol{\Phi}}}}_b\hat{\bar{\bar{\boldsymbol{\Phi}}}}_b^H. \tag{30}$$

However, the principle effect of using a finite number of snapshots N is that the influence of noise will not be completely cancelled out in the sample estimate $\hat{\bar{\boldsymbol{\Psi}}}$ in (30). As a result, when the number of snapshots N is not sufficiently large, the QR factor $\hat{\boldsymbol{R}}$ of $\hat{\bar{\boldsymbol{\Psi}}}$ will be perturbed from its true value $\boldsymbol{R}$ in (20) and may become an upper-triangular and nonsingular matrix with full rank due to the effect of estimation error. Consequently, the rank property in Theorem 1 will not be satisfied exactly, and the number of incident signals (i.e., the effective rank of $\hat{\bar{\boldsymbol{\Psi}}}$) could not be determined simply by comparing the magnitude relation between the diagonal elements of $\hat{\boldsymbol{R}}$. Here, we propose a ratio criterion based on the row elements of QR factor $\hat{\boldsymbol{R}}$ to estimate the number of signals without any ad hoc threshold setting.

Now, performing the QR decomposition with column pivoting to the matrix $\hat{\bar{\boldsymbol{\Psi}}}$ in (30) (e.g., [33]), we get

$$\hat{\bar{\boldsymbol{\Psi}}}\boldsymbol{\Pi} = \hat{\boldsymbol{Q}}\hat{\boldsymbol{R}} = \hat{\boldsymbol{Q}}\begin{bmatrix} \hat{\boldsymbol{R}}_{11}, & \hat{\boldsymbol{R}}_{12} \\ \boldsymbol{O}_{(M-\bar{p}-p)\times p}, & \hat{\boldsymbol{R}}_{22} \end{bmatrix} \begin{matrix} \}p \\ \}M-\bar{p}-p \end{matrix} \tag{31}$$

where $\boldsymbol{\Pi}$ is an $(M-\bar{p}) \times (M-\bar{p})$ permutation matrix, which is used to represent different methods of the QR decomposition with column interchanges (see Section IV-C for the choice of $\boldsymbol{\Pi}$). Then, by introducing an auxiliary quantity $\zeta(i)$ in terms of the nonzero elements of the ith row of QR factor $\hat{\boldsymbol{R}}$ as

$$\zeta(i) \triangleq \sum_{k=i}^{M-\bar{p}} |\hat{r}_{ik}| + \varepsilon, \quad \text{for } i = 1, 2, \cdots, M-\bar{p} \tag{32}$$

we can define a ratio criterion $\xi(i)$ as

$$\xi(i) \triangleq \frac{\zeta(i)}{\zeta(i+1)}, \quad \text{for } i = 1, 2, \cdots, M-\bar{p}-1 \tag{33}$$

where ε is an arbitrary and positive small constant (e.g., $\varepsilon = 10^{-10}$) for avoiding the possibly undetermined ratio of 0/0 in (33). Thus, the number of incident signals is determined as the value of the running index $i \in \{1, 2, \cdots, M-\bar{p}-1\}$ for which the criterion $\xi(i)$ is maximized, i.e.,

$$\hat{p} = \arg\max_i \xi(i). \tag{34}$$

Moreover, with an finite number of snapshots N, the detection rule in (29) is modified as

$$|\hat{r}_{11}| \underset{\mathcal{H}_1}{\overset{\mathcal{H}_0}{\lessgtr}} \hat{c}_{11}. \tag{35}$$

Then, if $\mathcal{H}_0$ is accepted, we can get $\hat{p} = 0$. Otherwise, we further have to estimate the number of signals with (34).

Therefore, the implementation of the proposed MENSE with finite array data $\{\boldsymbol{y}(n)\}_{n=1}^N$ and the computational complexity of each step can be summarized as follows.

1) Set the subarray size to $\bar{p} = \lfloor M/2 \rfloor$. .....1 flops
2) Calculate the sample correlation vector $\hat{\boldsymbol{\varphi}}$ between $\boldsymbol{y}(n)$ and $y_M^*(n)$ and those of $\hat{\bar{\boldsymbol{\varphi}}}$ between $\boldsymbol{y}(n)$ and $y_1^*(n)$ as

$$\hat{\boldsymbol{\varphi}} = \frac{1}{N}\sum_{n=1}^N \boldsymbol{y}(n)y_M^*(n), \quad \hat{\bar{\boldsymbol{\varphi}}} = \frac{1}{N}\sum_{n=1}^N \boldsymbol{y}(n)y_1^*(n) \tag{36}$$

where $\hat{\boldsymbol{\varphi}} = [\hat{c}_{1M}, \hat{c}_{2M}, \cdots, \hat{c}_{MM}]^T$, and $\hat{\bar{\boldsymbol{\varphi}}} = [\hat{c}_{11}, \hat{c}_{21}, \cdots, \hat{c}_{M1}]^T$. .....$4(4N+3)M$ flops
3) Form the estimated Hankel correlation matrices $\hat{\bar{\boldsymbol{\Phi}}}_f$, $\hat{\bar{\bar{\boldsymbol{\Phi}}}}_f$, $\hat{\bar{\boldsymbol{\Phi}}}_b$, and $\hat{\bar{\bar{\boldsymbol{\Phi}}}}_b$, and hence the estimated matrix $\hat{\bar{\boldsymbol{\Phi}}}$ from $\hat{\boldsymbol{\varphi}}$ and $\hat{\bar{\boldsymbol{\varphi}}}$, as

$$\hat{\bar{\boldsymbol{\Phi}}}_f = \text{Hank}\{\boldsymbol{h}_c, \boldsymbol{h}_r\}, \quad \hat{\bar{\bar{\boldsymbol{\Phi}}}}_f = \text{Hank}\{\bar{\boldsymbol{h}}_c, \bar{\boldsymbol{h}}_r\} \tag{37}$$

$$\hat{\bar{\boldsymbol{\Phi}}}_b = \boldsymbol{J}_{M-\bar{p}}\hat{\bar{\bar{\boldsymbol{\Phi}}}}_f^*\boldsymbol{J}_{\bar{p}}, \quad \hat{\bar{\bar{\boldsymbol{\Phi}}}}_b = \boldsymbol{J}_{M-\bar{p}}\hat{\bar{\boldsymbol{\Phi}}}_f^*\boldsymbol{J}_{\bar{p}} \tag{38}$$

$$\hat{\bar{\boldsymbol{\Phi}}} = [\hat{\bar{\boldsymbol{\Phi}}}_f, \hat{\bar{\bar{\boldsymbol{\Phi}}}}_f, \hat{\bar{\boldsymbol{\Phi}}}_b, \hat{\bar{\bar{\boldsymbol{\Phi}}}}_b] \tag{39}$$

where $\boldsymbol{h}_c = [\hat{c}_{1M}, \hat{c}_{2M}, \cdots, \hat{c}_{M-\bar{p},M}]^T$, $\boldsymbol{h}_r = [\hat{c}_{M-\bar{p},M}, \hat{c}_{M-\bar{p}+1,M}, \cdots, \hat{c}_{M-1,M}]^T$, $\bar{\boldsymbol{h}}_c = [\hat{c}_{21}, \hat{c}_{31}, \cdots, \hat{c}_{L1}]^T$, $\bar{\boldsymbol{h}}_r = [\hat{c}_{L1}, \hat{c}_{L+1,1}, \cdots, \hat{c}_{M1}]^T$, and Hank$\{\cdot\}$ denotes the Hankel operation. .....$2(M-\bar{p})\bar{p}+2$ flops
4) Calculate the autoproduct $\hat{\bar{\boldsymbol{\Psi}}}$ with (30) and perform its QR decomposition with (31). .....$32(M-\bar{p})^2\bar{p} + 14(M-\bar{p})^3 - 6(M-\bar{p})^2 + 22(M-\bar{p}) - 5$ flops

TABLE I
COMPARISON BETWEEN THE COMPUTATIONAL COMPLEXITY OF THE IMPLEMENTATION OF THE MDL AND SS/FFSS-BASED MDL METHODS AND THAT OF THE MENSET ALGORITHM IN MATLAB FLOPS

Step	MDL	SS/FBSS-MDL	MENSE
Calculation of correlations	$(8N+6)M^2$	$(8N+6)(M+(m-1)(2M-m))$	$4(4N+3)M$
Formulation of needed matrix		$2m^2(M-m+1)+6m^2$	$2(M-\bar{p})\bar{p}+32(M-\bar{p})^2\bar{p}+2$
EVD/QR operation	$\mathcal{O}(M^3)$	$\mathcal{O}(m^3)$	$14(M-\bar{p})^3-6(M-\bar{p})^2+22(M-\bar{p})-5$
Calculation of criterion	$M(M+16)$	$m(m+16)$	$2.5(M-\bar{p}+1)(M-\bar{p})+2(M-\bar{p})-1$

5) If $|\hat{r}_{11}| < \hat{c}_{11}$, the number of signal is estimated as $\hat{p}=0$, and stop the detection procedure; otherwise $\hat{p} \neq 0$, and continue with the next step. ..... 0 flops
6) Calculate the ratio criterion $\xi(i)$ with (32) and (33) for $i = 1, 2, \cdots, M-\bar{p}-1$ and determine the number of incident signals with (34).
..... $2.5(M-\bar{p}+1)(M-\bar{p})+2(M-\bar{p})-1$ flops

The computational complexity above is roughly indicated in terms of the number of flops, where a flop is defined as a floating-point addition or multiplication operation as adopted by the MATLAB software, and a predetermined permutation matrix $\mathbf{\Pi}$ is used in Step 4) to simplify calculation.

Remark 5: Since the unitary matrix $\boldsymbol{Q}$ (or $\hat{\boldsymbol{Q}}$) of the QR decomposition is not necessary, the detection algorithm can be implemented more efficiently without computing this QR factor (see [45, App. A] for basic algorithm for Householder QR decomposition). □

Remark 6: The implementation of the benchmark MDL or AIC method [7] and its variants usually involve three major steps: i) calculation of the array covariance matrix or the spatially smoothed subarray covariance matrix (for the SS and forward-backward SS (FBSS) based MDL/AIC) and ii) computation of eigenvalues with EVD, and iii) calculation of the criterion. The computational complexities of the MDL, SS/FBSS-based MDL in MATLAB flops are roughly shown and compared with that of the MENSE algorithm in Table I, where m is the subarray size used in the SS/FBSS-based method with $m > p$ and $M-m+1 \geq p$ (for SS) or $2(M-m+1) \geq p$ (for FBSS) [27], [31].

By denoting the numbers of MATLAB flops required by the MDL, SS/FBSS-based MDL, and the MENSE algorithm as f_{MDL}, $f_{\mathrm{SS-MDL}}$, and f_{MENSE}, we can define the relative efficiency ratios as $f_{\mathrm{MENSE}}/f_{\mathrm{MDL}}$ and $f_{\mathrm{MENSE}}/f_{\mathrm{SS-MDL}}$, and these ratios in terms of the number of sensors M are shown in Fig. 1, where M is verified from $M=5$ to $M=100$, while $N=100$ and $m=\bar{p}=\lfloor M/2 \rfloor$. When the number of snapshots N is not significantly larger than the number of sensors M, the computational complexity is mostly dominated by the calculation of needed correlation matrix and the EVD or QR operation, and hence we approximately get

$$\frac{f_{\mathrm{MENSE}}}{f_{\mathrm{MDL}}} \approx \frac{5.75M^3+15M^2}{8M^3+\mathcal{O}(M^3)+7M^2}$$
$$\frac{f_{\mathrm{MENSE}}}{f_{\mathrm{SS-MDL}}} \approx \frac{5.75M^3+15M^2}{6.25M^3+\mathcal{O}(0.5M^3)+14M^2}$$

for $N \approx M$ and $m=\bar{p} \approx M/2$. When $N \gg M$, the computations of EVD or QR operation will be also are negligible, and we easily obtain $f_{\mathrm{MENSE}}/f_{\mathrm{MDL}} \approx 2/M$ and $f_{\mathrm{MENSE}}/f_{\mathrm{SS-MDL}} \approx 2M/(M+(2M-m)(m-1))$. Obviously f_{MENSE} is smaller than f_{MDL} and $f_{\mathrm{SS-MDL}}$, and these quantitative comparisons show that the MENSE algorithm is computationally efficient than the (SS/FBSS-based) MDL method with EVD. The computational complexity of the (SS/FBSS-based) AIC method is almost identical that of the (SS/FBSS-based) MDL method, except that the flops of last step is given by $M(M+13)$ (or $m(m+13)$), hence the behaviors of $f_{\mathrm{MENSE}}/f_{\mathrm{AIC}}$ and $f_{\mathrm{MENSE}}/f_{\mathrm{SS-AIC}}$ are similar to that shown in Fig. 1 and are omitted herein. □

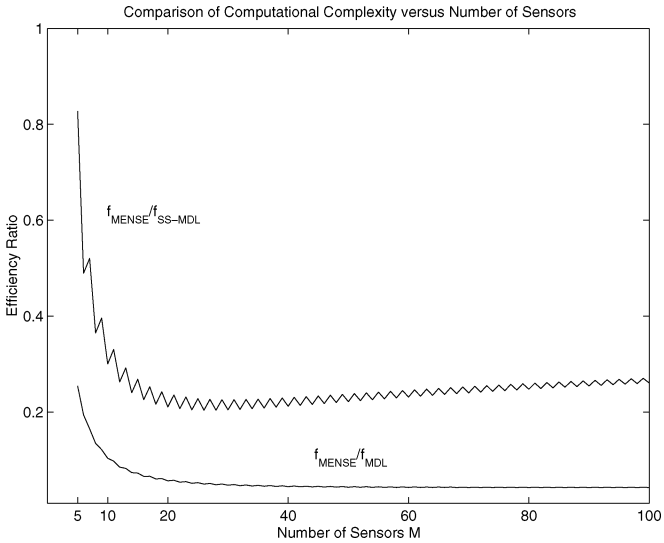

Fig. 1. Relative efficiency ratios between the estimated number of MATLAB flops required by the MENSE algorithm and that needed by the MDL and SS/FBSS-based MDL methods in terms of the number of sensors ($M=100$ and $m=\bar{p}=\lfloor M/2 \rfloor$).

