# Further Results on the Cramér-Rao Bound for Sparse Linear Arrays

Mianzhi Wang, Student Member, IEEE, Zhen Zhang, Student Member, IEEE, and Arye Nehorai, Fellow, IEEE

Abstract-With uniform linear arrays (ULAs), subspace-based direction of arrival (DOA) estimation algorithms, such as MUtiple SIgnal Classification (MUSIC), can only resolve M - 1 sources using M sensors. Sparse linear arrays, such as co-prime and nested arrays, can identify up  $O(M^2)$  uncorrelated sources using only O(M) sensors when such DOA estimation algorithms are applied to their difference coarray model. In our previous work, we derived closed-form asymptotic mean-squared error (MSE) expressions for two coarray based MUSIC algorithms and analyzed the Cramér-Rao bound (CRB) in high signal-tonoise ratio (SNR) regions, under the assumption of uncorrelated sources. In this paper, we provide further analysis of the CRB presented in our previous work, especially for co-prime and nested arrays. We first establish the connection between two CRBs, the CRB derived with the assumption that the sources are uncorrelated, and the classical stochastic CRB derived without this assumption. We show that they are asymptotically equal in high SNR regions for uncorrelated sources. Next, we analyze the behavior of the former CRB for co-prime and nested arrays with a large number of sensors. We show the effect of configuration parameters on this CRB and derive the optimal configuration parameters for co-prime and nested arrays with large number of sensors. We show that this CRB can decrease at a rate of  $O(M^{-5})$  for large values of M, while this rate is only  $O(M^{-3})$  for an *M*-sensor ULA. This finding theoretically demonstrates that co-prime and nested arrays can achieve better asymptotic estimation performance when the number of sensors is a limiting factor. We also show that for a fixed aperture, coprime and nested arrays require more snapshots to achieve the same performance as ULAs, showing the trade-off between the number of spatial samples and the number of temporal samples. Finally, we demonstrate our theoretical results with numerical experiments.

*Index Terms*—Cramér-Rao bound, performance analysis, sparse linear arrays, co-prime arrays, nested arrays

#### I. INTRODUCTION

**D** IRECTION-of-arrival (DOA) estimation is an important topic in array signal processing, finding wide applications in radar and sonar [1]–[3]. Traditionally, a uniform linear array (ULA) is deployed. Using classical subspace-based DOA estimation algorithms, such as MUtiple SIgnal Classification (MUSIC) [4]–[6], we can identify up to M - 1 sources using M sensors. However, if the sources are uncorrelated, sparse linear arrays, such as minimum redundancy arrays

(MRAs) [7]–[9], co-prime arrays [10]–[17], and nested arrays [18]–[25], can identify up to  $O(M^2)$  uncorrelated sources using only M sensors by exploiting their difference coarray structure (e.g., applying MUSIC to the difference coarray model).

Such an attractive property makes it very interesting to analyze the statistical performance of sparse linear arrays that utilize the difference coarray model. In [26]-[28], Stoica and Nehorai conducted a thorough statistical performance analysis of ULAs. The authors derived the closed-form asymptotic mean-squared error (MSE) expression of the MUSIC estimator and analyzed its asymptotic statistical efficiency. The same authors also derived the Cramér-Rao bounds (CRBs) for both the conditional model and the stochastic model, as well as established their connections. In [29], Li et. al analyzed the performance of common subspace-based DOA estimators (e.g., MUSIC, root-MUSIC [5], and ESPRIT [30]) and derived a unified MSE expression. However, these analyses are usually based on the physical array model of ULAs. They cannot be directly extended to sparse linear arrays where the difference coarray model is utilized. In [31], the authors derived the CRB for arbitrary arrays in the one-source case, and numerically analyzed this CRB for various sparse linear arrays. In our previous work [32], we derived closed-form asymptotic MSE expressions of DA-MUSIC [33] and SS-MUSIC [18], two commonly used MUSIC variants that utilize the difference coarray model of sparse linear arrays. We also analyzed the CRB derived with the assumption that the sources are uncorrelated, where we demonstrated its unusual behavior in high signal-to-noise ratio (SNR) regions when the number of sources is greater than the number of sensors. It is worth noting that, in [34] and [35], the authors independently discovered the same phenomenon. Moreover, in [34], the authors showed that this CRB can remain valid even if the number of sources is greater than the number of sensors, which theoretically explains why sparse linear arrays can identify up to  $O(M^2)$ uncorrelated sources using only M sensors.

In this paper, we will take another step and conduct further analysis of the CRB presented in our previous paper [32]. In Section II, we will provide a brief review of sparse linear arrays, the stochastic signal model, and our CRB. In Section III, we will establish the connection between our CRB and the classical stochastic CRB [28], which is derived without the assumption that the sources are uncorrelated. In Section IV, we will analyze the behavior of our CRB for co-prime and nested arrays with large number of sensors. We will analytically show that this CRB can decrease at a rate of  $O(M^{-5})$  when the number of sensors, M, is large and the number of source, K, is

This work was supported by the ONR Grant No. N00014-13-1-0050 and No. N00014-17-1-2371. This work appeared in part in Performance Analysis of Coarray-based MUSIC and the Cramr-Rao Bound in the *Proceedings of the 2017 IEEE International Conference on Acoustics, Speech and Signal Processing.* 

M. Wang, Z. Zhang, and A. Nehorai are with the Preston M. Green Department of Electrical and Systems Engineering, Washington University in St. Louis, St. Louis, MO 63130 USA (e-mail: mianzhi.wang@wustl.edu; zhen.zhang@wustl.edu; nehorai@ese.wustl.edu)

less than M. This result theoretically shows that co-prime and nested arrays can achieve much better asymptotic performance than ULAs with the same number of sensors. Additionally, our analytical results give the optimal configuration parameters for co-prime and nested arrays with large number of sensors. Finally, we will use numerical examples to demonstrate our theoretical results in Section V, and draw concluding remarks in Section VI.

In this paper, we make use of the following notations. Given a matrix A, we use  $A^T$ ,  $A^H$ , and  $A^*$  to denote the transpose, the Hermitian transpose, and the conjugate of A, respectively. We use  $A_{ij}$  to denote the (i, j)-th element of A, and  $a_i$  to denote the *i*-th column of A. Let  $A = [a_1 a_2 \dots a_N] \in \mathbb{C}^{M \times N}$ , and we define the vectorization operation as  $\operatorname{vec}(A) = [a_1^T a_2^T \dots a_N^T]^T$ . We use  $\otimes$ ,  $\odot$ , and  $\circ$  to denote the Kronecker product, the Khatri-Rao product (i.e., the column-wise Kronecker product), and the Hadamard product (i.e., the element-wise product), respectively. We denote by  $\Re(A)$  and  $\Im(A)$  the real and the imaginary parts of A. If A is a square matrix, we denote its trace by tr(A). We use  $\operatorname{diag}(a_1, a_2, \ldots, a_n)$  to denote the diagonal matrix constructed from the diagonal elements  $a_1, \ldots, a_n$ . Given a matrix A, we use diag(A) to denote the column vector constructed from its main diagonal. If A is full column rank, we define its pseudo inverse as  $A^{\dagger} = (A^{H}A)^{-1}A^{H}$ . We also define the projection matrix onto the null space of A as  $\Pi_A^{\perp} = I - A A^{\dagger}$ .

#### II. A REVIEW OF THE COARRAY SIGNAL MODEL

We consider a sparse linear array whose sensors are placed on a grid with grid size  $d_0^{1}$ . We can denote the sensor locations as  $\mathcal{D} = \{\bar{d}_1 d_0, \bar{d}_2 d_0, \ldots, \bar{d}_M d_0\}$ , where  $\bar{d}_i$  are integers and M denotes the number of sensors. Typical sparse linear arrays include minimum redundancy arrays (MRAs) [7], [8], nested arrays [18], co-prime arrays [11], and their extensions [15], [22], [23].

Assume that K far-field narrowband source are impinging on the array from the directions  $\boldsymbol{\theta} = [\theta_1, \theta_2, \dots, \theta_K]^T$ . The N snapshots received by the array can be expressed as

$$\boldsymbol{y}(t) = \boldsymbol{A}(\boldsymbol{\theta})\boldsymbol{x}(t) + \boldsymbol{n}(t), \ t = 1, 2, \dots, N,$$
(1)

where  $A(\theta), x(t)$ , and n denote the steering matrix, the source signals, and the additive noise, respectively. More specifically,  $A(\theta) = [a(\theta_1), a(\theta_2), \dots, a(\theta_K)]$ , where

$$\boldsymbol{a}(\theta_k) = \left[e^{j\frac{2\pi}{\lambda}\bar{d}_1 d_0 \sin \theta_k} \quad \cdots \quad e^{j\frac{2\pi}{\lambda}\bar{d}_M d_0 \sin \theta_k}\right], \quad (2)$$

and  $\lambda$  denotes the carrier wavelength of the impinging signals.

To simplify notations in the following discussion, we define  $\omega_k = (2\pi d_0 \sin \theta_k)/\lambda$  and use  $\boldsymbol{\omega} = [\omega_1, \omega_2, \ldots, \omega_K]^T$  to represent the DOAs. Because there exists a one-to-one mapping between  $\omega_k$  and  $\theta_k$  for every  $\theta_k \in (-\pi/2, \pi/2)$ , there is no loss of information. Typically,  $d_0$  is chosen to be  $\lambda/2$ , and we have  $\omega_k \in (-\pi, \pi)$ .

We adapt the stochastic model [28], where the following assumptions are made:

A1 Both the source signals and the noise are white circularlysymmetric Gaussian.

A2 The source DOAs are distinct (i.e.,  $\omega_k \neq \omega_l \ \forall k \neq l$ ).

A3 The source signals and the noise are uncorrelated.

A4 The snapshots are temporally uncorrelated.

Additionally, we assume that the sources are uncorrelated. Given the above assumptions, we can express the covariance matrix as

$$\boldsymbol{R} = \mathbb{E}[\boldsymbol{y}(t)\boldsymbol{y}^{H}(t)] = \boldsymbol{A}\boldsymbol{P}\boldsymbol{A}^{H} + \sigma\boldsymbol{I}, \qquad (3)$$

where  $\mathbf{P} = \mathbb{E}[\mathbf{x}(t)\mathbf{x}^{H}(t)]$  denotes the source covariance matrix, and  $\sigma$  denotes the variance of the additive noise. Under the assumption that the sources are uncorrelated,  $\mathbf{P}$ reduces to a diagonal matrix, which can be expressed as  $\mathbf{P} = \text{diag}(p_1, p_2, \dots, p_K).$ 

Vectorizing  $\boldsymbol{R}$  leads to

$$\boldsymbol{r} = \operatorname{vec}(\boldsymbol{R}) = \boldsymbol{A}_{\mathrm{d}}\boldsymbol{p} + \sigma\operatorname{vec}(\boldsymbol{I}),$$
 (4)

where  $A_d = A^* \odot A$  and  $p = [p_1, p_2, \dots, p_K]^T$ . Prior work has shown that  $A_d$  can be viewed as an steering matrix of a difference coarray whose sensor locations are given by  $\mathcal{D}_{co} = \{(\bar{d}_m - \bar{d}_n)d_0|m, n = 1, 2, \dots, M\}$  [18]. Therefore rcan be viewed as measurement vector of the difference coarray with a deterministic source signal p plus a deterministic noise term  $\sigma \operatorname{vec}(I)$ , and (4) is usually referred to as the difference coarray model. For carefully designed sparse linear arrays,  $\mathcal{D}_{co}$  contains more unique sensor locations than  $\mathcal{D}$ , and an augmented sample covariance matrix can be constructed from the estimate of r. By applying DOA estimation algorithms, such as MUSIC, to this augmented sample covariance matrix, we are able to resolve more sources than the number of sensors [32].

Because both the difference coarray model and the resulting augmented covariance matrix are constructed from the samples from the original model, we still make use of the statistical properties of the original signal model (1) when conducting performance analysis of such DOA estimation algorithms. Therefore, it is crucial that we thoroughly analyze the CRB based on the signal model (1).

Because P is a diagonal matrix, the number of unknown parameters to be estimated is 2K + 1. Using the property that  $tr(ABCD) = vec(A^T)^T (D^T \otimes B) vec(C)$ , the CRB for the DOAs can then be expressed in the following compact form [32], [34], [35]:

 $\boldsymbol{B}_{(\text{sto-uc})}(\boldsymbol{\omega}) = \frac{1}{N} (\boldsymbol{M}_{\boldsymbol{\omega}}^{H} \boldsymbol{\Pi}_{\boldsymbol{M}_{\boldsymbol{s}}}^{\perp} \boldsymbol{M}_{\boldsymbol{\omega}})^{-1},$ (5)

where

$$\boldsymbol{M}_{\boldsymbol{\omega}} = (\boldsymbol{R}^T \otimes \boldsymbol{R})^{-1/2} \dot{\boldsymbol{A}}_{\mathrm{d}} \boldsymbol{P}, \tag{6}$$

$$\boldsymbol{M_s} = (\boldsymbol{R}^T \otimes \boldsymbol{R})^{-1/2} [\boldsymbol{A}_{\rm d} \ \operatorname{vec}(\boldsymbol{I}_M)], \tag{7}$$

$$\dot{\boldsymbol{A}}_{\mathrm{d}} = \dot{\boldsymbol{A}}^* \odot \boldsymbol{A} + \boldsymbol{A}^* \odot \dot{\boldsymbol{A}}, \qquad (8)$$

$$\boldsymbol{A}_{\mathrm{d}} = \boldsymbol{A}^* \odot \boldsymbol{A},\tag{9}$$

$$\dot{\boldsymbol{A}} = \begin{bmatrix} \frac{\partial \boldsymbol{a}(\omega_1)}{\partial \omega_1} & \frac{\partial \boldsymbol{a}(\omega_1)}{\partial \omega_2} & \cdots & \frac{\partial \boldsymbol{a}(\omega_1)}{\partial \omega_K} \end{bmatrix}.$$
(10)

In our prior work [32], we analyzed the unusual behavior of  $B_{(\text{sto-uc})}$  when  $K \ge M$ . We showed that when  $K \ge M$ ,

<sup>&</sup>lt;sup>1</sup>Alternatively, such a sparse linear array can be viewed as a thinned ULA with  $d_0$  as its inter-element spacing

 $B_{(\text{sto-uc})}$  remains positive definite even if SNR  $\rightarrow \infty$ . In the following sections, we focus on the K < M case. We will first establish the connection between  $B_{(\text{sto-uc})}$  and the classical stochastic CRB derived by Stoica et. al in [28]. Then, we will analyze  $B_{(\text{sto-uc})}$  when M is sufficiently large. These analyses will provide more insights into the performance limits of sparse linear arrays.

## III. CONNECTION TO THE CLASSICAL STOCHASTIC CRB

For the stochastic model [28], we can derive the CRB either without the prior assumption that the sources are uncorrelated, or with it. Hence, we can obtain two different CRBs under the two different assumptions, and both of them are applicable for ULAs and sparse linear arrays when K < M. In the following discussion, we first identify their differences, and then investigate the connection between them.

Without prior knowledge that the sources are uncorrelated, the unknown parameters consist of the DOAs,  $\omega$ , the real and imaginary parts of P, and the noise variance  $\sigma$ . Because Pis Hermitian, there are  $K^2 + K + 1$  unknown parameters. In this case, the CRB of the DOAs is given by [28], [36]:

$$\boldsymbol{B}_{(\text{sto})}(\boldsymbol{\omega}) = \frac{\sigma}{2N} \left\{ \Re[(\dot{\boldsymbol{A}}^{H} \boldsymbol{\Pi}_{\boldsymbol{A}}^{\perp} \dot{\boldsymbol{A}}) \circ (\boldsymbol{P} \boldsymbol{A}^{H} \boldsymbol{R}^{-1} \boldsymbol{A} \boldsymbol{P})^{T}] \right\}^{-1}.$$
(11)

We refer to  $B_{(sto)}$  as the classical stochastic CRB.

With the prior assumption that the sources are uncorrelated, the CRB of the DOAs is given by  $B_{(\text{sto-uc})}$ , as we derived in (5) in the previous section.

Using (5), we observe that the existence of  $B_{(\text{sto-uc})}$  depends on  $A_d$ , the steering matrix of the difference coarray. It has been shown that  $B_{(\text{sto-uc})}$  remains valid for carefully designed sparse linear arrays, even if the number of sources exceeds the number of sensors [34]. On the other hand, according to (11),  $B_{(\text{sto})}$  is valid only when the number of sources is less than the number of sensors. Otherwise  $A^T A$  becomes singular and the projection matrix  $\Pi_A^{\perp}$  is no longer well-defined.

While the compact form (5) of  $B_{(\text{sto-uc})}$  provides great convenience when analyzing the maximum number of resolvable sources [34], it is not well-suited for our asymptotic analysis in the following discussion. Therefore, we provide a brief review of its more traditional form, obtained by block-wise computation of the Fisher information matrix (FIM). Under the assumption that the sources are uncorrelated, the FIM of the stochastic model is given by [2]

$$\boldsymbol{J}_{(\text{sto-uc})} = N \begin{bmatrix} \boldsymbol{J}_{\boldsymbol{\omega}\boldsymbol{\omega}} & \boldsymbol{J}_{\boldsymbol{\omega}\boldsymbol{p}} & \boldsymbol{J}_{\boldsymbol{\omega}\boldsymbol{\sigma}} \\ \boldsymbol{J}_{\boldsymbol{p}\boldsymbol{\omega}} & \boldsymbol{J}_{\boldsymbol{p}\boldsymbol{p}} & \boldsymbol{J}_{\boldsymbol{p}\boldsymbol{\sigma}} \\ \boldsymbol{J}_{\boldsymbol{\sigma}\boldsymbol{\omega}} & \boldsymbol{J}_{\boldsymbol{\sigma}\boldsymbol{p}} & \boldsymbol{J}_{\boldsymbol{\sigma}\boldsymbol{\sigma}} \end{bmatrix},$$
(12)

where

$$\begin{aligned} J_{\omega\omega} =& 2\Re[(\dot{A}^{H}R^{-1}\dot{A})^{*} \circ (PA^{H}R^{-1}AP) \\ &+ (\dot{A}^{H}R^{-1}A)^{*} \circ (PA^{H}R^{-1}\dot{A}P)], \\ J_{pp} =& (A^{H}R^{-1}A)^{*} \circ (A^{H}R^{-1}A), \\ J_{\sigma\sigma} =& \operatorname{tr}(R^{-2}), \\ J_{\omega p} =& 2\Re[(\dot{A}^{H}R^{-1}A)^{*} \circ (PA^{H}R^{-1}A)], \\ J_{\omega\sigma} =& 2\Re[\operatorname{diag}(P\dot{A}^{H}R^{-2}A)], \\ J_{p\sigma} =& \operatorname{diag}(A^{H}R^{-2}A), \end{aligned}$$

and  $J_{p\omega} = J^H_{\omega p}, J_{\sigma\omega} = J^H_{\omega\sigma}, J_{\sigma p} = J^H_{p\sigma}.$ 

By inverting  $J_{(\text{sto-uc})}$ , we obtain the alternative expression of  $B_{(\text{sto-uc})}$ . While this expression seems much more complicated than the one in (5), it can be shown that they are equivalent via Lemma 1 in Appendix A. In the following derivations, we make extensive use of (12) instead of (5).

When P is diagonal, there is a subtle distinction between  $B_{(\text{sto})}$  and  $B_{(\text{sto-uc})}$ .  $B_{(\text{sto})}$  gives the CRB when the sources are uncorrelated and this knowledge is not known a priori.  $B_{(\text{sto-uc})}$  gives the CRB when the sources are uncorrelated and this knowledge is known a priori. This subtle distinction implies that  $B_{(\text{sto})}$  and  $B_{(\text{sto-uc})}$  are not equal. In fact, it is straightforward to show that  $B_{(\text{sto-uc})} \leq B_{(\text{sto})}$ , implying that incorporating the prior knowledge reduces uncertainties in estimating the DOAs. If we compare (11) with (12), we can observe that the term  $PA^HR^{-1}AP$  appears in both expressions, suggesting a potential connection between  $B_{(\text{sto})}$  and  $B_{(\text{sto-uc})}$ . We reveal this connection in Theorem 1.

Theorem 1: Assume that the K sources are uncorrelated and that K < M. If we fix the diagonal matrix  $\mathbf{P} \succ 0$ ,  $\mathbf{B}_{(\text{sto})} \doteq \mathbf{B}_{(\text{sto-uc})}$  as  $\sigma \to 0$ , where  $\doteq$  denotes that the equality is up to the first order with respect to  $\sigma$ .

Proof: See Appendix B.

Theorem 1 shows that when the sources are uncorrelated and the number of sources is less than the number of sensors,  $B_{(sto)}$  and  $B_{(sto-uc)}$  are approximately equal when the SNR is large. This result agrees with our intuition. When the SNR is larger, we can clearly identify the signals, and incorporating the prior knowledge will not give much improvement in estimation performance. When the SNR is low, the signals cannot be clearly distinguished from the noise, and we are more uncertain about whether the sources are correlated. In this case, incorporating the prior knowledge will help improve the estimation performance.