IV. STATISTICAL ANALYSIS

A. Consistency of MENSE Detection Criterion

Because the finite-sample properties of detection methods are difficult to study analytically, an important characteristic of a detection scheme is its ability to provide an unbiased estimate of the number of signals for a large number of snapshots N. For investigating the consistency of the MENSE criterion, we first consider the asymptotical error of $\hat{\mathbf{\Psi}}$.

Theorem 2: By dividing the total array into $\bar{p}+1$ overlapping "virtual" forward and backward subarrays with $M-\bar{p}$ sensors, the asymptotical error of the autoproduct $\hat{\mathbf{\Psi}}$ in (30) is given by

$$E\{\hat{\mathbf{\Psi}}-\mathbf{\Psi}\} = \frac{1}{N}\left(c_{MM}\left(\boldsymbol{C}_1^f+\boldsymbol{C}_2^b\right)+c_{11}\left(\boldsymbol{C}_2^f+\boldsymbol{C}_1^b\right)\right) \quad (40)$$

where $\boldsymbol{C}_1^f$ and $\boldsymbol{C}_2^f$ are the spatially summed covariance matrices of the first and last $\bar{p}$ "virtual" forward subarrays, while $\boldsymbol{C}_1^b$ and $\boldsymbol{C}_2^b$ are those of the first and last $\bar{p}$ "virtual" backward subarrays.

Proof: See Appendix B. ■

Thus, as the number of snapshots N tends to infinity, the asymptotical error of the estimated autoproduct $\hat{\boldsymbol{\Psi}}$ in (30) approaches zero. Then, we easily obtain the following theorem on the consistency of the proposed criterion.

Theorem 3: As the number of snapshots N tends toward infinity, the estimated number of incident signals obtained with the MENSE detection criterion is consistent.

Proof: See Appendix C. ■

B. Asymptotical Threshold for Detection

Now, by defining an asymptotical autoproduct $\boldsymbol{\Psi}_{\text{as}}$ of the sample estimate $\hat{\boldsymbol{\Psi}}$ in (30) as its expectation and substituting (B5)–(B8) into (40), we obtain

$$\boldsymbol{\Psi}_{\text{as}} = \bar{\boldsymbol{\Psi}} + \frac{1}{N}\left(c_{MM}\left(\bar{\boldsymbol{C}}_1^f + \bar{\boldsymbol{C}}_2^b\right) + c_{11}\left(\bar{\boldsymbol{C}}_2^f + \bar{\boldsymbol{C}}_1^b\right) + 2\bar{p}\sigma^2(c_{MM} + c_{11})\boldsymbol{I}_{M-\bar{p}}\right) \quad (41)$$

where $c_{ii} = c_s|\bar{\rho}_i|^2 + \sigma^2$, and the ranks of four noise-free summed "virtual" subarray covariance matrices equal the number of signals p. Then, from the facts that $\text{rank}(\boldsymbol{AB}) \le \min(\text{rank}(\boldsymbol{A}), \text{rank}(\boldsymbol{B}))$ and $\text{rank}(\boldsymbol{A}+\boldsymbol{B}) \le \text{rank}(\boldsymbol{A})+\text{rank}(\boldsymbol{B})$ (e.g., [48]), the rank of matrix $\boldsymbol{\Psi}_{\text{as}}$ is given by

$$\text{rank}(\bar{\boldsymbol{\Psi}}_{\text{as}}) = \begin{cases} p, & \text{for } \sigma^2 = 0 \text{ and/or } N \to \infty \\ M - \bar{p}, & \text{others} \end{cases}. \quad (42)$$

Evidently, the estimated number of signals is perturbed by the noise variance σ^2 and the number of snapshots N. Thus, by substituting $\boldsymbol{\Psi}_{\text{as}}$ into the proposed algorithm, we can obtain a theoretical detection threshold for the SNR, number of snapshots, or angular separation for correct detection and gain some insights into the detection performance.

Especially when no incident signal is present, from (24) and (41), we easily get

$$c_{11} = \sigma^2, \quad \text{and} \quad \boldsymbol{\Psi}_{\text{as}} = \frac{4}{N}\bar{p}\sigma^4\boldsymbol{I}_{M-\bar{p}} \quad (43)$$

and the element r_{11}^{as} of the QR factor $\boldsymbol{R}_{\text{as}}$ of $\boldsymbol{\Psi}_{\text{as}}$ is given by

$$r_{11}^{as} = \frac{4}{N}\bar{p}\sigma^4. \quad (44)$$

Then, in view of the hypothesis $\mathcal{H}_0$ given in (28) and (29), we can theoretically obtain the low threshold for the number of snapshots $\bar{N}$ for correctly detecting the absence of incident signal(s) as

$$\bar{N} > 4\bar{p}\sigma^4. \quad (45)$$

C. Choice of QR Permutation Matrix

Although the permutation matrix $\boldsymbol{\Pi}$ in (31) could be determined by a successive dynamic swapping of the columns of $\hat{\boldsymbol{\Psi}}$ during the process of QR decomposition according to the QR decomposition with column pivoting (QRP) [33] or RRQR [43], this data-dependent pivoting scheme may be undesirable for some real-time applications due to its extra computational cost [41]. To avoid the dynamic column shuffling procedure, here we propose a way of predetermining $\boldsymbol{\Pi}$ by considering the asymptotical autoproduct $\boldsymbol{\Psi}_{\text{as}}$ in (41).

By reexpressing the matrix $\boldsymbol{\Psi}_{\text{as}}$ in its column vectors $\{\bar{\boldsymbol{\psi}}_i\}$, the Gramian matrix $\boldsymbol{\Lambda}$ of $\boldsymbol{\Psi}_{\text{as}}$ is given by

$$\boldsymbol{\Lambda} \triangleq \boldsymbol{\Psi}_{\text{as}}^H\boldsymbol{\Psi}_{\text{as}} = \{\lambda_{ik}\} \quad (46)$$

where $\lambda_{ik} \triangleq \bar{\boldsymbol{\psi}}_i^H\bar{\boldsymbol{\psi}}_k$, and $\boldsymbol{\Lambda}$ is centrosymmetric, Hermitian, and persymmetric as well as $\boldsymbol{\Psi}_{\text{as}}$. Then a quantitative measure of linear independence of columns $\{\bar{\boldsymbol{\psi}}_i\}$ is the dependency coefficient $\bar{\lambda}_{ik}$ defined as

$$\bar{\lambda}_{ik} \triangleq \frac{|\lambda_{ik}|}{\sqrt{\lambda_{ii}\lambda_{kk}}} \quad (47)$$

where $\bar{\lambda}_{ik} = \bar{\lambda}_{ki} = \bar{\lambda}_{M-\bar{p}-k,M-\bar{p}-i} = \bar{\lambda}_{M-\bar{p}-i,M-\bar{p}-k}$, and $\bar{\lambda}_{ik} \le 1$ under Cauchy-Schwarz inequality. Thus, by comparing the elements of a dependency coefficient matrix $\bar{\boldsymbol{\Lambda}} \triangleq \{\bar{\lambda}_{ik}\}$, we can form a permutation matrix $\boldsymbol{\Pi}$ to ensure the minimum linear dependency between the adjacent columns of $\boldsymbol{\Psi}_{\text{as}}\boldsymbol{\Pi}$. Further, this $\boldsymbol{\Pi}$ (referred as the QRPA) can be used in (31) to improve the detection performance by lowering the detection threshold without increased computational complexity.

Remark 7: It should be emphasized that the proposed determination of the permutation matrix may not be feasible in some applications, when the knowledge of asymptotical autoproduce $\boldsymbol{\Psi}_{\text{as}}$ is unavailable. Then, an alternative *a priori* manner is the column index maximum-difference bisection rule-based scheme [37], [41] (referred as the QRPP), where $\boldsymbol{\Pi}$ is given by the equation, shown at the bottom of the page. This data-independent scheme could possibly provide the shuffled columns with relatively small dependency by taking the special symmetries of the matrices $\boldsymbol{\Psi}$ (i.e., $\hat{\boldsymbol{\Psi}}$) and $\bar{\boldsymbol{\Lambda}}$ into account. □

$$\boldsymbol{\Pi} = \begin{cases} \left[\boldsymbol{e}_1, \boldsymbol{e}_{M-\bar{p}}, \boldsymbol{e}_2, \boldsymbol{e}_{M-\bar{p}-1}, \cdots, \boldsymbol{e}_{(M-\bar{p})/2}, \boldsymbol{e}_{(M-\bar{p})/2+1}\right], & \text{for } M-\bar{p}: \text{even} \\ \left[\boldsymbol{e}_1, \boldsymbol{e}_{M-\bar{p}}, \boldsymbol{e}_{\lceil (M-\bar{p})/2 \rceil}, \boldsymbol{e}_2, \boldsymbol{e}_{M-\bar{p}-1}, \cdots, \boldsymbol{e}_{\lfloor (M-\bar{p})/2 \rfloor}, \boldsymbol{e}_{\lceil (M-\bar{p})/2 \rceil+1}\right], & \text{for } M-\bar{p}: \text{odd} \end{cases}.$$

V. Numerical Examples

The effectiveness of the proposed MENSE algorithm in estimating the number of signals is evaluated through numerical examples. The ULA with M sensors is separated by a half-wavelength, and the SNR is defined as the ratio of the power of the source signals to that of the additive noise at each sensor. The AIC and MDL methods with EVD [7], [10], the SS- and FBSS-based AIC and MDL methods [12], [27], [31], and the QR-based MDL method [37], which is modified by combining the FBSS-MDL criterion [12], [31] with the QRPP decomposition [37], [41], are carried out for performance comparison. The simulation results shown below are based on 1000 independent trials.

Example 1—Performance Versus SNR: First we examine the detection performance of the MENSE algorithm with respect to the SNR, where two coherent signals with equal power $(c_s = 1)$ arrive from $\theta_1 = 5^\circ$ and $\theta_2 = 12^\circ$, and their SNR is varied from -10 to 15 dB. The numbers of sensors and of snapshots are $M = 10$ and $N = 64$. The subarray sizes are set at $\bar{p} = \lfloor M/2 \rfloor = 5$ and $m = 5$ for the MENSE and the SS/FBSS-based methods.

By examining the asymptotical autoproduct $\mathbf{\Psi}_{\rm as}$ in (41), the QRPA predetermined permutation matrix is given by $\mathbf{\Pi} = [\boldsymbol{e}_1, \boldsymbol{e}_5, \boldsymbol{e}_2, \boldsymbol{e}_4, \boldsymbol{e}_3]$. The probabilities of correct detection (i.e., $\hat{p} = p$) of the MENSE algorithm with QRPA, QRPP [37] (i.e., $\mathbf{\Pi} = [\boldsymbol{e}_1, \boldsymbol{e}_5, \boldsymbol{e}_3, \boldsymbol{e}_2, \boldsymbol{e}_4]$), QRP [33], and QR (i.e., $\mathbf{\Pi} = \boldsymbol{I}_{M-\bar{p}}$) in terms of the SNR are shown in Fig. 2. The performance of the SS-based AIC and MDL methods is poor at lower and medium SNRs due to the bias in the estimated eigenvalues, whereas the detection of the FBSS-based AIC and MDL methods is considerably enhanced because of the forward-backward averaging, which effectively doubles that the amount of available data and improves the estimation of eigenvalues [12]. The QR-based MDL method performs slightly better than the FBSS-based MDL method but worse than the FBSS-based AIC method. The MENSE algorithm with QR generally outperforms the SS/FBSS-based methods and the QR-based MDL method because of the reduced effect of additive noise, while it is slightly inferior to the FBSS-based AIC method at medium SNR. Further, by introducing the column pivoting into the QR decomposition to remedy the effect of additive noise, the MENSE algorithm with the QRP, QRPP, and QRPA can be significantly improved at lower SNR, and the proposed QRPA can ensure efficient detection similarly to the QRPP [37].

Moreover, by considering matrix $\mathbf{\Psi}_{\rm as}$ in (41), as shown in Fig. 2, we find that the low threshold $\overline{\rm SNR}$ for the MENSE with QR, QRP, or QRPP/QRPA is given by -3, -2, or -9 dB, respectively. Here the difference between the actual and theoretical thresholds for correct detection results from the statistical fluctuations, and the theoretical threshold based on the asymptotical analysis usually corresponds to the 40%$\sim$60% range of the ensemble-averaged probability of correct detection. Additionally the probabilities of missing and false alarm (i.e., $P_{\rm m}$ and $P_{\rm fa}$ defined in Appendix C) of the MENSE algorithm with QR permutation matrices versus the SNR are depicted in Fig. 3. Clearly $P_{\rm fa}$ is much smaller than $P_{\rm m}$ at low SNR, and these probabilities of incorrect detection can be decreased significantly at low SNR when the QRPP [37] and the proposed QRPA are used.

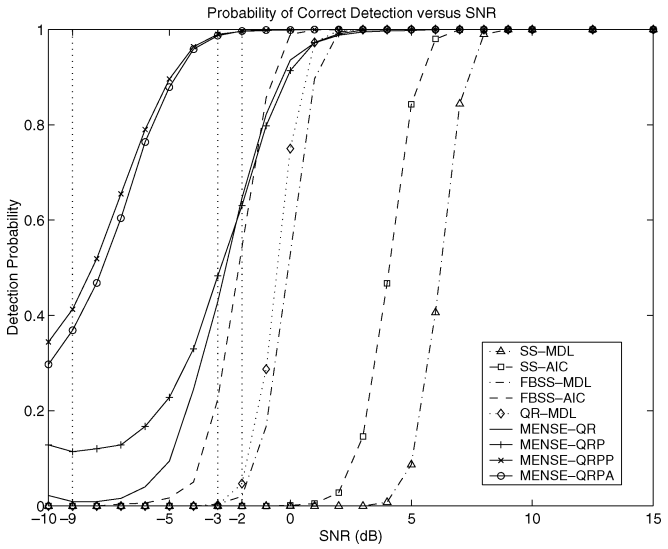

Fig. 2. Probability of correct detection versus the SNR for Example 1 (vertical dotted line: detection threshold; $N = 64$, $M = 10$, and $p = 2$).

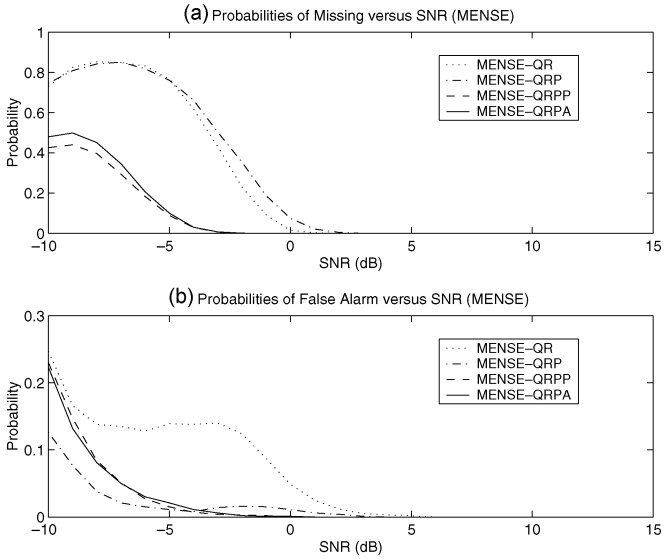

Fig. 3. (a) Probability of missing and (b) probability of false alarm versus the SNR with different QR permutation matrices for Example 1 ($N = 64$, $M = 10$, and $p = 2$).

Example 2—Performance Versus Number of Snapshots: Now we inspect the performance of the MENSE against the number of snapshots with the similar simulation conditions to those of Example 1, except that the SNR is set at 2.5 dB, and the number of snapshots is varied from $N = 1$ to $N = 1000$.

As shown in Fig. 4, evidently the small number of snapshots causes the estimated eigenvalues to be inaccurate and hence leads to degraded detection with the SS-based AIC and MDL methods. The MENSE algorithm with QR or QRP generally performs better than the FBSS-based MDL method and suffers

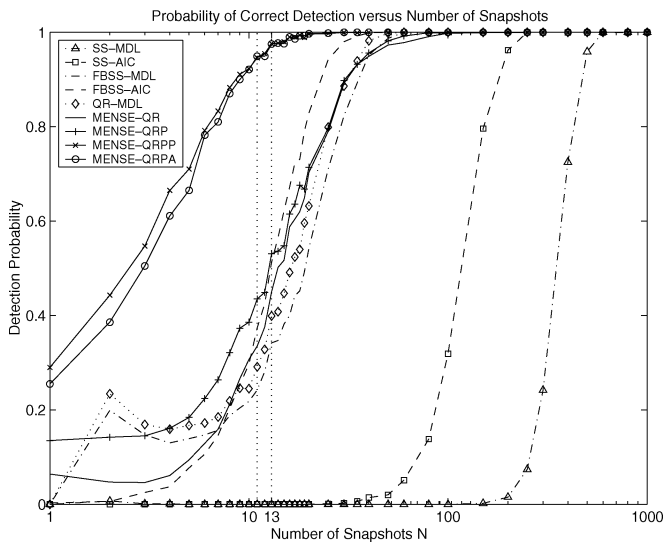

Fig. 4. Probability of correct detection versus the number of snapshots for Example 2 (vertical dotted line: detection threshold; SNR = 2.5 dB, $M = 10$, and $p = 2$).

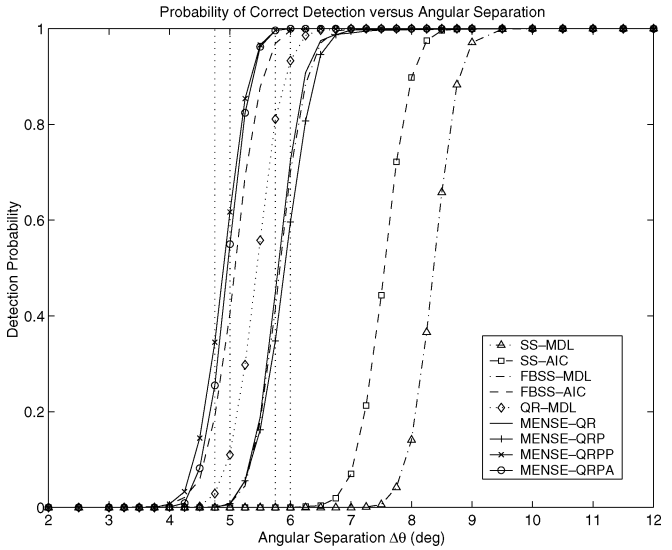

Fig. 5. Probability of correct detection versus the angular separation for Example 3 (vertical dotted line: detection threshold; SNR = 2.5 dB, $N = 64$, $M = 10$, and $p = 2$).

only slight degradation as compared with the FBSS-based AIC method for a relatively small number of snapshots. However, the MENSE algorithm with QRPP/QRPA is significantly superior to the other methods even with a small number of snapshots, where the QRPP or QRPA well offsets the influence of a small number of snapshots. Further, its probability of correct detection approaches one promptly even though the number of snapshots N is not "significantly large." In addition, the theoretical threshold for the number of snapshots is effectively reduced from $\bar{N} = 11$ (with QR) or $\bar{N} = 13$ (with QRP) to $\bar{N} = 1$ in this empirical scenario.

Example 3—Performance Versus Angular Separation: Here, we assess the performance of the MENSE with respect to the angular separation between two coherent signals. In this example, two coherent signals impinge on the array along $\theta_1 = 5^\circ$ and $\theta_2 = \theta_1 + \Delta\theta$, where $\Delta\theta$ is varified from $\Delta\theta = 2^\circ$ to $\Delta\theta = 12^\circ$, and the other simulation conditions are the same as those in Example 1, except that the SNR is fixed at 2.5 dB.

Although the MENSE algorithm with QR or QRP performs similarly to the FBSS-based MDL method, from Fig. 5, we can see that the MENSE algorithm with QRPP or QRPA can efficiently detect the closely spaced coherent signals and outperforms the other methods. We also find that the theoretical threshold for angular separation $\overline{\Delta\theta}$ is decreased from $\overline{\Delta\theta} = 6^\circ$ (with QR) or $\overline{\Delta\theta} = 5.75^\circ$ (with QRP) to $\overline{\Delta\theta} = 5^\circ$ (with QRPP) or $\overline{\Delta\theta} = 4.75^\circ$ (with QRPA) by studying the asymptotical autoproduct $\mathbf{\Psi}_{\rm as}$.

Example 4—Performance Versus Subarray Size: We then study the effect of the subarray size on the MENSE detection performance. The simulation conditions are similar to those in Example 1, except that the SNR is set at 2.5 dB, and the subarray size $\bar{p}$ is varied from $\bar{p} = 2$ to $\bar{p} = 7$ to satisfy the condition $p \le \bar{p} < M - p$, while the subarray size m for the SS/FBSS-based methods is varied $m = 3$ to $m = 10$, and

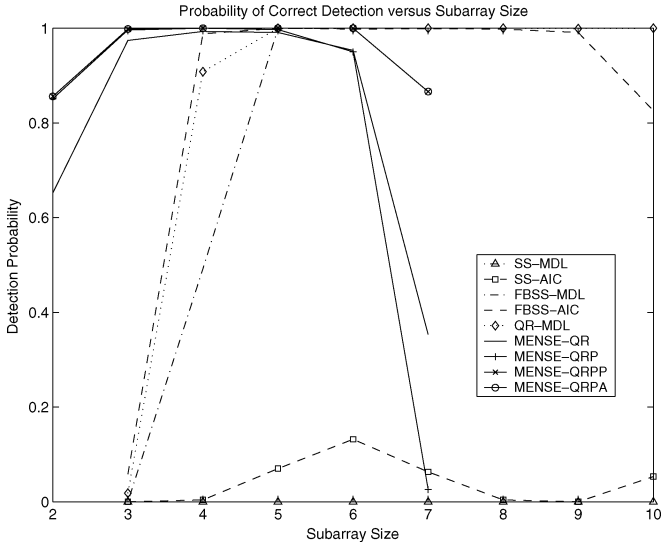

Fig. 6. Probability of correct detection versus the subarray size for Example 4 (SNR = 2.5 dB, $N = 64$, $M = 10$, and $p = 2$).

$\mathbf{\Pi} = [\boldsymbol{e}_1, \boldsymbol{e}_{M-\bar{p}}, \boldsymbol{e}_2, \boldsymbol{e}_{M-\bar{p}-1}, \boldsymbol{e}_3, \cdots]$ is chosen as the QRPA permutation matrix for the sake of simplicity.

Fig. 6 plots the probabilities of correct detection against the subarray size $\bar{p}$ or m. As the dimension $(M - \bar{p}) \times (M - \bar{p})$ of matrix $\hat{\mathbf{\Psi}}$ varies with the subarray size $\bar{p}$, the coherency decorrelation and the detection capability of the MENSE are affected by the subarray size similarly to that of the SS/FBSS-based methods, when the number of snapshots is small and/or the SNR is low. Clearly, a high probability of correct detection can be achieved by the MENSE algorithm with QRPA and QRPP for $\bar{p} = 3$, 4, 5, and 6 (the differences are almost indistinguishable), and the setting of subarray size $\bar{p} = \lfloor M/2 \rfloor$ is an appropriate and reasonable choice.

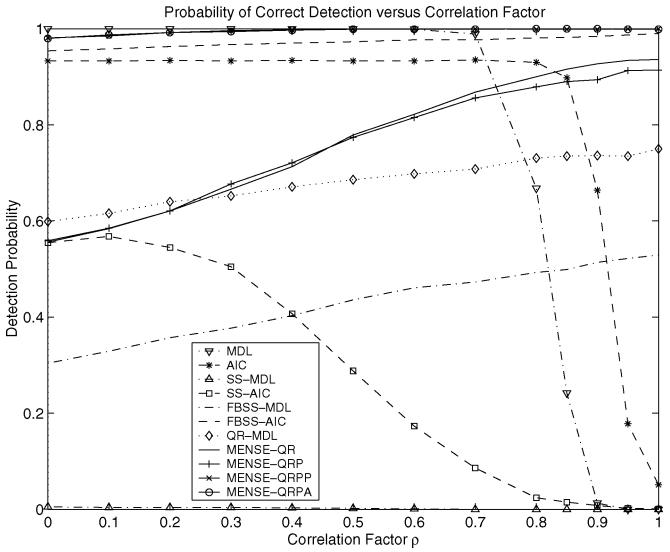

Fig. 7. Probability of correct detection versus the correlation factor for Example 5 (SNR = 0 dB, $N = 64$, $M = 10$, and $p = 2$).