## IV. THE CRAMÉR-RAO BOUND FOR CO-PRIME AND NESTED ARRAYS WITH LARGE NUMBER OF SENSORS

In this section, we analyze the behavior of  $B_{(\mathrm{sto-uc})}$  for ULAs, co-prime arrays, and nested arrays with large numbers of sensors. The expression of  $B_{(sto-uc)}$  is rather complicated and unrevealing. By adopting the assumption that the number of sensors is large, we are able to approximate  $B_{(sto-uc)}$  with a much simpler and more revealing expression, leading to more insights into the statistical performance of co-prime and nested arrays. In [37], our preliminary results showed that  $B_{(sto-uc)}$  for co-prime and nested arrays can decrease at a rate of  $O(M^{-5})$ . In this section, we will provide rigorous proofs and more thorough analysis. Our analysis can also be extended to other sparse linear arrays with closed-form solutions, such as generalized co-prime arrays [15]. While numerical simulations show that MRAs share behaviors similar to co-prime and nested arrays [37], we cannot obtain similar analytical results because MRA configurations do not have closed-form solutions. We will begin with ULAs and then proceed to analyze co-prime and nested arrays. Throughout this section, we will assume that the number of sources is strictly less than the number of sensors.

#### A. Uniform Linear Arrays

We begin by analyzing the behavior of  $B_{(\text{sto-uc})}$  for ULAs with large number of sensors, which will serve as a reference in later discussion. In [26], the authors showed that for an M-sensor ULA, the CRB of the deterministic signal model decreases at a rate of  $O(M^{-3})$  for large M. However, to the authors' best knowledge, it is not shown if  $B_{(\text{sto-uc})}$  shares the same behavior. In the following proposition, we show that  $B_{(\text{sto-uc})}$  indeed shares the same behavior.

Proposition 1: Assume that  $\text{SNR}_i^{-1} = \sigma/p_i \ll M$  for all  $i = 1, 2, \ldots, K$  and that K < M. Then for ULAs, as  $M \to \infty$ ,

$$\boldsymbol{B}_{(\text{sto-uc})}(\boldsymbol{\omega}) \approx \frac{6}{M^3 N} \sigma \boldsymbol{P}^{-1}.$$
 (13)

*Proof:* See Appendix C.

## B. Nested Arrays

Nested arrays are constructed by concatenating two uniform linear arrays with different inter-element spacings. The precise definition of nested arrays is stated as follows:

Definition 1: A nested array generated by the parameter pair  $(N_1, N_2)$  is given by  $\{1, \ldots, N_1\}d_0 \cup \{N_1 + 1, 2(N_1 + 1), \ldots, N_2(N_1 + 1)\}d_0$ .

Unlike ULAs, the physical array geometries of nested arrays can be drastically different, even if they share the same number of sensors<sup>2</sup>. To obtain a more thorough analysis, given the number of sensors, M, we let  $N_1 = \mu M$  and  $N_2 = (1-\mu)M$ . We can vary  $\mu \in (0,1)$  and M to obtain all possible nested array configurations<sup>3</sup>.

Due to the nonuniformity of nested arrays, the behavior of  $B_{(\text{sto-uc})}$  for nested arrays turns out to be more complicated than that of ULAs. We begin with the one-source case.

Theorem 2: Let the rational number  $\mu$  satisfy  $\mu \in (0, 1)$ . Consider a nested array generated by  $(N_1, N_2)$  satisfying  $N_1 + N_2 = M$  and  $N_1 = \mu M$ . Assume that K = 1 and that  $SNR^{-1} = \sigma/p \ll M$ . Then as  $M \to \infty$ ,

$$\boldsymbol{B}_{(\text{sto-uc})}(\omega) \approx \frac{1}{h_{\text{ne}}(\mu)} \frac{1}{N} \frac{1}{M^5} \frac{1}{\text{SNR}}, \quad (14)$$

where

$$h_{\rm ne}(\mu) = \frac{\mu^2 (1-\mu)^3 (1+3\mu)}{6}$$

Proof: See Appendix D-A.

Theorem 2 shows that, in the one-source case,  $B_{(\text{sto-uc})}$ of nested arrays can decrease at a rate of  $O(M^{-5})$  as the number of sensors, M, approaches infinity. This rate is much faster than that of the ULAs, which is  $O(M^{-3})$  as shown in Proposition 1. The coefficient  $h_{\text{ne}}(\mu)$  is determined by  $\mu$ , which encodes the effect of different M-sensor nested array configurations on  $B_{(\text{sto-uc})}$ . As a typical example, when we choose  $N_1 = N_2 = Q$ ,  $B_{(\text{sto-uc})}$  can be simplified to (15) as shown in the following corollary.



Fig. 1:  $(B_{(\text{sto-uc})}(\omega_1) + B_{(\text{sto-uc})}(\omega_2))/2$  computed from different combinations of  $(\omega_1, \omega_2)$  for a uniform linear array with 16 sensors and a nested array generated by the parameter pair (8,8). Both arrays consist of 16 sensors. Locations where  $(N_1 + 1)(\omega_1 - \omega_2) = 2k\pi$  for some non-zero integer k are marked with vertical dashed lines. The horizontal dash line denotes the approximation given by (16) in Theorem 3.

Corollary 1: Under the same assumptions as in Theorem 2, as  $Q \to \infty$ ,  $B_{(\text{sto-uc})}$  for nested arrays generated by the parameter pair (Q, Q) is given by

$$\boldsymbol{B}_{(\text{sto-uc})}(\omega) \approx \frac{12}{5} \frac{1}{Q^5} \frac{1}{N} \frac{1}{\text{SNR}}.$$
 (15)

*Proof:* Immediately obtained by setting M = 2Q and  $\mu = 0.5$  in (14).

We next consider multiple sources. Unlike ULAs, the interelement spacing of the second subarray of a nested array is  $(N_1 + 1)d_0$ , which is greater than  $d_0$ . Consequently, its unambiguous range is  $(-\pi/(N_1 + 1), \pi/(N_1 + 1))$ , which is much smaller than  $(-\pi,\pi)$ . If any two sources,  $\omega_m$  and  $\omega_n$ , satisfy  $(N_1 + 1)(\omega_m - \omega_n) = 2k\pi$  for some non-zero integer k, they cannot be resolved by the second subarray. For instance, when  $N_1 > 1$ , the two DOAs,  $\omega_1 = 0$  and  $\omega_2=2\pi/(N_1+1),$  cannot be resolved by this subarray because  $e^{j\omega_1q(N_1+1)}=e^{j\omega_2q(N_1+1)}$  for  $q=1,2,\ldots,N_2.$ Although such DOA pairs cannot be resolved by the second subarray alone, they can still be resolved by the whole nested array [34]. However, we expect degraded performance when such DOA pairs exist. To illustrate this behavior, we plot in Fig. 1 the values of  $m{B}_{( ext{sto-uc})}$  of two sources, denoted by  $\omega_1$  and  $\omega_2$ , against  $\omega_1 - \omega_2$  for a 16-sensor ULA and a nested array generated by the parameter pair (8,8). We observe that  $B_{(sto-uc)}$  of the nested array is much smaller than that of the ULA. However,  $\boldsymbol{B}_{(\text{sto-uc})}$  of the nested array is not as flat as that of the ULA, and shows large peaks when  $(N_1+1)(\omega_1-\omega_2)=2k\pi$  for some non-zero integer k. To further analyze such a behavior, we introduce the concepts of fully degenerate source placements and  $\delta$ -level non-degenerate source placements in Definition 2 and 3.

Definition 2: Let  $\omega_1, \omega_2, \ldots, \omega_K$  be K distinct DOAs within the range  $(-\pi, \pi)$ . These K DOAs are said to be fully degenerate with respect to a positive integer L if  $(\omega_m - \omega_n)L = 2k\pi$  for some non-zero  $k \in \mathbb{Z}$  whenever  $m \neq n$ ,  $m, n = 1, 2, \ldots, K$ .

<sup>&</sup>lt;sup>2</sup>For example, the nested arrays generated by (8, 2) and (5, 5) both have 10 sensors. However, the latter can achieve 30 degrees of freedom, while the former can achieve only 18 degrees of freedom.

<sup>&</sup>lt;sup>3</sup>Once M is given, the number of choices of  $\mu$  is finite because we need to ensure that both  $N_1$  and  $N_2$  are nonnegative integers.



Fig. 2: An illustration of  $\Omega_3^{0.4}$  within  $(-2\pi, 2\pi)$ . The intervals in  $\Omega_3^{0.4}$  are highlighted by bold segments along the *x*-axis.

Definition 3: Let  $\omega_1, \omega_2, \ldots, \omega_K$  be K distinct DOAs within the range  $(-\pi, \pi)$ . These K DOAs are said to be  $\delta$ -level nondegenerate with respect to a positive integer L if  $\omega_m - \omega_n \in \Omega_L^{\delta}$  whenever  $m \neq n, m, n = 1, 2, \ldots, K$ , where

 $\Omega_L^{\delta} = \{ \omega | \omega L/2 \in [k\pi + \arcsin \delta, (k+1)\pi - \arcsin \delta], k \in \mathbb{Z} \},\$ and  $0 < \delta < 1.$ 

According to Definition 3, if the K DOAs are  $\delta$ -level nondegenerate with respect to L, then for any DOA pair  $\omega_m$ and  $\omega_n$ ,  $|\sin((\omega_m - \omega_n)L/2)| \ge \delta$ . To better understand this statement, we illustrate the set  $\Omega_3^{0.4}$  in Fig. 2. We can observe that, as long as  $\omega \in \Omega_3^{0.4}$ ,  $|\sin(3\omega/2)| \ge 0.4$ . Specifically, if the K DOAs are  $\delta$ -level non-degenerate with respect to 1, then the wraparound distance between any two DOAs within  $(-\pi, \pi)$  is lower bounded by  $2 \arcsin \delta$ , which ensures that any two DOAs are not too close to each other. When the number of sensors is sufficiently large, we are able to conduct various approximations and greatly simplify  $B_{(\text{sto-uc})}$  for both the  $\delta$ -level non-degenerate and the fully degenerate source placements. The results are summarized in Theorem 3.

Theorem 3: Let the rational number  $\mu$  satisfy  $\mu \in (0, 1)$ . Consider a nested array generated by  $(N_1, N_2)$  satisfying  $N_1 + N_2 = M$  and  $N_1 = \mu M$ . Assume that K < M and that  $\text{SNR}_i^{-1} = \sigma/p_i \ll M$  for  $i = 1, 2, \ldots, K$ .

1) If the K DOAs are  $\delta$ -level non-degenerate with respect to 1 and  $N_1+1$  for some  $0 < \delta < 1$  and M is sufficiently large,

$$\boldsymbol{B}_{(\text{sto-uc})}(\boldsymbol{\omega}) \approx \frac{1}{h_{\text{ne}}(\mu)} \frac{1}{N} \frac{1}{M^5} \sigma \boldsymbol{P}^{-1}, \qquad (16)$$

where  $h_{\rm ne}(\mu)$  follows the same definition as in Theorem 2.