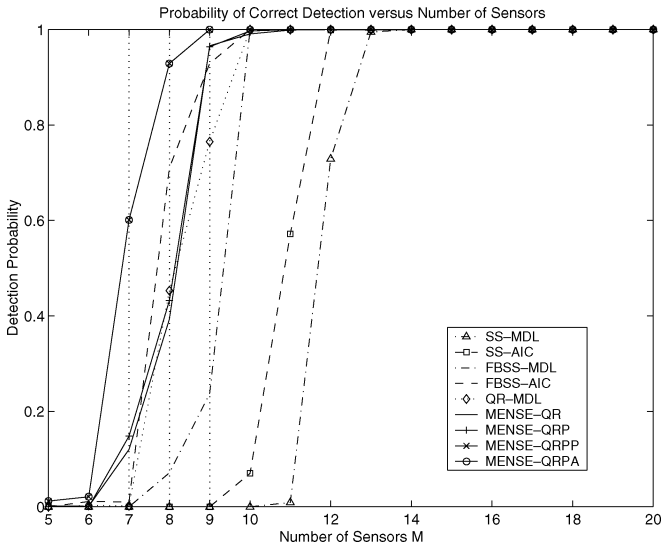

Fig. 8. Probability of correct detection versus the number of sensors for Example 6 (vertical dotted line: detection threshold; SNR = 2.5 dB, $N = 64$, and $p = 2$).

Example 5—Performance Versus Correlation Factor: Then, we evaluate the detection performance with respect to the correlation between the incident signals, where the correlation factor ρ between $s_1(n)$ and $s_2(n)$ is defined by [56]

$$\rho \triangleq \frac{E\{s_1(n)s_2^*(n)\}}{\sqrt{E\left\{|s_1(n)|^2\right\}E\left\{|s_2(n)|^2\right\}}}$$

where $|\rho| \leq 1$. Here, the correlation factor ρ is varied from 0 to 1 (its phase is assumed to be zero), the SNR is set at 0 dB, and the number of snapshots is fixed at $N = 64$, while the other parameters are the same as those in Example 1.

The simulation results in terms of the correlation factor are shown in Fig. 7. Because the maximum possible array aperture (i.e., $M = 10$) is used, the MDL method provides the best detection for uncorrelated and weakly correlated signals. But its performance degrades sharply with the increased signal correlation as that of the AIC method, whereas the SS-based AIC and MDL methods perform poorly due to the smaller array aperture (i.e., $m = 5$) and the slightly strict conditions of the low SNR and the small number of snapshots. While the FBSS-based AIC and MDL and the QR-based MDL methods perform better than the SS-based AIC and MDL methods for correlated signals, their probabilities of correct detection do not tend toward one despite the use of forward-backward averaging. Although the MENSE with QR or QRP is inferior to the FBSS-based AIC method, the MENSE algorithm with QRPP or QRPA is superior to the other methods in detecting strongly correlated signals, where the slight degradation for low correlations is due to the reduced dimension of the $(M - \bar{p}) \times (M - \bar{p})$ sample matrix $\hat{\boldsymbol{\Psi}}$, where $M - \bar{p} = 5 < M$. Thus, we can see that the MENSE with QRPP/QRPA is insensitive to the correlation between incident signals.

Example 6—Performance Versus Number of Sensors: Next, we test the detection performance against the number of sensors, where the number of sensors is varied from $M = 5$ to $M = 20$, and $\bar{p} = m = \lfloor M/2 \rfloor$ are chosen as the subarray sizes. The other simulation parameters are identical to those in Example 1, except that the SNR is fixed at 2.5 dB, and the QRPA permutation matrix is given by $\boldsymbol{\Pi} = [\boldsymbol{e}_1, \boldsymbol{e}_{M-\bar{p}}, \boldsymbol{e}_2, \boldsymbol{e}_{M-\bar{p}-1}, \boldsymbol{e}_3, \cdots]$ for the sake of simplicity.

As shown in Fig. 8, the MENSE algorithm with QRPP/QRPA has a higher probability of correct detection than the other methods for a small number of sensors. Further the probability of correct detection tends toward one when $M \geq 11$ no matter which QR permutation matrix (i.e., the QR, QRP, or QRPP/QRPA) is used, and the theoretical threshold for the number of sensors is given by $\bar{M} = 9$, 8, or 7, respectively.

Example 7—Detecting Absence of Incident Signal: Here we verify the performance of the MENSE in detecting the absence of impinging signals. The number of sensors is $M = 10$, and the number of snapshots is varied from $N = 1$ to $N = 100$.

The probability of false alarm P_{fa} of the MENSE with QRPP/QR versus the number of snapshots is plotted in Fig. 9 for several noise variances $\sigma^2 = 1$, 0.5, 0.3, and 0.2. Evidently, the AIC and MDL methods are not affected by the noise variance, but they pose an unacceptably high risk of false alarm for a small number of snapshots N (i.e., $N < M$) due to the bias in the estimated eigenvalues; further the AIC's P_{fm} tends toward 5% even for a larger number of snapshots. Generally, the MENSE algorithm has a rather small P_{fm} and outperforms for a smaller number of snapshots when $\sigma^2 \leq 0.3$. Additionally the theoretical threshold for the number of snapshots $\bar{N}$ is given by $\bar{N} = 21$, 11, 7, and 5 for the noise variances $\sigma^2 = 1$, 0.5, 0.3, and 0.2, respectively.

Example 8—Performance in the Presence of Array Uncertainties: Now we consider the performance of the MENSE versus the SNR, when the sensor gain and phase errors and the mutual coupling between sensors exist, which are often encountered in many practical situations (e.g., [51]). The received signal model in (2) is modified as

$$\boldsymbol{y}(n) = \boldsymbol{M}\boldsymbol{\Gamma}\boldsymbol{A}\boldsymbol{s}(n) + \boldsymbol{w}(n)$$

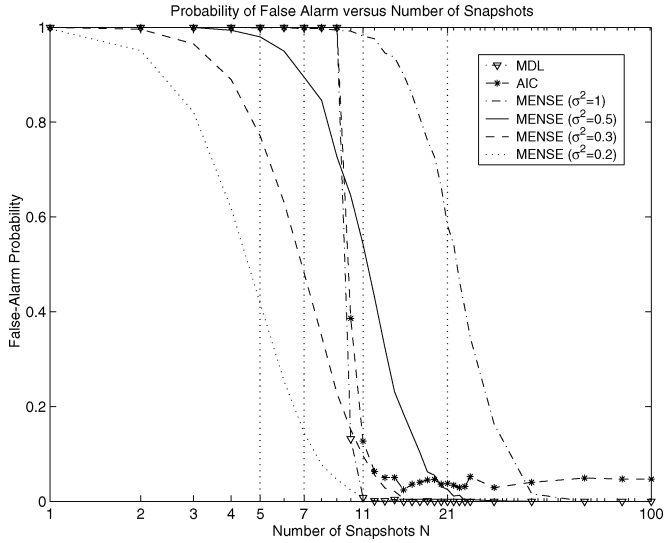

Fig. 9. Probability of false alarm versus the number of snapshots for several noise variances for Example 7 (vertical dotted line: detection threshold; $M = 10$, and $p = 0$).

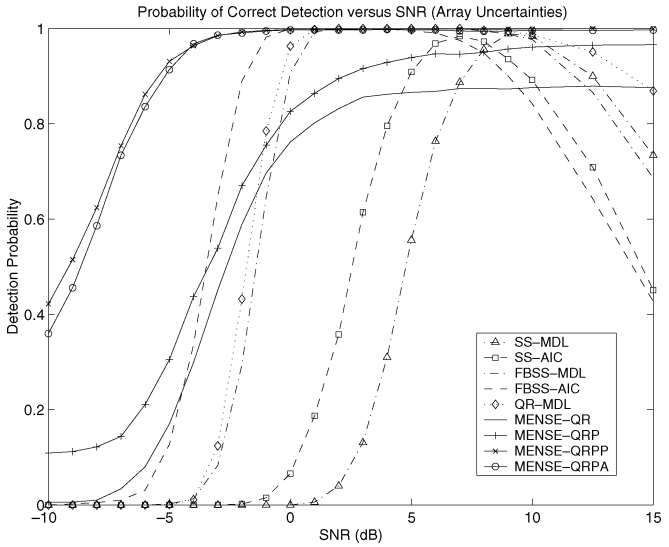

Fig. 10. Probability of correct detection versus the SNR in the presence of array uncertainties for Example 8 ($\sigma_\alpha = 0.1$, $\sigma_\phi = 5^\circ$, $\boldsymbol{\mu} = [1, 0.1, 0.025, 0, \cdots, 0]^T$, $N = 64$, $M = 10$, and $p = 2$).

where $\boldsymbol{M}$ is a banded Toeplitz matrix representing the effect of mutual coupling between the sensors, $\boldsymbol{\Gamma}$ is a diagonal matrix with the elements $\{\gamma_i\}$, $\gamma_i \triangleq \alpha_i e^{-j\phi_i}$, and α_i and ϕ_i denote the gain and phase of each sensor, which can be chosen according to the following relations [51]:

$$\alpha_i = \sqrt{12}\sigma_\alpha \bar{\alpha}_i + 1 \quad \text{and} \quad \phi_i = \sqrt{12}\sigma_\phi \bar{\phi}_i$$

where $\{\bar{\alpha}_i\}$ and $\{\bar{\phi}_i\}$ are independent random variables distributed uniformly in the interval $(-0.5, 0.5)$, and σ_α and σ_ϕ are the standard deviations of the gain α_i and phase ϕ_i. In this example, the Toeplitz mutual coupling matrix $\boldsymbol{M}$ has the first column as $\boldsymbol{\mu} = [1, 0.1, 0.025, 0, \cdots, 0]^T$, $\sigma_\alpha = 0.1$ and $\sigma_\phi = 5^\circ$, while the other simulation parameters are similar to those of Example 1.

The probabilities of correct detection versus the SNR are shown in Fig. 10. Clearly regardless of the sensor gain and phase errors and the mutual coupling effect between three adjacent sensors, the MENSE algorithm is robust to these uncertainties. Further the MENSE algorithm with QRPP and QRPA can determine the number of signals from the uncalibrated array data with high probability of correct detection and outperform the other methods even at low SNR.

Example 9—Performance Under Spatially Inhomogenous Noise: Finally we investigate the performance of the MENSE in the presence of the deviations from the spatial homogeneity of noise model given by (5), which appears in some practical applications (e.g., [52]–[54]), where the noise variance matrix is given by

$$\boldsymbol{C}_w = E\left\{\boldsymbol{w}(n)\boldsymbol{w}^H(n)\right\} = (\sigma \boldsymbol{I}_M + \nu \boldsymbol{K})^2$$

where ν is a positive constant, $\boldsymbol{K}$ is a diagonal matrix with the elements $\{\kappa_i\}$, and $\{\kappa_i\}$ are independent random variables distributed uniformly over the interval (0, 1). Here, we set $\nu = 1$, and the simulation conditions are similar to that of Example 1,

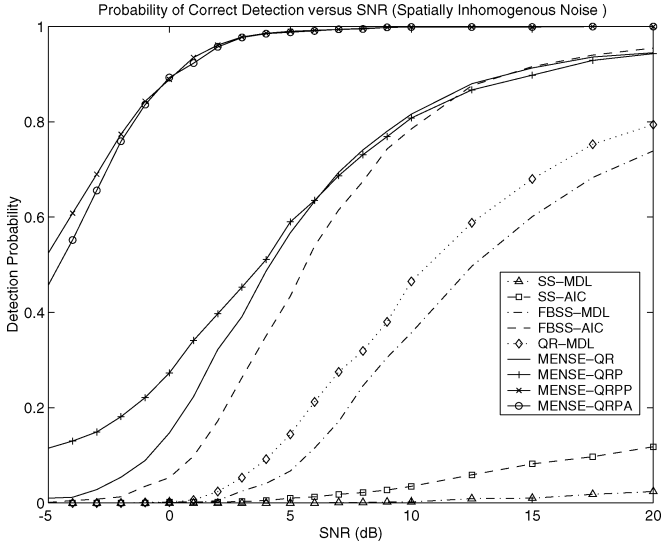

Fig. 11. Probability of correct detection versus the SNR for spatially inhomogenous noise model for Example 9 ($\nu = 1$, $N = 64$, $M = 10$, and $p = 2$).

except that the SNR is varied from -5 to 20 dB, where the SNR is calculated when $\nu = 0$ [54].

Fig. 11 depicts the probability of correct detection with respect to the SNR for this nonideal noise model. Because the AIC and MDL criterion can be interpreted as a test for determining the multiplicity of the smallest eigenvalues [7], and hence they performs poorly when the noise variance at each sensor is unequal [16], [17], [52]–[54], where the relation between the number of signals and this multiple becomes invalid. However, since the autocorrelation c_{ii} of array data, which contains the noise variance at each sensor, is not used in the combined Hankel correlation matrix $\boldsymbol{\Phi}$ and hence its autoproduct $\boldsymbol{\Psi}$ in principle, the MENSE algorithm is robust to the deviations from the normal noise model. As shown in Fig. 11, although the number of snapshots is small, we can see that the MENSE algorithm performs better than the (FBSS/SS-based) MDL and AIC

methods under the spatially inhomogenous noise model when the SNR is relatively low.

VI. Conclusion

A computationally efficient QR-based nonparametric method called MENSE was proposed for estimating the number of narrowband signals impinging on a ULA. The MENSE algorithm does not require the computationally cumbersome eigendecomposition and the evaluation of all correlations of the array data, where the effect of additive noise is remedied by appropriately choosing the used subarrays. Consequently, the MENSE algorithm is suitable for real-time implementation and is remarkably insensitive to the correlation of incident signals and flexible with the spatially correlated noise. The asymptotical consistency of the MENSE estimator was studied, and its ability to detection could be predicted by examining the QR decomposition of the asymptotical autoproduct matrix with different permutation matrix. The simulation results showed that the MENSE algorithm with a predetermined QR permutation matrix such as QRPA or QRPP is superior in detecting closely spaced signals with a small number of snapshots and/or at relatively low SNR and is robust against the array uncertainties and the deviations from the spatial homogenous assumption of additive noise.

Appendix A
Proof of Theorem 1

It follows from the basic assumptions that the ith row $\boldsymbol{b}_i^T(\theta)$ of matrix $\boldsymbol{A}$ in (2) can be expressed as a linear combination of the first p rows $\boldsymbol{b}_1(\theta), \boldsymbol{b}_2(\theta), \cdots, \boldsymbol{b}_p(\theta)$ with p complex coefficients $\gamma_{i1}, \gamma_{i2}, \cdots, \gamma_{ip}$ which are not all zero for $i = p+1, \cdots, M$, i.e.,

$$\boldsymbol{b}_i(\theta) = \gamma_{i1}^* \boldsymbol{b}_1(\theta) + \gamma_{i2}^* \boldsymbol{b}_2(\theta) + \cdots + \gamma_{ip}^* \boldsymbol{b}_p(\theta). \tag{A1}$$

By defining $\tilde{\boldsymbol{A}}$ be the $p \times p$ submatrix of $\boldsymbol{A}$ in (2) consisting of the first p rows, i.e., $\tilde{\boldsymbol{A}} \triangleq [\boldsymbol{b}_1(\theta), \boldsymbol{b}_2(\theta), \cdots, \boldsymbol{b}_p(\theta)]^T$, from (A1), we easily obtain the coefficients $\{\gamma_{ik}\}$

$$\boldsymbol{\gamma}_i = \tilde{\boldsymbol{A}}^{-H} \boldsymbol{b}_i^*(\theta) \tag{A2}$$

where $\boldsymbol{\gamma}_i \triangleq [\gamma_{i1}, \gamma_{i2}, \cdots, \gamma_{ip}]^T$, and the inverse matrix $\tilde{\boldsymbol{A}}^{-1}$ of the full rank matrix $\tilde{\boldsymbol{A}}$ exists. When the condition that $p \le \bar{p} < M - p$ is satisfied, from (18) and (A1), we obtain

$$\begin{aligned}\boldsymbol{\Psi} &= c_s^2 |\bar{\rho}_M|^2 \boldsymbol{T} [\boldsymbol{b}_1(\theta), \boldsymbol{b}_2(\theta), \cdots, \boldsymbol{b}_{M-\bar{p}}(\theta)]^* \\ &= c_s^2 |\bar{\rho}_M|^2 [\boldsymbol{g}_1, \boldsymbol{g}_2, \cdots, \boldsymbol{g}_p, \gamma_{p+1,1}\boldsymbol{g}_1 \\ &\quad + \gamma_{p+1,2}\boldsymbol{g}_2 + \cdots + \gamma_{p+1,p}\boldsymbol{g}_p, \cdots, \gamma_{M-\bar{p},1}\boldsymbol{g}_1 \\ &\quad + \gamma_{M-\bar{p},2}\boldsymbol{g}_2 + \cdots + \gamma_{M-\bar{p},p}\boldsymbol{g}_p]\end{aligned} \tag{A3}$$

where $\boldsymbol{T} \triangleq \bar{\boldsymbol{A}}\boldsymbol{B}\boldsymbol{F}\boldsymbol{F}^H\boldsymbol{B}^H$, $\boldsymbol{g}_k \triangleq \boldsymbol{T}\boldsymbol{b}_k^*(\theta) = [g_{k1}, g_{k2}, \cdots, g_{k,M-\bar{p}}]^T$, and (A1) is used implicitly.

Now we consider and choose the complex Householder transform matrix $\boldsymbol{Q}_1$ to zero out but the first element of the first column $\boldsymbol{g}_1$ of $\boldsymbol{\Psi}$ in (A3) as (see, e.g., [33] and [48])

$$\boldsymbol{Q}_1 \triangleq \boldsymbol{I}_{M-\bar{p}} - 2\frac{\boldsymbol{v}_1\boldsymbol{v}_1^H}{\boldsymbol{v}_1^H\boldsymbol{v}_1} \tag{A4}$$

where the Householder vector $\boldsymbol{v}_1$ is given by $\boldsymbol{v}_1 = \boldsymbol{g}_1 + \mathrm{sign}(g_{11})\|\boldsymbol{g}_1\|\boldsymbol{e}_1$, and $\boldsymbol{e}_1$ is an $(M-\bar{p}) \times 1$ unit vector with a unity element at the first location and zeros elsewhere, and the $\mathrm{sign}(x)$ for complex quantity x is chosen according the sign of its real part. Through some simple computations, we get

$$\boldsymbol{Q}_1\boldsymbol{g}_1 = -\mathrm{sign}(g_{11})\|\boldsymbol{g}_1\|\boldsymbol{e}_1. \tag{A5}$$

By defining $\boldsymbol{g}_k^{(1)} \triangleq \boldsymbol{Q}_1\boldsymbol{g}_k = [g_{k1}^{(1)}, g_{k2}^{(1)}, \cdots, g_{k,M-\bar{p}}^{(1)}]^T$ for $k = 1, 2, \cdots, p$ and pre-multiplying (A3) with $\boldsymbol{Q}_1$ in (A4), we have

$$\begin{aligned}\boldsymbol{Q}_1\boldsymbol{\Psi} = c_s^2|\bar{\rho}_M|^2 \Big[&g_{11}^{(1)}\boldsymbol{e}_1, \boldsymbol{g}_2^{(1)}, \cdots, \boldsymbol{g}_p^{(1)}, \gamma_{p+1,1}g_{11}^{(1)}\boldsymbol{e}_1 \\ &+ \gamma_{p+1,2}\boldsymbol{g}_2^{(1)} + \cdots + \gamma_{p+1,p}\boldsymbol{g}_p^{(1)}, \cdots, \\ &\gamma_{M-\bar{p},1}g_{11}^{(1)}\boldsymbol{e}_1 + \gamma_{M-\bar{p},2}\boldsymbol{g}_2^{(1)} + \cdots \\ &+ \gamma_{M-\bar{p},p}\boldsymbol{g}_p^{(1)}\Big]\end{aligned} \tag{A6}$$

where $g_{11}^{(1)} = -\mathrm{sign}(g_{11})\|\boldsymbol{g}_1\| \neq 0$. We then continue the process of QR decomposition by expressing the kth column of the matrix $\boldsymbol{Q}_{i-1}\cdots\boldsymbol{Q}_2\boldsymbol{Q}_1\boldsymbol{\Psi}$ as

$$\begin{aligned}\boldsymbol{g}_k^{(i-1)} &\triangleq \left[g_{k1}^{(1)}, g_{k2}^{(2)}, \cdots, g_{k,i-1}^{(i-1)}, g_{ki}^{(i-1)}, g_{k,i+1}^{(i-1)}, \cdots, g_{k,M-\bar{p}}^{(i-1)}\right]^T \\ &= \left[g_{k1}^{(1)}, g_{k2}^{(2)}, \cdots, g_{k,i-1}^{(i-1)}, \left(\bar{\boldsymbol{g}}_k^{(i-1)}\right)^T\right]^T\end{aligned} \tag{A7}$$

where $\bar{\boldsymbol{g}}_k^{(i-1)} \triangleq [g_{ki}^{(i-1)}, g_{k,i+1}^{(i-1)}, \cdots, g_{k,M-\bar{p}}^{(i-1)}]^T$, where $i = 2, 3, \cdots, p-1$, and $k = i, i+1, \cdots, p$. We can zero out the last $M - \bar{p} - i$ elements $g_{i,i+1}^{(i-1)}, g_{i,i+2}^{(i-1)}, \cdots, g_{i,M-\bar{p}}^{(i-1)}$ of the ith column $\bar{\boldsymbol{g}}_i^{(i-1)}$ of the matrix $\boldsymbol{Q}_{i-1}\cdots\boldsymbol{Q}_2\boldsymbol{Q}_1\boldsymbol{\Psi}$ by using the complex Householder transform matrix $\boldsymbol{Q}_i$ given by

$$\begin{aligned}\boldsymbol{Q}_i &\triangleq \boldsymbol{I}_{M-\bar{p}} - 2\frac{\tilde{\boldsymbol{v}}_1\tilde{\boldsymbol{v}}_1^H}{\tilde{\boldsymbol{v}}_1^H\tilde{\boldsymbol{v}}_1} \\ &= \begin{bmatrix}\boldsymbol{I}_{i-1}, & \boldsymbol{O}_{(i-1)\times(M-\bar{p}-i+1)} \\ \boldsymbol{O}_{(M-\bar{p}-i+1)\times(i-1)}, & \boldsymbol{H}_i\end{bmatrix}\end{aligned} \tag{A8}$$

for $i = 2, 3, \cdots, p$, where

$$\boldsymbol{H}_i \triangleq \boldsymbol{I}_{M-\bar{p}-i+1} - 2\frac{\boldsymbol{v}_i\boldsymbol{v}_i^H}{\boldsymbol{v}_i^H\boldsymbol{v}_i} \tag{A9}$$

$$\boldsymbol{v}_i \triangleq \bar{\boldsymbol{g}}_i^{(i-1)} + \mathrm{sign}\left(g_{ii}^{(i-1)}\right)\left\|\bar{\boldsymbol{g}}_i^{(i-1)}\right\|\boldsymbol{e}_1^{(i-1)} \tag{A10}$$

in which $\tilde{\boldsymbol{v}}_i \triangleq [\boldsymbol{0}_{(i-1)\times 1}^T, \boldsymbol{v}_i^T]^T$, and $\boldsymbol{e}_1^{(i-1)}$ is an $(M-\bar{p}-i)\times 1$ unit vector with a unity element at the first location and zeros elsewhere. As a result, we can obtain

$$\begin{aligned}\boldsymbol{Q}_i\boldsymbol{g}_k^{(i-1)} &= \left[g_{k1}^{(1)}, g_{k2}^{(2)}, \cdots, g_{k,i-1}^{(i-1)}, \left(\boldsymbol{H}_i\bar{\boldsymbol{g}}_k^{(i-1)}\right)^T\right]^T \\ &= \left[g_{k1}^{(1)}, g_{k2}^{(2)}, \cdots, g_{k,i-1}^{(i-1)}, g_{ki}^{(i)}, \left(\bar{\boldsymbol{g}}_k^{(i)}\right)^T\right]^T\end{aligned} \tag{A11}$$

for $k= i, i+1, \cdots, p$, where

$$\begin{aligned}\bar{\boldsymbol{g}}_k^{(i)} &\triangleq \boldsymbol{H}_i[g_{k,i+1}^{(i-1)}, g_{k,i+2}^{(i-1)}, \cdots, g_{k,M-\bar{p}}^{(i-1)}]^T \\ &\triangleq [g_{k,i+1}^{(i)}, g_{k,i+2}^{(i)}, \cdots, g_{k,M-\bar{p}}^{(i)}]^T\end{aligned}$$

and especially

$$g_{ii}^{(i)} = -\operatorname{sign}\left(g_{ii}^{(i-1)}\right)\left\|\bar{\boldsymbol{g}}_i^{(i-1)}\right\| \neq 0 \tag{A12}$$

$$g_{i,i+1}^{(i)} = g_{i,i+2}^{(i)} = \cdots = g_{i,M-\bar{p}}^{(i)} = 0. \tag{A13}$$