2) If the K DOAs are fully degenerate with respect to  $N_1$  + 1, then when M is sufficiently large,

$$\boldsymbol{B}_{(\text{sto-uc})}(\boldsymbol{\omega}) \approx \frac{1}{h_{\text{ne-d}}(\mu)} \frac{1}{N} \frac{1}{M^5} \sigma \boldsymbol{P}^{-1}, \qquad (17)$$

where

$$h_{\text{ne-d}}(\mu) = \frac{\mu^2 (1-\mu)^3}{6} \frac{4\mu + (1-\mu)K}{\mu + (1-\mu)K}.$$
 (18)

Proof: See Appendix E.

*Remark 1:* Theorem 3.1 gives the best-case approximation of  $B_{(\text{sto-uc})}$  for nested arrays with large number of sensors, while Theorem 3.2 provides the approximation of  $B_{(\text{sto-uc})}$ for fully degenerate source placements. When K = 1, (17) reduces to (16). Additionally,  $h_{\text{ne-d}}(\mu)$  decreases as the number of sources, K, increases. Hence  $B_{(\text{sto-uc})}$  increases as the number of fully degenerate sources increases. Recall that in Fig. 1,  $B_{(\text{sto-uc})}$  fluctuates for different combinations of DOAs. We are now able to approximate the range of such fluctuations via (16) and (17).

Remark 2: According to [18], given a fixed number of sensors, M, the maximum degrees of freedom is achieved when  $\mu \approx 0.5$  (i.e.,  $N_1 = N_2$  when M is even and  $N_1 + 1 = N_2$  when M is odd). On the other hand, according to Theorem 2 and Theorem 3.1, for a fixed M, different nested array configurations only affect the coefficient  $h_{\rm ne}(\mu)$ . Therefore, the best-case  $B_{\rm (sto-uc)}$  is minimized when  $h_{\rm ne}(\mu)$  is maximized. Interesting,  $h_{\rm ne}(\mu)$  is maximized at  $\mu^* \approx 0.4625$ , which is slightly different from 0.5. This discrepancy implies that when M is large, a nested array configuration cannot achieve the maximum degrees of freedom and the optimal estimation performance at the same time. We will further illustrate this interesting finding in Section V with numerical experiments.

#### C. Co-prime Arrays

Next, we consider co-prime arrays, whose the definition is given as follows:

Definition 4: A co-prime array generated by the co-prime pair  $(N_1, N_2)$  is given by  $\{0, N_1, \ldots, (N_2 - 1)N_1\}d_0 \cup \{N_2, 2N_2, \ldots, (2N_1 - 1)N_2\}d_0$ .

Similar to nested arrays, given a fixed number of sensors, M, there exist multiple co-prime arrays configurations. To obtain a more thorough analysis, we let  $N_1 = \mu(M+1)$  and  $N_2 = (1-2\mu)(M+1)$ . It can be verified that  $2N_1 + N_2 - 1 = M$  is satisfied. By varying both  $\mu$  and M, we can obtain all possible co-prime array configurations. Note that once M is fixed, the number of choices of  $\mu$  is finite because the following conditions must be satisfied:

- Both μ(M+1) and (1-2μ)(M+1) are positive integers;
   μ(M+1) and (1-2μ)(M+1) are co-prime;
- 3)  $\mu(M+1) < (1-2\mu)(M+1).$

Therefore a valid choice of  $\mu$  must be within the interval (0, 1/3).

We now proceed to consider the one-source case for coprime arrays. The results are summarized in Theorem 4.

Theorem 4: Let the rational number  $\mu$  satisfy  $\mu \in (0, 1/3)$ . Consider a co-prime array generated by the co-prime pair  $(N_1, N_2)$  satisfying  $N_1 = \mu(M+1)$  and  $N_2 = (1-2\mu)(M+1)$ . Assume that K = 1 and that  $\text{SNR}^{-1} = \sigma/p \ll M$ . Then as  $M \to \infty$ ,

$$\boldsymbol{B}_{(\text{sto-uc})}(\omega) \approx \frac{1}{h_{\text{cp}}(\mu)} \frac{1}{N} \frac{1}{M^5} \frac{1}{\text{SNR}},$$
(19)

where

$$h_{\rm cp}(\mu) = \frac{\mu^2 (1 - 2\mu)^2 (1 + 12\mu - 12\mu^2)}{6}.$$

*Proof:* See Appendix D-B.

Theorem 4 shows that, in the one-source case,  $B_{(\text{sto-uc})}$  of co-prime arrays is similar to that of nested arrays. The only difference lies in the coefficient,  $h_{cp}(\mu)$ , which is determined

by the co-prime parameters. As a typical example, when we choose  $N_1 = Q$  and  $N_2 = Q + 1$ ,  $B_{(\text{sto-uc})}$  can be simplified to (20) as shown in the following corollary.

Corollary 2: Under the same assumptions as in Theorem 4, as  $Q \to \infty$ ,  $B_{(\text{sto-uc})}$  for co-prime arrays generated by the co-prime pair (Q, Q + 1) is given by

$$\boldsymbol{B}_{(\text{sto-uc})}(\omega) \approx \frac{6}{11} \frac{1}{Q^5} \frac{1}{N} \frac{1}{\text{SNR}}.$$
 (20)

*Proof:* Immediately obtained by setting M = 3Q and  $\mu \rightarrow 1/3$  in (19).

We next consider multiple sources. Unlike nested arrays, co-prime arrays consist of two subarrays whose inter-element spacings are greater than  $d_0$ . Given a co-prime pair,  $(N_1, N_2)$ , the unambiguous range of the two subarrays are given by  $(-\pi/N_1,\pi/N_1)$  and  $(-\pi/N_2,\pi/N_2)$ , both of which do not cover the full range  $(-\pi, \pi)$ . If two DOAs,  $\omega_m$  and  $\omega_n$ , satisfy  $N_1(\omega_m - \omega_n) = 2k\pi$  for some non-zero integer k, they cannot be resolved by the first subarray. Similarly, if they satisfy  $N_2(\omega_m - \omega_n) = 2k\pi$  for some non-zero integer k, they cannot be resolved by the second subarray. Nevertheless, the full co-prime array can still resolve the DOAs within the full range  $(-\pi, \pi)$  without ambiguity [34], [38]. However, we expect degraded estimation performance when a DOA pair is ambiguous to either of the subarrays. To demonstrate such performance degradation, we plot in Fig. 3 the values of  $B_{(\text{sto-uc})}$  of two sources, denoted by  $\omega_1$  and  $\omega_2$ , against  $\omega_1 - \omega_2$ for a 16-sensor ULA and a co-prime array generated by the parameter pair (5,7). Although  $B_{(sto-uc)}$  of the co-prime array is much smaller than that of the ULA, it is not as flat as that of the ULA. There exist small peaks near the locations given by  $N_1(\omega_1 - \omega) = 2k\pi$  and  $N_1(\omega_1 - \omega) = 2k\pi$ , where k is a non-zero integer. Unlike the results in Fig. 1,  $B_{(sto-uc)}$  of the co-prime array exhibits more peaks clustered around the dashed lines, and the peaks locations are not aligned with the dashed lines. Consequently, we are unable to derive a simple and DOA-independent approximation of  $m{B}_{(\mathrm{sto-uc})}$  similar to (17) under the fully degenerate case. Nevertheless, we can still obtain similar results for the non-degenerate case, which are summarized in Theorem 5.

Theorem 5: Let the rational number  $\mu$  satisfy  $\mu \in (0, 1/3)$ . Consider a co-prime array generated by the co-prime pair  $(N_1, N_2)$  satisfying  $N_1 = \mu(M+1)$  and  $N_2 = (1-2\mu)(M+1)$ . Assume that K < M and that  $\text{SNR}_i^{-1} = \sigma/p_i \ll M$  for  $i = 1, 2, \ldots, K$ . If the K DOAs are  $\delta$ -level non-degenerate with respect to both  $N_1$  and  $N_2$  for some  $0 < \delta < 1$  and M is sufficiently large,

$$\boldsymbol{B}_{(\text{sto-uc})}(\boldsymbol{\omega}) \approx \frac{1}{h_{\text{cp}}(\mu)} \frac{1}{N} \frac{1}{M^5} \sigma \boldsymbol{P}^{-1}, \qquad (21)$$

where  $h_{\rm ne}(\mu)$  follows the same definition as in Theorem 4.

*Proof:* The proof follows exactly the same route as in Appendix E, except that the steering vectors are replaced with those of the co-prime arrays'. The details are omitted due to page limitations.

*Remark 3:* According to [11], a co-prime array generated by the co-prime pair  $(N_1, N_2)$  can achieve  $O(N_1N_2)$  degrees of freedom. Therefore, under the constraint that  $2N_1 + N_2 -$ 



Fig. 3:  $(\boldsymbol{B}_{(\text{sto-uc})}(\omega_1) + \boldsymbol{B}_{(\text{sto-uc})}(\omega_2))/2$  computed from different combinations of  $(\omega_1, \omega_2)$  for a uniform linear array with 16 sensors and a co-prime array generated by the coprime pair (5,7). Both arrays consist of 16 sensors. Locations where  $N_1(\omega_1 - \omega_2) = 2k\pi$  or  $N_2(\omega_1 - \omega_2) = 2k\pi$  for some non-zero integer k are marked with vertical dashed lines. The horizontal dash line denotes the approximation given by (21) in Theorem 5.

1 = M, the maximum degrees of freedom is achieved when  $2N_1 = N_2$ , or  $\mu \approx 0.25$ . Interestingly,  $h_{\rm cp}(\mu)$  is maximized at  $\mu^* \approx 0.2747$ , which is slightly different from 0.25. Therefore, similar to the nested array case, when M is large, a co-prime array cannot achieve the maximum degrees of freedom and the optimal estimation performance at the same time. We will further illustrate this interesting finding in Section V with numerical experiments.

## D. Discussion

Theorem 2–5 lead to the following three important implications for co-prime and nested arrays:

- 1) Given the same number of sensors, co-prime and nested arrays can achieve a much better estimation performance than ULAs.
- Given the same aperture, co-prime and nested arrays need many more snapshots to achieve the same estimation performance of ULAs.
- Co-prime and nested arrays with large number of sensors cannot attain the maximum degrees of freedom and the minimal CRB at the same time.

The first implication comes directly from Theorem 2–5. Given the same number of sensors, M,  $B_{(\text{sto-uc})}$  of co-prime and nested arrays can decrease at a rate of  $O(M^{-5})$ , which is much faster than  $O(M^{-3})$ . In addition to their attractive ability to identify  $K \ge M$  uncorrelated sources, co-prime and nested arrays can also achieve much better estimation performance than ULAs with the same number of sensors when K < M.