Therefore, by forming a sequence of transformations $\boldsymbol{Q}_1, \boldsymbol{Q}_2, \cdots, \boldsymbol{Q}_p$, from (A11)–(A13), the QR factors of $\boldsymbol{\Psi}$ can be obtained

$$\boldsymbol{Q} = \boldsymbol{Q}_1\boldsymbol{Q}_2\cdots\boldsymbol{Q}_p, \quad \boldsymbol{R} = \begin{bmatrix}\boldsymbol{R}_{11}, & \boldsymbol{R}_{12} \\ \boldsymbol{R}_{21}, & \boldsymbol{R}_{22}\end{bmatrix}\begin{matrix}\}p \\ \}M-\bar{p}-p\end{matrix} \tag{A14}$$

where [see equations (A15)–(A17), shown at the bottom of the page], the facts that the Household transform matrix $\boldsymbol{Q}_i$ (and $\boldsymbol{H}_i$) is unitary and $\boldsymbol{Q}_i^H = \boldsymbol{Q}_i$ are used implicitly. From (A12), since $g_{ii}^{(i)} \neq 0$ for $i = 1, 2, \cdots, p$ and $c_s^2 \neq 0$ (i.e., $p \geq 1$), obviously we can find that $\dim(\mathcal{R}(\boldsymbol{\Psi})) = \operatorname{rank}(\boldsymbol{R}) = \operatorname{rank}(\boldsymbol{R}_{11}) = p$, i.e., the number of incident signals can be determined from the rank of QR factor $\boldsymbol{R}$ of the matrix $\boldsymbol{\Psi}$. ■

Appendix B
Proof of Theorem 2

As the signal vectors $\tilde{\boldsymbol{y}}_{fl}(n)$ and $\tilde{\boldsymbol{y}}_{bl}(n)$ of the lth "virtual" forward and backward subarrays are given by

$$\begin{aligned}\tilde{\boldsymbol{y}}_{fl}(n) &\triangleq [y_l(n), y_{l+1}(n), \cdots, y_{l+L-2}(n)]^T \\ &= \bar{\boldsymbol{A}}\boldsymbol{D}^{l-1}\boldsymbol{s}(n) + \tilde{\boldsymbol{w}}_{fl}(n)\end{aligned} \tag{B1}$$

$$\begin{aligned}\tilde{\boldsymbol{y}}_{bl}(n) &\triangleq [y_{M-l+1}(n), y_{M-l}(n), \cdots, y_{\bar{p}-l+2}(n)]^H \\ &= \bar{\boldsymbol{A}}\boldsymbol{D}^{-(M-l)}\boldsymbol{s}^*(n) + \tilde{\boldsymbol{w}}_{bl}(n)\end{aligned} \tag{B2}$$

where $\tilde{\boldsymbol{w}}_{fl}(n) \triangleq [w_l(n), w_{l+1}(n), \cdots, w_{l+L-2}(n)]^T$, and $\tilde{\boldsymbol{w}}_{bl}(n) \triangleq [w_{M-l+1}(n), w_{M-l}(n), \cdots, w_{\bar{p}-l+2}(n)]^H$, for $l = 1, 2, \cdots, \bar{p}+1$, we can get the covariance matrices $\boldsymbol{C}_{ll}^f$ and $\boldsymbol{C}_{ll}^b$ of the lth "virtual" forward and backward subarray as

$$\begin{aligned}\boldsymbol{C}_{ll}^f &\triangleq E\left\{\tilde{\boldsymbol{y}}_{fl}(n)\tilde{\boldsymbol{y}}_{fl}^H(n)\right\} \\ &= c_s\bar{\boldsymbol{A}}\boldsymbol{D}^{l-1}\boldsymbol{\beta}\boldsymbol{\beta}^H\boldsymbol{D}^{-(l-1)}\bar{\boldsymbol{A}}^H + \sigma^2\boldsymbol{I}_{M-\bar{p}}\end{aligned} \tag{B3}$$

$$\begin{aligned}\boldsymbol{C}_{ll}^b &\triangleq E\left\{\tilde{\boldsymbol{y}}_{bl}(n)\tilde{\boldsymbol{y}}_{bl}^H(n)\right\} \\ &= c_s\bar{\boldsymbol{A}}\boldsymbol{D}^{-(M-l)}\boldsymbol{\beta}^*\boldsymbol{\beta}^T\boldsymbol{D}^{M-l}\bar{\boldsymbol{A}}^H + \sigma^2\boldsymbol{I}_{M-\bar{p}}.\end{aligned} \tag{B4}$$

Further, we easily obtain

$$\boldsymbol{C}_1^f \triangleq \sum_{l=1}^{\bar{p}}\boldsymbol{C}_{ll}^f = \bar{\boldsymbol{C}}_1^f + \bar{p}\sigma^2\boldsymbol{I}_{M-\bar{p}} \tag{B5}$$

$$\boldsymbol{C}_2^f \triangleq \sum_{l=2}^{\bar{p}+1}\boldsymbol{C}_{ll}^f = \bar{\boldsymbol{C}}_2^f + \bar{p}\sigma^2\boldsymbol{I}_{M-\bar{p}} \tag{B6}$$

$$\boldsymbol{C}_1^b \triangleq \sum_{l=1}^{\bar{p}}\boldsymbol{C}_{ll}^b = \bar{\boldsymbol{C}}_1^b + \bar{p}\sigma^2\boldsymbol{I}_{M-\bar{p}} \tag{B7}$$

$$\boldsymbol{C}_2^b \triangleq \sum_{l=2}^{\bar{p}+1}\boldsymbol{C}_{ll}^b = \bar{\boldsymbol{C}}_2^b + \bar{p}\sigma^2\boldsymbol{I}_{M-\bar{p}} \tag{B8}$$

where the noise-free counterparts are given by $\bar{\boldsymbol{C}}_1^f \triangleq c_s\bar{\boldsymbol{A}}\boldsymbol{B}\bar{\boldsymbol{A}}_1^T\bar{\boldsymbol{A}}_1^*\boldsymbol{B}^*\bar{\boldsymbol{A}}^H$, $\bar{\boldsymbol{C}}_2^f \triangleq c_s\bar{\boldsymbol{A}}\boldsymbol{B}\boldsymbol{D}\bar{\boldsymbol{A}}_1^T\bar{\boldsymbol{A}}_1^*\boldsymbol{D}^{-1}\boldsymbol{B}^*\bar{\boldsymbol{A}}^H$, $\bar{\boldsymbol{C}}_1^b \triangleq c_s\bar{\boldsymbol{A}}\boldsymbol{D}^{-(M-1)}\boldsymbol{B}^*\bar{\boldsymbol{A}}_1^T\bar{\boldsymbol{A}}_1^*\boldsymbol{B}\boldsymbol{D}^{M-1}\bar{\boldsymbol{A}}^H$, and $\bar{\boldsymbol{C}}_2^b \triangleq c_s\bar{\boldsymbol{A}}\boldsymbol{D}^{-(M-2)}\boldsymbol{B}^*\bar{\boldsymbol{A}}_1^T\bar{\boldsymbol{A}}_1^*\boldsymbol{B}\boldsymbol{D}^{M-2}\bar{\boldsymbol{A}}^H$.

From (36)–(38), the estimated Hankel correlation matrices $\hat{\boldsymbol{\Phi}}_f$, $\hat{\bar{\boldsymbol{\Phi}}}_f$, $\hat{\boldsymbol{\Phi}}_b$, and $\hat{\bar{\boldsymbol{\Phi}}}_b$ can be reexpressed by (cf. [44])

$$\hat{\boldsymbol{\Phi}}_f = \frac{1}{N}\sum_{n=1}^{N}\boldsymbol{Y}_f(n)y_M^*(n),\ \hat{\bar{\boldsymbol{\Phi}}}_f = \frac{1}{N}\sum_{n=1}^{N}\bar{\boldsymbol{Y}}_f(n)y_1^*(n) \tag{B9}$$

$$\hat{\boldsymbol{\Phi}}_b = \frac{1}{N}\sum_{n=1}^{N}\boldsymbol{Y}_b(n)y_1(n),\ \hat{\bar{\boldsymbol{\Phi}}}_b = \frac{1}{N}\sum_{n=1}^{N}\bar{\boldsymbol{Y}}_b(n)y_M(n) \tag{B10}$$

$$\boldsymbol{R}_{11} = c_s^2|\bar{\rho}_M|^2\begin{bmatrix} g_{11}^{(1)}, & g_{21}^{(1)}, & g_{31}^{(1)}, & \cdots & g_{p1}^{(1)} \\ 0, & g_{22}^{(2)}, & g_{32}^{(2)}, & \cdots & g_{p2}^{(2)} \\ 0, & 0, & g_{33}^{(3)}, & \cdots & g_{p3}^{(3)} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0, & 0, & 0, & \cdots & g_{pp}^{(p)} \end{bmatrix} \tag{A15}$$

$$\boldsymbol{R}_{12} = c_s^2|\bar{\rho}_M|^2\begin{bmatrix} \gamma_{p+1,1}g_{11}^{(1)} + \gamma_{p+1,2}g_{21}^{(1)} + \cdots + \gamma_{p+1,p}g_{p1}^{(1)}, & \cdots & \gamma_{M-\bar{p},1}g_{11}^{(1)} + \gamma_{M-\bar{p},2}g_{21}^{(1)} + \cdots + \gamma_{M-\bar{p},p}g_{p1}^{(1)} \\ \gamma_{p+1,2}g_{22}^{(2)} + \cdots + \gamma_{p+1,p}g_{p2}^{(2)}, & \cdots & \gamma_{M-\bar{p},2}g_{22}^{(2)} + \cdots + \gamma_{M-\bar{p},p}g_{p2}^{(2)} \\ \gamma_{p+1,3}g_{33}^{(3)} + \cdots + \gamma_{p+1,p}g_{p3}^{(3)}, & \cdots & \gamma_{M-\bar{p},3}g_{33}^{(3)} + \cdots + \gamma_{M-\bar{p},p}g_{p3}^{(3)} \\ \vdots & \ddots & \vdots \\ \gamma_{p+1,p}g_{pp}^{(p)}, & \cdots & \gamma_{M-\bar{p},p}g_{pp}^{(p)} \end{bmatrix} \tag{A16}$$

$$\boldsymbol{R}_{21} = \boldsymbol{O}_{(M-\bar{p}-p)\times p}, \quad \boldsymbol{R}_{22} = \boldsymbol{O}_{(M-\bar{p}-p)\times(M-\bar{p}-p)} \tag{A17}$$

where $\boldsymbol{Y}_f(n) \triangleq [\tilde{\boldsymbol{y}}_{f1}(n), \tilde{\boldsymbol{y}}_{f2}(n), \cdots, \tilde{\boldsymbol{y}}_{f\bar{p}}(n)]$, $\bar{\boldsymbol{Y}}_f(n) \triangleq [\tilde{\boldsymbol{y}}_{f2}(n), \tilde{\boldsymbol{y}}_{f3}(n), \cdots, \tilde{\boldsymbol{y}}_{f\bar{p}+1}(n)]$, $\boldsymbol{Y}_b(n) \triangleq [\tilde{\boldsymbol{y}}_{b1}(n), \tilde{\boldsymbol{y}}_{b2}(n), \cdots, \tilde{\boldsymbol{y}}_{b\bar{p}}(n)]$, and $\bar{\boldsymbol{Y}}_b(n) \triangleq [\tilde{\boldsymbol{y}}_{b2}(n), \tilde{\boldsymbol{y}}_{b3}(n), \cdots, \tilde{\boldsymbol{y}}_{b\bar{p}+1}(n)]$. Under the basic assumptions and by using the formula for the expectation of the product of four complex Gaussian random variables and vectors with zero-mean [49], we can get