To understand the second implication, we consider a ULA with  $M^2$  sensors. From Proposition 1, we know that  $B_{(\text{sto-uc})}$  of this ULA is  $O(M^{-6})$ . To achieve the same aperture, we need a co-prime (or nested) array with only O(M) sensors. However, according Theorem 2–5, the resulting  $B_{(\text{sto-uc})}$  of this co-prime (or nested) array will be only  $O(M^{-5})$ . Therefore, we need O(M) times more snapshots to achieve the

same estimation performance as the ULA. By thinning a ULA into a co-prime (or nested) array, we can reduce the number of sensors from  $O(M^2)$  to O(M), while keeping the array's ability to identify up to  $O(M^2)$  uncorrelated sources. However, this thinning operation indeed comes with a cost: the variance of the estimated DOAs can be M times larger. The second implication shows the trade-off between the number of spatial samples and the number of temporal samples.

The third implication results from Remark 1 and 3. For M-sensor nested arrays, the optimal ratio between  $N_1$  and  $N_2$  to minimize  $\boldsymbol{B}_{(\text{sto-uc})}$  of non-degenerate source placements is approximately 0.8605, which is slightly smaller than 1, the optimal ratio to maximize degrees of freedom. For M-sensor co-prime arrays, the optimal ratio between  $N_1$  and  $N_2$  to minimize  $\boldsymbol{B}_{(\text{sto-uc})}$  of non-degenerate source placements is approximately 0.6096, which is slightly larger than 0.5, the optimal ratio to maximize degrees of freedom.

Remark 4: In the above analysis, the number of sources, K, is assumed to be smaller than the number of sensors, M. Because co-prime and nested arrays can identify more sources than the number of sensors, it would be interesting to conduct a similar analysis for the  $K \ge M$  case. However, when M is very large and  $K \ge M$  holds, the sources become densely located within  $(-\pi/2, \pi/2)$ . In this case,  $\omega_i - \omega_j$  is close to zero for any two different sources i and j, rendering the approximations in Appendix E invalid. Therefore, the results in Theorem 3 and 5 cannot be directly extended to the cases when  $K \ge M$ .

# V. NUMERICAL ANALYSIS

In this section, we demonstrate our results in Theorem 1–5 using numerical experiments. In all the following experiments, we normalize the number of snapshots to 1 and define the SNR as

$$\mathrm{SNR} = 10 \log_{10} \frac{\min_{k=1,2,\dots,K} p_k}{\sigma}.$$

When there are K > 1 sources, we use the mean values,  $\frac{1}{K} \sum_{k=1}^{K} \boldsymbol{B}_{(\text{sto-uc})}(\omega_k)$ , instead of the individual values,  $\boldsymbol{B}_{(\text{sto-uc})}(\omega_i)$ , when making comparisons.

We start this section by demonstrate Theorem 1. We consider the following four different sparse linear arrays:

- Co-prime (3,5):  $[0,3,5,6,9,10,12,15,20,25]d_0$ ;
- MRA 10 [8]: [0, 1, 4, 10, 16, 22, 28, 30, 33, 35]d<sub>0</sub>;
- Nested (4,6):  $[0, 1, 2, 3, 4, 9, 14, 19, 24, 29]d_0;$
- Nested (5,5):  $[0, 1, 2, 3, 4, 5, 11, 17, 23, 29]d_0$ .

We consider six sources with equal power, whose the DOAs,  $\theta_k$ , are given by  $\theta_k = -\pi/3 + 2(k-1)\pi/15$ , k = 1, 2, ..., 6. We vary the SNR from -20 dB to 20 dB and plot the relative difference between  $B_{(sto)}$  and  $B_{(sto-uc)}$  in Fig. 4. It can be observed that when the SNR is above 0 dB, the relative difference between  $B_{(sto)}$  and  $B_{(sto-uc)}$  for all four sparse linear arrays drastically decreases to zero as SNR increases. When the SNR is below 0 dB,  $B_{(sto-uc)}$  becomes more optimistic and deviates from  $B_{(sto)}$ . These observations agree with our theoretical results in Theorem 1.

We next demonstrate Theorem 2 and Theorem 4 via numerical experiments. We consider co-prime arrays generated



Fig. 4:  $|\operatorname{tr}(\boldsymbol{B}_{(\operatorname{sto})} - \boldsymbol{B}_{(\operatorname{sto-uc})})| / \operatorname{tr}(\boldsymbol{B}_{(\operatorname{sto-uc})})$  for the four arrays under different SNRs.



Fig. 5:  $B_{(\text{sto-uc})}$  vs. Q for (a) co-prime arrays; (b) nested arrays. One-source case. The solid lines represent accurate values computed using (5), while the dashed lines represent approximations given by Corollary 1 and 2.

by the co-prime pair (Q, Q + 1), and nested arrays generated by the parameter pair (Q, Q), where we vary Q between 3 and 20. With such configurations,  $B_{(sto-uc)}$  of co-prime and nested arrays can be approximated with even simpler expressions as shown in Corollary 1 and 2. We consider four different SNR settings: -20 dB, -10 dB, 0 dB, and 10 dB, and consider



Fig. 6:  $B_{(\text{sto-uc})}$  vs. M for M-sensor ULAs, nested arrays coprime arrays. K = 5. The markers represent the actual values of  $B_{(\text{sto-uc})}$  obtained from 1000 randomly generated source placements. The dashed lines represent approximations given by Proposition 1, Theorem 3 and Theorem 5.

one source placed at the the origin. The results are plotted in Fig. 5. Give large enough Q values and sufficient SNR, our approximation is very close to the accurate value of  $B_{(sto-uc)}$  for both co-prime and nested arrays. When the SNR is slow, the noise variance term can no longer be neglected and our approximation deviates from the true values. When the value of Q is small, the contribution of the terms with lower degrees with respect to Q is no longer negligible, and our approximation is no longer accurate.

Next, we consider the multiple-source case and demonstrate that  $B_{(sto-uc)}$  for co-prime and nested arrays can indeed decrease at a rate of  $O(M^{-5})$ . We consider three groups of arrays: (i) M-sensor ULAs with  $M = 10, 11, \dots, 100;$ (ii) nested arrays generated by the parameter pairs (Q, Q),  $Q = 5, 6, \dots, 50$ ; (iii) co-prime arrays generated by the coprime pairs  $(Q, Q+1), Q = 5, 6, \dots, 33$ . We consider 5 sources with equal power and set SNR = 0 dB. For each array configuration with M sensors, we randomly generate 1000 source placements within  $(-4\pi/5, 4\pi/5)$  and ensure that the minimal source separation is no less than  $2\pi/M$ . We compute the actual values of  $B_{(sto-uc)}$  for these source placements and compare them with the approximations given in Proposition 1, Theorem 3 and Theorem 5. The results are plotted in Fig. 6. The actual CRB values computed from random source placement, while not fall exactly on the dash lines, cluster closely to the dashed lines as the number of sensors M grows. The observation confirms our approximation of  $B_{(sto-uc)}$  for a sufficient large M. In addition, the CRB values of co-prime and nested arrays decrease much faster than those of ULAs, because they decrease at a rate of  $O(M^{-5})$ instead of  $O(M^{-3})$ .

In the previous experiments, we assume that the sources have equal power. However, Theorem 3 and 5 does not require all sources share the same power. Therefore, we conduct addition experiments for the multiple-source case when the source powers are not equal. We consider four sources with  $p = [2, 10, 30, 50]\sigma$ . The results are plotted in Fig. 7. We ob-



Fig. 7:  $B_{(\text{sto-uc})}$  of individual sources vs. Q for (a) co-prime arrays and (b) nested arrays. Four sources with different powers are considered. The solid lines represent accurate values computed using (5), while the dashed lines represent approximations given by (16) and (21).

serve that the actual CRBs closely follow the approximations given by (16) and (21) for all four sources.

Next, we analyze how different nested and co-prime array configurations affect  $B_{(sto-uc)}$  when the number of sensors M is fixed. We consider six sources with equal power and set SNR = 0 dB. For each array configuration, we randomly generate 5000 source placements within the range  $(-0.9\pi, 0.9\pi)$ while ensuring the minimal source separation is no less than  $\pi/M$ . We compute the values of  $B_{(\text{sto-uc})}$  for these source placements and compare them with the approximations given by Theorem 3 and 5. We first consider nested arrays with 50 sensors and vary  $N_1$  from 10 to 40. Correspondingly, the values of  $\mu$  vary from 0.2 to 0.8. The results are plotted in Fig. 8. It can be observed that most of the CRB values cluster around the approximation given by (16), demonstrating that our result is applicable to a wide range of nested-array configurations. It can also be observed that most of the CRB values fall under the dashed line, showing that the approximation of the fully degenerate case given by (17) provides a reasonable estimate of the degenerated performance. Additionally, we observe that



Fig. 8:  $B_{(\text{sto-uc})}$  vs.  $\mu$  for nested arrays with 50 sensors. K = 6. The solid line represents the approximation obtained from (16), and the dashed line represents the approximation obtained from (17). The markers represent the distribution of actual values of  $B_{(\text{sto-uc})}$  obtained from randomly generated source placements. The marker size is proportional to the number of  $B_{(\text{sto-uc})}$  values concentrating around the marker center.



Fig. 9:  $B_{(\text{sto-uc})}$  vs.  $\mu$  for co-prime arrays with 81 sensors. K = 6. The solid line represents the approximation obtained from (21). The markers represent the distribution of actual values of  $B_{(\text{sto-uc})}$  obtained from randomly generated source placements. The marker size is proportional to the number of  $B_{(\text{sto-uc})}$  values concentrating around the marker center.

the lowest CRB values are found around  $\mu \approx 0.46$  instead of  $\mu = 0.5$ , which agrees with our theoretical prediction in Remark 1. We next consider co-prime arrays with 81 sensors and vary  $N_1$  from 3 to 27. Correspondingly, the values of  $\mu$ vary from 0.037 to 0.329. The maximum degrees of freedom is achieved when  $N_1 = 21, N_2 = 40$  with  $\mu \approx 0.2561$ . The results are plotted in Fig. 9. It can be observed that (21) provides a reasonable estimate of the CRB values for a wide range of co-prime array configurations. Similar to the nested array case, the lowest CRB values are found around  $\mu \approx 0.28$ instead of 0.2561, which agrees with our theoretical prediction in Remark 3.

We close this section by addressing the comments in Remark 4 using numerical experiments. We consider a co-prime array generated by the co-prime pair (20, 21). The resulting co-prime array has M = 60 sensors. We uniformly place the DOAs,  $\omega_k$ , at  $\omega_k = -\pi/3 + 2(k-1)\pi/(3K-3), k =$  $1, 2, \ldots, K$ . We vary the number of sources, K, from 2 to 61. We plot the actual CRB,  $B_{(sto-uc)}$ , together with the



Fig. 10:  $B_{(\text{sto-uc})}$  and the approximation given by (16) versus the number of sources. The co-prime array has 60 sensors. The solid lines represent accurate values computed using (5), while the dashed lines represent approximations given by (21).

approximation given by (21) in Fig. 10. The real CRB values are denoted by solid lines, and the approximations given by (21) are denoted by dashed lines. We can observe that, when the number of sources is small, the actual CRB values are very close to our approximations, despite some fluctuations. However, as the number of sources increases, the actual CRB values begin to deviate from our approximations. In such cases, these sources become very close to each other. To satisfy the assumption that the DOAs are  $\delta$ -level non-degenerate with respect to 20 and 21,  $\delta$  must be chosen to be very small. Hence,  $\delta^{-1}$  will be large enough such that  $M = 60 \gg \delta^{-1}$ no longer holds. Consequently, our approximation (21) is no longer accurate and the actual CRBs start to deviate from our approximation.

#### VI. CONCLUSIONS AND FUTURE WORK

In this paper, we conducted further analysis of the CRB presented in our previous work [32], denoted by  $B_{(sto-uc)}$ . We first showed that, when the SNR is high,  $B_{(\text{sto-uc})}$  coincides with the classical stochastic CRB,  $B_{(sto)}$ . We next analyzed the behavior of  $B_{(sto-uc)}$  for co-prime and nested arrays with a large number of sensors. We showed how different configuration parameters affect  $B_{(sto-uc)}$  and derived the optimal configuration parameters for co-prime and nested arrays with large number of sensors. We showed that given a fixed number of sensors, co-prime and nested arrays significantly outperform ULAs. This finding theoretically confirmed the advantage of using co-prime and nested arrays when the number of sensors is a limiting factor. We also showed that when the aperture is fixed, co-prime and nested arrays need many more snapshots to achieve the same performance as ULAs, demonstrating the trade-off between the number of spatial samples and the number of temporal samples. These results show both the pros and cons of sparse linear arrays and will aid in choosing between sparse linear arrays and ULAs in practical problems. Potential future work involves: (i) investigating the behavior of the CRB in cases of degenerate source placements for coprime arrays; (ii) analyzing the CRB for co-prime and nested arrays with large number of sensors when there are more sources than the number of sensors.

## APPENDIX A USEFUL LEMMAS

Lemma 1: Let A, B, C, D, E, and F be compatible matrices. Then

$$(\boldsymbol{A} \odot \boldsymbol{B})^{H} (\boldsymbol{C} \otimes \boldsymbol{D}) (\boldsymbol{E} \odot \boldsymbol{F}) = (\boldsymbol{A}^{H} \boldsymbol{C} \boldsymbol{E}) \circ (\boldsymbol{B}^{H} \boldsymbol{D} \boldsymbol{F}).$$
 (22)

Proof: The left-hand side of (22) can be expanded as

$$\begin{bmatrix} \boldsymbol{a}_{1}^{H} \otimes \boldsymbol{b}_{1}^{H} \\ \vdots \\ \boldsymbol{a}_{M}^{H} \otimes \boldsymbol{b}_{M}^{H} \end{bmatrix} (\boldsymbol{C} \otimes \boldsymbol{D}) \begin{bmatrix} \boldsymbol{e}_{1} \otimes \boldsymbol{f}_{1} & \cdots & \boldsymbol{e}_{N} \otimes \boldsymbol{f}_{N} \end{bmatrix}, \quad (23)$$

whose (i, j)-th element is given by

$$(\boldsymbol{a}_i^H \otimes \boldsymbol{b}_i^H)(\boldsymbol{C} \otimes \boldsymbol{D})(\boldsymbol{e}_j \otimes \boldsymbol{f}_j) = (\boldsymbol{a}_i^H \boldsymbol{C} \boldsymbol{e}_j)(\boldsymbol{b}_i^H \boldsymbol{D} \boldsymbol{f}_j).$$

Observing that  $a_i^H C e_j$  is the (i, j)-th element of  $A^H C E$ , and that  $b_i^H D f_j$  is the (i, j)-th element of  $B^H D F$ , we immediately conclude that the left-hand side is equal to the right-hand side in (22).

Lemma 2 (Woodbury matrix inversion lemma [39]):

$$(A+UCV)^{-1} = A^{-1} - A^{-1}U(C^{-1}+VA^{-1}U)^{-1}VA^{-1}$$

Lemma 3: Let A be nonsingular and B have a sufficiently small norm. Then

$$(A+B)^{-1} \approx A^{-1} - A^{-1}BA^{-1}.$$
 (24)

*Proof:* For **B** with a sufficiently small norm, the spectral radius of  $A^{-1}B$  will be less than one, and the Taylor series expansion of  $(A + B)^{-1}$  converges [39, P. 421]. Therefore, (24) is just the first-order Taylor approximation. *Lemma 4:* 

$$\sum_{k=0}^{n-1} \sin kd = \frac{\sin \frac{nd}{2}}{\sin \frac{d}{2}} \sin \frac{(n-1)d}{2},$$
$$\sum_{k=0}^{n-1} \cos kd = \frac{\sin \frac{nd}{2}}{\sin \frac{d}{2}} \cos \frac{(n-1)d}{2}.$$

*Proof:* Obtained by considering the real and imaginary parts of the sum  $\sum_{k=0}^{n-1} e^{jkd}$ .

Lemma 5:

$$\sum_{k=0}^{n-1} k \sin kd = \frac{(n-1)\sin(nd) - n\sin((n-1)d)}{2(\cos d - 1)},$$
$$\sum_{k=0}^{n-1} k \cos kd = \frac{(n-1)\cos(nd) - n\cos((n-1)d) + 1}{2(\cos d - 1)}.$$

*Proof:* Obtained by differentiating both sides of the equations in Lemma 4 with respect to d.

Lemma 6:

$$\sum_{k=0}^{n-1} k^2 \sin kd = \frac{n(n-1)[\sin(nd) - \sin((n-1)d)]}{2(\cos d - 1)} - \frac{\sin d[(n-1)\cos(nd) - n\cos((n-1)d) + 1]}{2(\cos d - 1)^2},$$
$$\sum_{k=0}^{n-1} k^2 \cos kd = \frac{n(n-1)[\cos(nd) - \cos((n-1)d)]}{2(\cos d - 1)} + \frac{\sin d[(n-1)\sin(nd) - n\sin((n-1)d)]}{2(\cos d - 1)^2}.$$

*Proof:* Obtained by differentiating both sides of the equations in Lemma 5 with respect to d.

## APPENDIX B Proof of Theorem 1

Without loss of generality, we assume that N = 1. We already know that when P is diagonal, the following inequalities hold:

$$J_{\omega\omega}^{-1} \preceq B_{(\text{sto-uc})} \preceq B_{(\text{sto})}.$$
(25)

It suffices to show that  $J_{\omega\omega}^{-1} \doteq B_{(\text{sto})}$ . We will make use of the following lemma:

*Lemma 7:* For sufficiently small  $\sigma$ ,  $\sigma R^{-1} = \Pi_A^{\perp} + O(\sigma)$ . *Proof:* By Lemma 2, we have

$$\sigma \boldsymbol{R}^{-1} = \boldsymbol{I} - \boldsymbol{A} (\sigma \boldsymbol{P}^{-1} + \boldsymbol{A}^{H} \boldsymbol{A})^{-1} \boldsymbol{A}^{H}.$$
 (26)

Because  $\mathbf{A}^{H}\mathbf{A}$  is full rank, by Lemma 3,  $(\sigma \mathbf{P}^{-1} + \mathbf{A}^{H}\mathbf{A})^{-1} = (\mathbf{A}^{H}\mathbf{A})^{-1} + O(\sigma)$ .

Using the above lemma, we observe that

$$\sigma \Re[(\dot{A}^{H}R^{-1}\dot{A})^{*} \circ (PA^{H}R^{-1}AP)]$$
  
=  $\Re[(\dot{A}^{H}(\sigma R^{-1})\dot{A})^{*} \circ (PA^{H}R^{-1}AP)]$   
=  $\Re[(\dot{A}^{H}\Pi_{A}^{\perp}\dot{A})^{*} \circ (PA^{H}R^{-1}AP) + O(\sigma)].$ 

Because that  $\Pi_A^{\perp} A = 0$ , we have

$$\sigma \Re[(\dot{A}^{H}R^{-1}A)^{*} \circ (PA^{H}R^{-1}\dot{A}P)] = \Re[(\dot{A}^{H}(\sigma R^{-1})A)^{*} \circ (PA^{H}R^{-1}\dot{A}P)] = O(\sigma).$$

Combined with the fact that  $\Re(\mathbf{X}) = \Re(\mathbf{X}^*)$ , we have

$$\begin{aligned} \boldsymbol{J}_{\boldsymbol{\omega}\boldsymbol{\omega}}^{-1} &= \frac{\sigma}{2} \big\{ \sigma \Re[(\dot{\boldsymbol{A}}^{H}\boldsymbol{R}^{-1}\dot{\boldsymbol{A}})^{*} \circ (\boldsymbol{P}\boldsymbol{A}^{H}\boldsymbol{R}^{-1}\boldsymbol{A}\boldsymbol{P})] \\ &+ \sigma \Re[(\dot{\boldsymbol{A}}^{H}\boldsymbol{R}^{-1}\boldsymbol{A})^{*} \circ (\boldsymbol{P}\boldsymbol{A}^{H}\boldsymbol{R}^{-1}\dot{\boldsymbol{A}}\boldsymbol{P})] \big\}^{-1} \\ &= \frac{\sigma}{2} \big\{ \Re[(\dot{\boldsymbol{A}}^{H}\boldsymbol{\Pi}_{\boldsymbol{A}}^{\perp}\dot{\boldsymbol{A}})^{*} \circ (\boldsymbol{P}\boldsymbol{A}^{H}\boldsymbol{R}^{-1}\boldsymbol{A}\boldsymbol{P}) + O(\sigma)] \big\}^{-1} \\ &= \frac{\sigma}{2} \big\{ \Re[(\dot{\boldsymbol{A}}^{H}\boldsymbol{\Pi}_{\boldsymbol{A}}^{\perp}\dot{\boldsymbol{A}}) \circ (\boldsymbol{P}\boldsymbol{A}^{H}\boldsymbol{R}^{-1}\boldsymbol{A}\boldsymbol{P})^{T}] + \Re[O(\sigma)] \big\}^{-1} \end{aligned}$$

By Lemma 3, we obtain that  $J_{\omega\omega}^{-1} \doteq B_{(sto)}$ .

$$A^{H}R^{-2}A = \sigma^{-2}A^{H}A(\sigma P^{-1} + A^{H}A)^{-1}[\sigma P^{-1} + A^{H}A - 2A^{H}A + A^{H}A(\sigma P^{-1} + A^{H}A)^{-1}A^{H}A]$$
  