$$\begin{aligned}
&E\left\{\hat{\bar{\boldsymbol{\Phi}}}_f \hat{\bar{\boldsymbol{\Phi}}}_f^H\right\} \\
&= \frac{1}{N^2} E\left\{\left(\sum_{n=1}^{N} \boldsymbol{Y}_f(n) y_M^*(n)\right)\left(\sum_{t=1}^{N} \boldsymbol{Y}_f(t) y_M^*(t)\right)^H\right\} \\
&= \frac{1}{N^2} \sum_{l=1}^{\bar{p}} \sum_{n=1}^{N} \sum_{t=1}^{N} E\left\{y_M^*(n) y_M(t) \tilde{\boldsymbol{y}}_{fl}(n) \tilde{\boldsymbol{y}}_{fl}^H(t)\right\} \\
&= \frac{1}{N^2} \sum_{l=1}^{\bar{p}} \sum_{n=1}^{N} \sum_{t=1}^{N} \Big(E\left\{y_M^*(n) y_M(t)\right\} E\left\{\tilde{\boldsymbol{y}}_{fl}(n) \tilde{\boldsymbol{y}}_{fl}^H(t)\right\} \\
&\qquad + E\left\{y_M^*(n) \tilde{\boldsymbol{y}}_{fl}(n)\right\} E\left\{y_M(t) \tilde{\boldsymbol{y}}_{fl}^H(t)\right\} \\
&\qquad + E\left\{y_M(t) \tilde{\boldsymbol{y}}_{fl}(n)\right\} E\left\{y_M^*(n) \tilde{\boldsymbol{y}}_{fl}^H(t)\right\}\Big) \\
&= \frac{1}{N^2} \sum_{l=1}^{\bar{p}} \sum_{n=1}^{N} \sum_{t=1}^{N} \Big(c_{MM} \delta_{n,t} \boldsymbol{C}_{ll}^f + \boldsymbol{\psi}_{fl} \boldsymbol{\psi}_{fl}^H \\
&\qquad + \boldsymbol{O}_{(M-\bar{p})\times(M-\bar{p})}\Big) \\
&= \frac{1}{N} c_{MM} \sum_{l=1}^{\bar{p}} \boldsymbol{C}_{ll}^f + \sum_{l=1}^{\bar{p}} \boldsymbol{\psi}_{fl} \boldsymbol{\psi}_{fl}^H \\
&= \frac{1}{N} c_{MM} \boldsymbol{C}_1^f + \boldsymbol{\Phi}_f \boldsymbol{\Phi}_f^H
\end{aligned} \tag{B11}$$

where $\boldsymbol{\psi}_{fl} \triangleq E\{\tilde{\boldsymbol{y}}_{fl}(n) y_M^*(n)\}$, and $\boldsymbol{\Phi}_f = [\boldsymbol{\psi}_{f1}, \boldsymbol{\psi}_{f2}, \cdots, \boldsymbol{\psi}_{f\bar{p}}]$. Similarly, we obtain

$$\begin{aligned}
&E\left\{\hat{\bar{\boldsymbol{\Phi}}}_f \hat{\bar{\boldsymbol{\Phi}}}_f^H\right\} \\
&= \frac{1}{N^2} \sum_{l=1}^{\bar{p}} \sum_{n=1}^{N} \sum_{t=1}^{N} E\left\{y_1^*(n) y_1(t) \tilde{\boldsymbol{y}}_{fl+1}(n) \tilde{\boldsymbol{y}}_{fl+1}^H(t)\right\} \\
&= \frac{1}{N^2} \sum_{l=1}^{\bar{p}} \sum_{n=1}^{N} \sum_{t=1}^{N} \Big(c_{11} \delta_{n,t} \boldsymbol{C}_{l+1,l+1}^f + \bar{\boldsymbol{\psi}}_{fl+1} \bar{\boldsymbol{\psi}}_{fl+1}^H \\
&\qquad + \boldsymbol{O}_{(M-\bar{p})\times(M-\bar{p})}\Big) \\
&= \frac{1}{N} c_{11} \boldsymbol{C}_2^f + \bar{\boldsymbol{\Phi}}_f \bar{\boldsymbol{\Phi}}_f^H
\end{aligned} \tag{B12}$$

$$\begin{aligned}
&E\left\{\hat{\bar{\boldsymbol{\Phi}}}_b \hat{\bar{\boldsymbol{\Phi}}}_b^H\right\} \\
&= \frac{1}{N^2} \sum_{l=1}^{\bar{p}} \sum_{n=1}^{N} \sum_{t=1}^{N} E\left\{y_1(n) y_1^*(t) \tilde{\boldsymbol{y}}_{bl}(n) \tilde{\boldsymbol{y}}_{bl}^H(t)\right\} \\
&= \frac{1}{N^2} \sum_{l=1}^{\bar{p}} \sum_{n=1}^{N} \sum_{t=1}^{N} \left(c_{11} \delta_{n,t} \boldsymbol{C}_{ll}^b + \boldsymbol{\psi}_{bl} \boldsymbol{\psi}_{bl}^H + \boldsymbol{O}_{(M-\bar{p})\times(M-\bar{p})}\right) \\
&= \frac{1}{N} c_{11} \boldsymbol{C}_1^b + \bar{\boldsymbol{\Phi}}_b \bar{\boldsymbol{\Phi}}_b^H
\end{aligned} \tag{B13}$$

$$\begin{aligned}
&E\left\{\hat{\bar{\boldsymbol{\Phi}}}_b \hat{\bar{\boldsymbol{\Phi}}}_b^H\right\} \\
&= \frac{1}{N^2} \sum_{l=1}^{\bar{p}} \sum_{n=1}^{N} \sum_{t=1}^{N} E\left\{y_M(n) y_M^*(t) \tilde{\boldsymbol{y}}_{bl+1}(n) \tilde{\boldsymbol{y}}_{bl+1}^H(t)\right\} \\
&= \frac{1}{N^2} \sum_{l=1}^{\bar{p}} \sum_{n=1}^{N} \sum_{t=1}^{N} \Big(c_{MM} \delta_{n,t} \boldsymbol{C}_{l+1,l+1}^b + \bar{\boldsymbol{\psi}}_{bl+1} \bar{\boldsymbol{\psi}}_{bl+1}^H \\
&\qquad + \boldsymbol{O}_{(M-\bar{p})\times(M-\bar{p})}\Big) \\
&= \frac{1}{N} c_{MM} \boldsymbol{C}_2^b + \bar{\boldsymbol{\Phi}}_b \bar{\boldsymbol{\Phi}}_b^H
\end{aligned} \tag{B14}$$

where $\bar{\boldsymbol{\psi}}_{fl} \triangleq E\{\tilde{\boldsymbol{y}}_{fl}(n) y_1^*(n)\}$, $\boldsymbol{\psi}_{bl} \triangleq E\{y_1(n) \tilde{\boldsymbol{y}}_{bl}(n)\}$, $\bar{\boldsymbol{\psi}}_{bl} \triangleq E\{y_M(n) \tilde{\boldsymbol{y}}_{bl}(n)\}$, $\bar{\boldsymbol{\Phi}}_f = [\bar{\boldsymbol{\psi}}_{f2}, \bar{\boldsymbol{\psi}}_{f3}, \cdots, \bar{\boldsymbol{\psi}}_{f\bar{p}+1}]$, $\boldsymbol{\Phi}_b = [\boldsymbol{\psi}_{b1}, \boldsymbol{\psi}_{b2}, \cdots, \boldsymbol{\psi}_{b\bar{p}}]$, and $\bar{\boldsymbol{\Phi}}_b = [\bar{\boldsymbol{\psi}}_{b2}, \bar{\boldsymbol{\psi}}_{b3}, \cdots, \bar{\boldsymbol{\psi}}_{b\bar{p}+1}]$. From (18) and (30), by combining (B11)–(B14) together, we straightforwardly get

$$E\{\hat{\boldsymbol{\Psi}}\} = \boldsymbol{\Psi} + \frac{1}{N}\left(c_{MM}\left(\boldsymbol{C}_1^f + \boldsymbol{C}_2^b\right) + c_{11}\left(\boldsymbol{C}_2^f + \boldsymbol{C}_1^b\right)\right). \tag{B15}$$

Therefore, the asymptotical error $E\{\hat{\boldsymbol{\Psi}} - \boldsymbol{\Psi}\}$ in (40) can be obtained immediately from (B15). ■

APPENDIX C
PROOF OF THEOREM 3

To established the consistency of proposed criterion, it suffices to prove that $\xi(i) - \xi(p) < 0$ almost sure (a.s.) with probability one (w.p.1) when $N \to \infty$ for $i \neq p$ (cf. [7] and [11]).

Because the predetermined permutation matrix $\boldsymbol{\Pi}$ does not affect the statistical property of $\hat{\boldsymbol{\Psi}}\boldsymbol{\Pi}$ and the estimate $\hat{\boldsymbol{\Psi}}$ is asymptotically consistent, we easily find that the QR factor $\hat{\boldsymbol{R}}$ in (31) is asymptotically consistent, i.e., $\hat{\boldsymbol{R}} \overset{N\to\infty}{\longrightarrow} \boldsymbol{R}$ a.s. w.p.1. Then, from (20), (31), and (32), we can get $\zeta(i) \overset{N\to\infty}{\longrightarrow} \bar{\zeta}_i + \varepsilon$ for $1 \le i \le p$ while $\zeta(i) \overset{N\to\infty}{\longrightarrow} \varepsilon$ for $p < i \le M - \bar{p}$, where $\bar{\zeta}_i \triangleq \sum_{k=i}^{M-\bar{p}} |r_{ik}|$. Hence, from (33), we have

$$\lim_{N\to\infty} \xi(i) = \begin{cases} \frac{\bar{\zeta}_i + \varepsilon}{\bar{\zeta}_{i+1} + \varepsilon} \approx \frac{\bar{\zeta}_i}{\bar{\zeta}_{i+1}} \triangleq \bar{c}_i, & \text{for } 1 \le i < p \\ \frac{\bar{\zeta}_p}{\varepsilon} + 1 \to \infty, & \text{for } i = p \\ \frac{\varepsilon}{\varepsilon} = 1, & \text{for } p < i \le M - \bar{p} - 1 \end{cases} \tag{C1}$$

where $\bar{c}_i$ is a positive constant. Thus, it follows that $\xi(i) - \xi(p) < 0$ for $i < p$ and $i > p$ as $N \to \infty$; consequently the maximum is achieved at $i = p$. That is the probability of missing $P_m \triangleq \mathrm{Prob}(\hat{p} < p)$ and that of false alarm $P_{\mathrm{fa}} \triangleq \mathrm{Prob}(\hat{p} > p)$ goes to zero asymptotically while the probability of correct detection approaches to one, when $N \to \infty$. Therefore, we can conclude that the estimate $\hat{p}$ in (34) is asymptotically consistent, i.e., $\hat{p} = p$ w.p.1 as $N \to \infty$. ■

ACKNOWLEDGMENT

The authors would like to thank the anonymous reviewers and the Associate Editor Prof. A. Leyman for the careful review, insight comments, and constructive suggestions and especially for the invaluable performance verification with the experimental data.

REFERENCES

[1] H. Krim and M. Viberg, "Two decades of array signal processing research: The parametric approach," *IEEE Signal Process. Mag.*, vol. 13, no. 4, pp. 67–94, 1996.

[2] H. L. Van Trees, *Optimum Array Processing, Part IV of Detection, Estimation, and Modulation Theory*. New York: Wiley, 2002.

[3] J. Xin and A. Sano, "MSE-based regularization approach to direction estimation of coherent narrowband signals using linear prediction," *IEEE Trans. Signal Process.*, vol. 49, no. 11, pp. 2481–2497, Nov. 2001.

[4] H. Akaike, "A new look at the statistical model identification," *IEEE Trans. Autom. Control*, vol. 19, no. 6, pp. 716–723, 1974.

[5] J. Rissanen, "Modeling by shortest data description," *Automatica*, vol. 14, no. 5, pp. 465–471, 1978.

[6] G. Schwarz, "Estimating the dimension of a model," *Ann. Stat.*, vol. 6, no. 2, pp. 461–464, 1978.

[7] M. Wax and T. Kailath, "Detection of signals by information theoretic criteria," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 33, no. 2, pp. 387–392, Apr. 1985.

[8] M. Kaveh, H. Wang, and H. Hung, "On the theoretical performance of a class of estimators of the number of narrowband sources," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 35, no. 9, pp. 1350–1352, 1987.

[9] Q.-T. Zhang, K. M. Wong, P. C. Yip, and J. P. Reilly, "Statistical analysis of the performance of information theoretic criteria in the detection of the number of signals in array processing," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 37, no. 10, pp. 1557–1567, 1989.

[10] K. M. Wong, Q.-T. Zhang, J. P. Reilly, and P. C. Yip, "On information theoretic criteria for determining the number of signals in high resolution array processing," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 38, no. 11, pp. 1959–1971, 1990.

[11] E. Fishler, M. Grosmann, and H. Messer, "Detection of signals by information theoretic criteria: General asymptotic performance analysis," *IEEE Trans. Signal Process.*, vol. 50, no. 5, pp. 1027–1036, May 2002.

[12] G. Xu, R. H. Roy, III, and T. Kailath, "Detection of number of sources via exploitation of centro-symmetry property," *IEEE Trans. Signal Process.*, vol. 42, no. 1, pp. 102–112, Jan. 1994.

[13] Y. Q. Yin and P. R. Krishnaiah, "On some nonparameteric methods for detection of the number of signals," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 35, no. 11, pp. 1533–1538, 1987.

[14] W. Chen, K. M. Wong, and J. P. Reilly, "Detection of the number of signals: a predicted eigen-threshold approach," *IEEE Trans. Signal Process.*, vol. 39, no. 5, pp. 1088–1098, May 1991.