=  $\sigma^{-2}A^{H}A(\sigma P^{-1} + A^{H}A)^{-1}[\sigma P^{-1} - A^{H}A(\sigma P^{-1} + A^{H}A)^{-1}(\sigma P^{-1} + A^{H}A - A^{H}A)]$   
=  $\sigma^{-2}A^{H}A(\sigma P^{-1} + A^{H}A)^{-1}(\sigma P^{-1} + A^{H}A - A^{H}A)(\sigma P^{-1} + A^{H}A)^{-1}\sigma P^{-1}$   
=  $A^{H}A(\sigma P^{-1} + A^{H}A)^{-1}P^{-1}(\sigma P^{-1} + A^{H}A)^{-1}P^{-1}.$  (27)

# APPENDIX C PROOF OF PROPOSITION 1

Following [26, Appendix G], for ULAs with a large number of sensors, M, we have

$$\frac{1}{M}\boldsymbol{A}^{H}\boldsymbol{A}\approx\boldsymbol{I},\ \frac{1}{M^{2}}\boldsymbol{A}^{H}\dot{\boldsymbol{A}}\approx\frac{j}{2}\boldsymbol{I},\ \frac{1}{M^{3}}\dot{\boldsymbol{A}}^{H}\dot{\boldsymbol{A}}\approx\frac{1}{3}\boldsymbol{I}.$$
 (28)

Applying Lemma 2, the inverse of R can be rewritten as

$$\mathbf{R}^{-1} = \sigma^{-1} [\mathbf{I} - \mathbf{A} (\sigma \mathbf{P}^{-1} + \mathbf{A}^{H} \mathbf{A})^{-1} \mathbf{A}^{H}].$$
(29)

Combined with the assumption that  $SNR_i^{-1} = \sigma/p_i \ll M$ , we have

$$A^{H}R^{-1}A$$

$$=\sigma^{-1}A^{H}A[I - (\sigma P^{-1} + A^{H}A)^{-1}A^{H}A]$$

$$=\sigma^{-1}A^{H}A(\sigma P^{-1} + A^{H}A)^{-1}[\sigma P^{-1} + A^{H}A - A^{H}A]$$

$$=A^{H}A(\sigma P^{-1} + A^{H}A)^{-1}P^{-1}$$

$$\approx P^{-1},$$
(30)

$$\dot{A}^{H}R^{-1}A$$

$$=\sigma^{-1}[\dot{A}^{H}A - \dot{A}^{H}A(\sigma P^{-1} + A^{H}A)^{-1}A^{H}A]$$

$$=\sigma^{-1}\dot{A}^{H}A(\sigma P^{-1} + A^{H}A)^{-1}(\sigma P^{-1} + A^{H}A - A^{H}A)$$

$$=\dot{A}^{H}A(\sigma P^{-1} + A^{H}A)^{-1}P^{-1}$$

$$\approx -j\frac{M}{2}P^{-1},$$
(31)

and

$$\dot{A}^{H} \mathbf{R}^{-1} \dot{A}$$

$$= \sigma^{-1} [\dot{A}^{H} \dot{A} - \dot{A}^{H} \mathbf{A} (\sigma \mathbf{P}^{-1} + \mathbf{A}^{H} \mathbf{A})^{-1} \mathbf{A}^{H} \dot{A}] \qquad (32)$$

$$\approx \frac{M^{3}}{12} \sigma^{-1} \mathbf{I}.$$

Substituting (30)–(32) into the expression of  $J_{\omega\omega}$ , we obtain

$$\boldsymbol{J}_{\boldsymbol{\omega}\boldsymbol{\omega}} = \frac{M^3}{6} \sigma^{-1} \boldsymbol{P}.$$

Using similar tricks, we can obtain the following:

$$\operatorname{tr}(\boldsymbol{R}^{-2}) \approx \sigma^{-2}(M - K),\tag{33}$$

$$\boldsymbol{A}^{H}\boldsymbol{R}^{-2}\boldsymbol{A} = \boldsymbol{A}^{H}\boldsymbol{A}[(\boldsymbol{\sigma}\boldsymbol{P}^{-1} + \boldsymbol{A}^{H}\boldsymbol{A})\boldsymbol{P}]^{-2} \approx \frac{\boldsymbol{P}^{-2}}{M}, \quad (34)$$
$$\dot{\boldsymbol{A}}^{H}\boldsymbol{R}^{-2}\boldsymbol{A} = \dot{\boldsymbol{A}}^{H}\boldsymbol{A}[(\boldsymbol{\sigma}\boldsymbol{P}^{-1} + \boldsymbol{A}^{H}\boldsymbol{A})\boldsymbol{P}]^{-2} \approx -j\frac{\boldsymbol{P}^{-2}}{2}.$$
$$(35)$$

The detailed derivation of (34) is summarized in (27). The derivation of (35) follows the same idea.

By (30), (34) and (33), we obtain that  $J_{pp} \approx P^{-2}$ ,  $J_{p\sigma} \approx \text{diag}(P^{-2})/M$  and that  $J_{\sigma\sigma} \approx \sigma^{-2}(M-K)$ . Therefore, using block-wise inversion, we get

$$egin{bmatrix} egin{aligned} egi$$

Denote the above inverse as  $K_1^{-1}$  and define  $K_2 = [J_{\omega p} \ J_{\omega \sigma}]$ . We have

$$oldsymbol{B}_{( ext{sto-uc})}(oldsymbol{\omega}) = (oldsymbol{J}_{oldsymbol{\omega}oldsymbol{\omega}} - oldsymbol{K}_2oldsymbol{K}_1^{-1}oldsymbol{K}_2^H)^{-1}$$

According to (31) and (35), in the expressions of  $J_{\omega p}$ and  $J_{\omega \sigma}$ , the terms inside the  $\Re(\cdot)$  operator will be almost imaginary. Therefore, both  $J_{\omega p}$  and  $J_{\omega \sigma}$  will be approximately zeros for large values of M. Combined with our previous approximations of  $K_1^{-1}$  and  $J_{\omega \omega}$ , we conclude that  $K_2^H K_1^{-1} K_2$  becomes negligible compared with  $J_{\omega \omega}$  for large values of M. Therefore,

$$B_{(\text{sto-uc})}(\boldsymbol{\omega}) \approx \frac{1}{N} J_{\boldsymbol{\omega}\boldsymbol{\omega}}^{-1} = \frac{6}{M^3 N} \sigma \boldsymbol{P}^{-1}.$$
Appendix D
PROOF OF THEOREM 2 and 4

#### A. The Nested Array Case

Given a nested array configured with the parameter pair  $(N_1, N_2)$ , its steering vector for the one-source case is given by  $\boldsymbol{a} = [\boldsymbol{a}_1^T \boldsymbol{a}_2^T]^T$ , where

$$\boldsymbol{a}_{1}^{T} = \begin{bmatrix} e^{j\omega} & e^{j2\omega} & \cdots & e^{jN_{1}\omega} \end{bmatrix}, \tag{36}$$

$$\boldsymbol{a}_{2}^{I} = \begin{bmatrix} e^{j(N_{1}+1)\omega} & e^{j2(N_{1}+1)\omega} & \cdots & e^{jN_{2}(N_{1}+1)\omega} \end{bmatrix}.$$
(37)

With respect to  $\boldsymbol{\omega}$ , the derivative vector  $\dot{\boldsymbol{a}}$  is given by  $\dot{\boldsymbol{a}} = j\boldsymbol{D}\boldsymbol{a}$ , where  $\boldsymbol{D} = \text{diag}(\boldsymbol{D}_1, \boldsymbol{D}_2)$ , and

$$D_1 = \text{diag}(1, 2, \dots, N_1),$$
  

$$D_2 = \text{diag}(N_1 + 1, 2(N_1 + 1), \dots, N_2(N_1 + 1)).$$
(38)

By letting  $N_1 = \mu M$ ,  $N_2 = (1 - \mu)M$ , we can approximate the following terms as

$$a^{H}a = M,$$

$$\dot{a}^{H}a = -ja^{H}Da = -j\left[\sum_{q=1}^{\mu M} q + \sum_{q=1}^{(1-\mu)M} q(\mu M + 1)\right]$$

$$\approx -j\frac{1}{2}\mu(1-\mu)^{2}M^{3},$$
(40)

$$\dot{a}^{H}\dot{a} = a^{H}D^{2}a = \sum_{q=1}^{\mu M} q^{2} + \sum_{q=1}^{(1-\mu)M} [q(\mu M+1)]^{2}$$
$$\approx \frac{1}{3}\mu^{2}(1-\mu)^{3}M^{5}, \qquad (41)$$

where the approximations are obtained by removing terms that are at least one-order smaller than the highest order terms.

We can calculate the inverse of  $\boldsymbol{R}$  from Lemma 2 as

$$\boldsymbol{R}^{-1} = \sigma^{-1} \left[ \boldsymbol{I} - \frac{\boldsymbol{a} \boldsymbol{a}^{H}}{\sigma p^{-1} + M} \right].$$
(42)

Hence, under the assumption that  $SNR^{-1} \ll M$ , we have

$$\begin{split} \boldsymbol{a}^{H}\boldsymbol{R}^{-1}\boldsymbol{a} &= \sigma^{-1}\Big[\boldsymbol{a}^{H}\boldsymbol{a} - \frac{\boldsymbol{a}^{H}\boldsymbol{a}\boldsymbol{a}^{H}\boldsymbol{a}}{\sigma p^{-1} + M}\Big] \\ &= \sigma^{-1}\Big[\boldsymbol{M} - \frac{\boldsymbol{M}^{2}}{\sigma p^{-1} + M}\Big] \\ &\approx p^{-1}, \\ \dot{\boldsymbol{a}}^{H}\boldsymbol{R}^{-1}\boldsymbol{a} &= \sigma^{-1}\Big[\dot{\boldsymbol{a}}^{H}\boldsymbol{a} - \frac{\dot{\boldsymbol{a}}^{H}\boldsymbol{a}\boldsymbol{a}^{H}\boldsymbol{a}}{\sigma p^{-1} + M}\Big] \\ &\approx p^{-1}\frac{-j\mu(1-\mu)^{2}M^{3}}{2(\sigma p^{-1} + M)} \\ &\approx -j\frac{1}{2}\mu(1-\mu)^{2}M^{2}p^{-1}, \end{split}$$

and

$$\begin{split} \dot{a}^{H} \mathbf{R}^{-1} \dot{a} = &\sigma^{-1} \Big[ \dot{a}^{H} \dot{a} - \frac{\dot{a}^{H} a a^{H} \dot{a}}{\sigma p^{-1} + M} \Big] \\ \approx &\sigma^{-1} \left[ \frac{1}{3} \mu^{2} (1 - \mu)^{3} M^{5} - \frac{\mu^{2} (1 - \mu)^{4} M^{6}}{4 (\sigma p^{-1} + M)} \right] \\ = &\frac{1}{12} \mu^{2} (1 - \mu)^{3} (1 + 3\mu) M^{5} \sigma^{-1}. \end{split}$$