[15] H. Lee and F. Li, "An eigenvector technique for detecting the number of emitters in a cluster," *IEEE Trans. Signal Process.*, vol. 42, no. 9, pp. 2380–2388, Sep. 1994.

[16] W. Xu, J. Pierre, and M. Kaveh, "Practical detection with calibrated arrays," in *Proc. IEEE 6th SP Workshop Statistical Signal Array Processing*, Victoria, Canada, Oct. 1992, pp. 82–85.

[17] M. Nezafat, M. Kaveh, and W. Xu, "Estimation of the number of sources based on the eigenvectors of the covariance matrix," in *Proc. IEEE 3rd Sensor Array Multichannel Signal Process. Workshop*, Barcelona, Spain, Jul. 2004, S10-5.

[18] H.-T. Wu, J.-F. Yang, and F.-K. Chen, "Source number estimators using transformed Gerschgorin radii," *IEEE Trans. Signal Process.*, vol. 43, no. 6, pp. 1325–1333, Jun. 1995.

[19] K. Konstantinides and K. Yao, "Statistical analysis of effective singular values in matrix rank matrix," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 36, no. 5, pp. 757–763, 1988.

[20] M. Wax and I. Ziskind, "Detection of the number of coherent signals by MDL principle," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 37, no. 8, pp. 1190–1196, 1989.

[21] Q. Wu and D. R. Fuhrmann, "A parametric method for determining the number of signals in narrowband direction finding," *IEEE Trans. Signal Process.*, vol. 39, no. 8, pp. 1848–1857, Aug. 1991.

[22] M. Wax, "Detection and localization of multiple sources via the stochastic signals model," *IEEE Trans. Signal Process.*, vol. 39, no. 11, pp. 2450–2456, Nov. 1991.

[23] M. Viberg, B. Ottersten, and T. Kailath, "Detection and estimation in sensor arrays using weighted subspace fitting," *IEEE Trans. Signal Process.*, vol. 39, no. 11, pp. 2436–2449, Nov. 1991.

[24] B. Ottersten, M. Viberg, P. Stoica, and A. Nehorai, "Exact and large sample maximum likelihood techniques for parameter estimation and detection in array processing," in *Radar Array Processing*, S. Haykin, J. Litva, and T. J. Shepherd, Eds. Berlin, Germany: Springer-Verlag, 1993, pp. 99–151.

[25] C.-M. Cho and P. M. Djuric, "Detection and estimation of DOA's of signals via Bayesian predictive densities," *IEEE Trans. Signal Process.*, vol. 42, no. 11, pp. 3051–3060, Nov. 1994.

[26] S. Valaee and P. Kabal, "An information theoretic approach to source enumeration in array signal processing," *IEEE Trans. Signal Process.*, vol. 52, no. 5, pp. 1171–1178, Nov. 2004.

[27] T.-J. Shan, M. Wax, and T. Kailath, "On spatial smoothing of estimation of coherent signals," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 33, no. 4, pp. 806–811, 1985.

[28] T.-J. Shan, A. Paulraj, and T. Kailath, "On smoothed rank profile tests in eigenstructure methods for directions-of-arrival estimation," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 35, no. 10, pp. 1377–1385, 1987.

[29] J. H. Cozzens and M. J. Sousa, "Source enumeration in a correlated signal environment," *IEEE Trans. Signal Process.*, vol. 42, no. 2, pp. 304–317, Feb. 1994.

[30] H. Krim and J. H. Cozzens, "A data-based enumeration technique for fully correlated signals," *IEEE Trans. Signal Process.*, vol. 42, no. 7, pp. 1662–1668, Jul. 1994.

[31] S. U. Pillai and B. H. Kwon, "Forward/backward spatial smoothing techniques for coherent signals identification," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 37, no. 1, pp. 8–15, 1989.

[32] R. F. Brcich, A. M. Zoubir, and P. Pelin, "Detection of sources using bootstrap techniques," *IEEE Trans. Signal Process.*, vol. 50, no. 2, pp. 206–215, Feb. 2002.

[33] G. H. Golub and C. F. Van Loan, *Matrix Computations*, 2nd ed. Baltimore, MD: The John Hopkins Univ. Press, 1989.

[34] P. Comon and G. H. Golub, "Tracking a few extreme singular values and vectors in signal processing," *Proc. IEEE*, vol. 78, no. 8, pp. 1327–1343, 1990.

[35] J. P. Reilly, "A real-time high-resolution technique for angle-of-arrival estimation," *Proc. IEEE*, vol. 75, no. 12, pp. 1692–1694, 1987.

[36] P. Comon, L. Kopp, and J. P. Reilly, "Comments, with reply, on 'A real-time high-resolution technique for angle-of-arrival estimation'," *Proc. IEEE*, vol. 77, no. 3, pp. 492–494, 1989.

[37] J. P. Reilly, W. G. Chen, and K. M. Wong, "A fast QR-based array-processing algorithm," in *Proc. SPIE*, 1988, vol. 975, pp. 36–47.

[38] S. Prasad and B. Chandna, "Direction-of-arrival estimation using rank revealing QR factorization," *IEEE Trans. Signal Process.*, vol. 39, no. 5, pp. 1224–1229, May 1991.

[39] P. Costa, J. Grouffaud, P. Larzabal, and H. Clergeot, "Estimation of the number of signals from features of the covariance matrix: A supervised approach," *IEEE Trans. Signal Process.*, vol. 47, no. 11, pp. 3108–3115, 1999.

[40] A. Swindlehurst, "Alternative algorithm for maximum likelihood DOA estimation and detection," *Proc. Inst. Elect. Eng.—Radar, Sonar Navig.*, vol. 141, no. 6, pp. 293–299, 1994.

[41] S. F. Hsieh, K. J. R. Liu, and K. Yao, "Estimation of multiple sinusoids frequencies using truncated least squares methods," *IEEE Trans. Signal Process.*, vol. 41, no. 2, pp. 990–994, Feb. 1993.

[42] W. M. Gentleman, "Matrix triangularization by systolic arrays," in *Proc. SPIE*, 1981, vol. 298, pp. 19–26.

[43] T. F. Chan, "Rank revealing QR factorization," *Linear Algebra Its Appl.*, vol. 88/89, pp. 67–82, 1987.

[44] J. Xin and A. Sano, "Computationally efficient subspace-based method for direction-of-arrival estimation without eigendecomposition," *IEEE Trans. Signal Process.*, vol. 52, no. 4, pp. 876–893, Apr. 2004.

[45] ——, "Efficient subspace-based algorithm for adaptive bearing estimation and tracking," *IEEE Trans. Signal Process.*, vol. 53, no. 12, pp. 4485–4505, Dec. 2005.

[46] M. Wax and I. Ziskind, "On unique localization of multiple sources by passive sensor arrays," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 38, no. 7, pp. 996–1000, 1989.

[47] J. Xin and A. Sano, "Linear prediction approach to direction estimation of cyclostationary signals in multipath environment," *IEEE Trans. Signal Process.*, vol. 49, no. 4, pp. 710–720, Apr. 2001.

[48] T. K. Moon and W. C. Stirling, *Mathematical Methods and Algorithms for Signal Processing*. Upper Saddle River, NJ: Prentice-Hall, 2000.

[49] P. H. M. Janssen and P. Stoica, "On the expectation of the product of four matrix-valued Gaussian random variables," *IEEE Trans. Autom. Control*, vol. 33, no. 9, pp. 867–870, 1988.

[50] R. E. Bethel and K. L. Bell, "Maximum likelihood approach to joint array detection/estimation," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 40, no. 3, pp. 1060–1072, 2004.

[51] B. Friedlander and A. J. Weiss, "Direction finding in the presence of mutual coupling," *IEEE Trans. Antennas Propag.*, vol. 39, no. 3, pp. 273–284, 1991.

[52] J. Friedmann, E. Fishler, and H. Messer, "General asymptotic analysis of the generalized likelihood ratio test for a Gaussian point source under statistical or spatial mismodeling," *IEEE Trans. Signal Process.*, vol. 50, no. 11, pp. 2617–2631, Nov. 2002.
[53] E. Fishler and H. V. Poor, "Estimation of the number of sources in unbalanced arrays via information theoretic criteria," *IEEE Trans. Signal Process.*, vol. 53, no. 9, pp. 3543–3553, Sep. 2005.
[54] M. Nezafat and M. Kaveh, "Estimation of the number of sources with DOA priors," in *Proc. 13th IEEE Workshop on Statistical Signal Processing*, Bordeaux, France, Jul. 2005, pp. 603–608.
[55] S. Haykin and A. Steinhardt, Eds., *Adaptive Radar Detection and Estimation*. New York: Wiley, 1992.
[56] S. U. Pillai, *Array Signal Processing*. New York: Springer-Verlag, 1989.

Jingmin Xin (S'92–M'96–SM'06) received the B.E. degree in information and control engineering from Xi'an Jiaotong University, Xi'an, China, in 1988 and the M.E. and Ph.D. degrees in electrical engineering from Keio University, Yokohama, Japan, in 1993 and 1996, respectively.

From 1988 to 1990, he was with the Tenth Institute of Ministry of Posts and Telecommunications (MPT) of China, Xi'an. He was with the Communications Research Laboratory (currently the National Institute of Information and Communications Technology), Japan, as an Invited Research Fellow of the Telecommunications Advancement Organization of Japan (TAO) (currently the National Institute of Information and Communications Technology) from 1996 to 1997 and as a Postdoctoral Fellow of the Japan Science and Technology Corporation (JST) (currently the Japan Science and Technology Agency) from 1997 to 1999. He was also a Guest (Senior) Researcher with YRP Mobile Telecommunications Key Technology Research Laboratories Company, Limited, Yokosuka, Japan, from 1999 to 2001. From 2002 to 2007, he was with Fujitsu Laboratories Limited, Yokosuka, Japan. Since 2007, he has been a Professor with the Institute of Artificial Intelligence and Robotics at Xi'an Jiaotong University. His research interests are in the areas of signal processing, including adaptive filtering, statistical signal processing, sensor array processing and spectral estimation, system identification, and their applications to mobile communication systems.

Nanning Zheng (SM'93–F'06) graduated from the Department of Electrical Engineering, Xi'an Jiaotong University, Xi'an, China, in 1975 and received the M.E. degree in information and control engineering from Xi'an Jiaotong University in 1981 and the Ph.D. degree in electrical engineering from Keio University, Yokohama, Japan, in 1985.

In 1975, he joined Xi'an Jiaotong University, where he is currently a Professor and the Director of the Institute of Artificial Intelligence and Robotics. Since August 2003, he has been the President of Xi'an Jiaotong University. His research interests include computer vision, pattern recognition, machine vision and image processing, neural networks, and hardware implementation of intelligent systems.

Dr. Zheng became a member of the Chinese Academy of Engineering in 1999 and has been the Chief Scientist and the Director of the Information Technology Committee of the China National High Technology Research and Development Program since 2001. He was General Chair of the International Symposium on Information Theory and Its Applications and General Co-Chair of the International Symposium on Nonlinear Theory and Its Applications, both in 2002. He is a member of the Board of Governors of the IEEE ITS Society and the Chinese Representative on the Governing Board of the International Association for Pattern Recognition. He also serves as an Executive Deputy Editor of the *Chinese Science Bulletin*.

Akira Sano (M'89) received the B.E., M.E., and Ph.D. degrees in mathematical engineering and information physics from the University of Tokyo, Japan, in 1966, 1968, and 1971, respectively.

In 1971, he joined the Department of Electrical Engineering, Keio University, Yokohama, Japan, where he is currently a Professor with the Department of System Design Engineering. He was a Visiting Research Fellow at the University of Salford, Salford, U.K., from 1977 to 1978. Since 1995, he has been a Visiting Research Fellow with the Communication Research Laboratory (currently the National Institute of Information and Communications Technology), Japan. His current research interests are in adaptive modeling and design theory in control, signal processing and communication, and applications to control of sounds and vibrations, mechanical systems and mobile communication systems. He is a coauthor of the textbook *State Variable Methods in Automatic Control* (New York: Wiley, 1988).

Dr. Sano received the Kelvin Premium from the Institute of Electrical Engineering in 1986. He is a Fellow of the Society of Instrument and Control Engineers and is a Member of the Institute of Electrical Engineering of Japan and the Institute of Electronics, Information and Communications Engineers of Japan. He was General Co-Chair of the 1999 IEEE Conference of Control Applications and an IPC Chair of 2004 IFAC Workshop on Adaptation and Learning in Control and Signal Processing. He served as Chair of IFAC Technical Committee on Modeling and Control of Environmental Systems from 1996 to 2001. He has also been Vice-Chair of IFAC Technical Committee on Adaptive Control and Learning since 1999 and has been Chair of IFAC Technical Committee on Adaptive and Learning Systems since 2002. He is also presently on the Editorial Board of *Signal Processing*.