Observing that  $\dot{a}^{H}R^{-1}a = ja^{H}DR^{-1}a$  and  $\dot{a}^{H}R^{-2}a = ja^{H}DR^{-2}a$  are both purely imaginary, and that  $a^{H}R^{-1}a$  is real, we immediately know that  $J_{\omega p}$  and  $J_{\omega \sigma}$  are exactly zero. Hence, the FIM takes the following form:

$$\boldsymbol{J} = N \begin{bmatrix} J_{\omega\omega} & 0 & 0\\ 0 & * & *\\ 0 & * & * \end{bmatrix}.$$
 (43)

Therefore, to obtain  $B_{(\text{sto-uc})}(\omega)$ , we need to evaluate only  $J_{\omega\omega}$ , which is given by

$$J_{\omega\omega} = 2\Re[(\dot{a}^{H}R^{-1}\dot{a})^{*} \circ (p^{2}a^{H}R^{-1}a) + (\dot{a}^{H}R^{-1}a)^{*} \circ (p^{2}a^{H}R^{-1}\dot{a})]$$
  
$$= 2\Re\left[\frac{\mu^{2}}{12}(1-\mu)^{3}(1+3\mu)M^{5}p\sigma^{-1} + \frac{\mu^{2}}{4}(1-\mu)^{4}M^{4}\right]$$
  
$$\approx \frac{1}{6}\mu^{2}(1-\mu)^{3}(1+3\mu)M^{5}p\sigma^{-1}.$$

Therefore,

$$\boldsymbol{B}_{(\text{sto-uc})}(\omega) = \frac{1}{N} J_{\omega\omega}^{-1} \approx \frac{6}{\mu^2 (1-\mu)^3 (1+3\mu)} \frac{1}{N} \frac{1}{M^5} \frac{1}{\text{SNR}}.$$

## B. The Co-prime Array Case

In the one-source case, the steering matrix A of a co-prime array reduces to a vector  $a = [a_1^T \ a_2^T]^T$ , where

$$a_1^T = \begin{bmatrix} 1 & e^{jN_1\omega} & \cdots & e^{j(N_2-1)N_1\omega} \end{bmatrix},$$
(44)  
$$a_2^T = \begin{bmatrix} e^{jN_2\omega} & e^{j2N_2\omega} & \cdots & e^{j(2N_1-1)N_2\omega} \end{bmatrix}.$$
(45)

With respect to  $\boldsymbol{\omega}$ , the derivative vector  $\dot{\boldsymbol{a}}$  is given by  $\dot{\boldsymbol{a}} = j\boldsymbol{D}\boldsymbol{a}$ , where  $\boldsymbol{D} = \text{diag}(\boldsymbol{D}_1, \boldsymbol{D}_2)$ , and

$$D_1 = \text{diag}(0, N_1, \dots, (N_2 - 1)N_1),$$
  

$$D_2 = \text{diag}(N_2, 2N_2, \dots, (2N_1 - 1)N_2).$$
(46)

Similar to the nested array case, by setting  $N_1 = \mu(M+1)$ and  $N_2 = (1-2\mu)(M+1)$ , we can approximate the following terms:

$$\begin{split} \mathbf{a}^{H}\mathbf{a} &= M, \\ \dot{\mathbf{a}}^{H}\mathbf{a} &= -j \bigg[ \sum_{q=1}^{(1-2\mu)(M+1)-1} q\mu(M+1) + \\ & \sum_{q=1}^{2\mu(M+1)-1} q(1-2\mu)(M+1) \bigg] \\ &\approx -j\frac{1}{2}\mu(1-4\mu^{2})M^{3}. \\ \dot{\mathbf{a}}^{H}\dot{\mathbf{a}} &= \sum_{q=1}^{(1-2\mu)(M+1)-1} q^{2}(\mu(M+1))^{2} + \\ & \sum_{q=1}^{2\mu(M+1)-1} q^{2}(1-2\mu)^{2}(M+1)^{2} \\ &\approx \frac{1}{3}\mu^{2}(1-2\mu)^{2}(1+6\mu)M^{5}. \end{split}$$

Combined with the assumption that  $SNR^{-1} \ll M$ , we obtain that  $a^H \mathbf{P}^{-1} \mathbf{a} \sim n^{-1}$ 

$$\dot{\boldsymbol{a}}^{H} \boldsymbol{R}^{-1} \boldsymbol{a} \approx -j \frac{1}{2} \mu (1-4\mu^2) M^2 p^{-1},$$

and

$$\dot{a}^{H} \mathbf{R}^{-1} \dot{a} \approx \frac{1}{12} \mu^{2} (1 - 2\mu)^{2} (1 + 12\mu - 12\mu^{2}) M^{5} \sigma^{-1}$$

Similar to the nested array case, to obtain  $B_{(\text{sto-uc})}(\omega)$ , we need to evaluate only  $J_{\omega\omega}$ , which is given by

$$J_{\omega\omega} = 2\Re \Big[ \frac{1}{12} \mu^2 (1 - 2\mu)^2 (1 + 12\mu - 12\mu^2) M^5 p \sigma^{-1} + \frac{1}{4} \mu^2 (1 - 4\mu^2)^2 M^4 \Big]$$
$$\approx \frac{1}{6} \mu^2 (1 - 2\mu)^2 (1 + 12\mu - 12\mu^2) M^5 p \sigma^{-1}.$$

Therefore,

$$\boldsymbol{B}_{(\text{sto-uc})}(\omega) \approx \frac{6}{\mu^2 (1 - 2\mu)^2 (1 + 12\mu - 12\mu^2)} \frac{1}{N} \frac{1}{M^5} \frac{1}{\text{SNR}}$$

## APPENDIX E Proof of Theorem 3

## A. The Non-degenerate Case

We begin by proving the first part of the theorem. In the multiple-source case, the steering matrix of the nested array generated by the parameter pair  $(N_1, N_2)$  can be expressed as  $\boldsymbol{A} = [\boldsymbol{A}_1^T \ \boldsymbol{A}_2^T]^T$ , where

$$\boldsymbol{A}_1 = \begin{bmatrix} \boldsymbol{a}_1(\omega_1) & \boldsymbol{a}_1(\omega_2) & \cdots & \boldsymbol{a}_1(\omega_K) \end{bmatrix}, \\ \boldsymbol{A}_2 = \begin{bmatrix} \boldsymbol{a}_2(\omega_1) & \boldsymbol{a}_2(\omega_2) & \cdots & \boldsymbol{a}_2(\omega_K) \end{bmatrix},$$
 (47)

and  $a_1$ ,  $a_2$  follow the same definitions as those in (36), (37). We also have  $\dot{A} = jDA$ , where  $D = \text{diag}(D_1, D_2)$  follows the same definition as that in (38). Therefore,

$$\begin{bmatrix} \mathbf{A}^{H} \mathbf{A} \end{bmatrix}_{m,n} = \begin{bmatrix} \mathbf{A}_{1}^{H} \mathbf{A}_{1} \end{bmatrix}_{m,n} + \begin{bmatrix} \mathbf{A}_{2}^{H} \mathbf{A}_{2} \end{bmatrix}_{m,n}$$
$$= \sum_{q=1}^{N_{1}} e^{jq(\omega_{m}-\omega_{n})} + \sum_{q=1}^{N_{2}} e^{jq(N_{1}+1)(\omega_{m}-\omega_{n})}.$$
(48)

Here  $[\cdot]_{m,n}$  denotes the (m, n)-th element.

The diagonal elements of  $A^H A$  simply reduces to M. If we can bound the off-diagonal elements, we can then approximate  $A^H A$  with MI for a sufficiently large M. If  $\dot{A}^H A$  and  $\dot{A}^H \dot{A}$  can be approximated in a similar way, we can approximate  $B_{(\text{sto-uc})}(\omega)$  with an approach similar to the one we used in Appendix C. To make such approximations possible, we introduce the following lemma.

*Lemma* 8: Let Q, L be positive integers and  $0 < \delta < 1$ . Let  $\omega \in (-2\pi, 2\pi) \cap \Omega_L^{\delta}$ , where  $\Omega_L^{\delta}$  follows the same definition as in Definition 3. Then

$$\begin{split} \left| \sum_{q=0}^{Q-1} e^{jq\omega L} \right| &\leq \delta^{-1}, \left| \sum_{q=0}^{Q-1} q e^{jq\omega L} \right| \leq \frac{Q}{2} \delta^{-2}, \\ \left| \sum_{q=0}^{Q-1} q^2 e^{jq\omega L} \right| &\leq \frac{1}{2} (Q^2 \delta^{-2} + Q \delta^{-3}). \end{split}$$

*Proof:* By the definition of  $\Omega_L^{\delta}$ , we know that  $|\sin(\omega L/2)|^{-1} \leq \delta^{-1}$ . By Lemma 4, the left-hand-side (LHS) of the first inequality follows

$$LHS = \left| \frac{\sin \frac{\omega QL}{2}}{\sin \frac{\omega L}{2}} e^{j \frac{(Q-1)\omega L}{2}} \right| \le \left| \sin \frac{\omega L}{2} \right|^{-1} \le \delta^{-1}.$$

By Lemma 5, the LHS of the second inequality follows

$$\begin{split} \text{LHS} &= \left| \frac{(Q-1)e^{jQ\omega L} - Qe^{j(Q-1)\omega L} + 1}{2(\cos(\omega L) - 1)} \right| \\ &\leq \frac{|(Q-1)e^{jQ\omega L}| + |Qe^{j(Q-1)\omega L}| + 1}{2|\cos(\omega L) - 1|} \\ &= \frac{Q}{2|\sin^2 \frac{\omega L}{2}|} \\ &\leq \frac{Q}{2}\delta^{-2}. \end{split}$$

By Lemma 6, the LHS of the third equality follows

$$\begin{split} \mathsf{LHS} &\leq \left| \frac{Q(Q-1)(e^{jQ\omega L} - e^{j(Q-1)\omega L})}{2(\cos(\omega L) - 1)} \right| \\ &+ \left| \frac{(Q-1)e^{j(Q\omega L - \frac{\pi}{2})} - Qe^{j((Q-1)\omega L - \frac{\pi}{2})} + j}{2(\cos(\omega L) - 1)^2(\sin(\omega L))^{-1}} \right| \\ &\leq \left| \frac{Q^2}{\cos(\omega L) - 1} \right| + \left| \frac{Q\sin(\omega L)}{(\cos(\omega L) - 1)^2} \right| \\ &\leq \frac{Q^2}{2\sin^2 \frac{\omega L}{2}} + \frac{Q}{2\left| \sin^3 \frac{\omega L}{2} \right|} \\ &\leq \frac{1}{2}(Q^2 \delta^{-2} + Q\delta^{-3}). \end{split}$$

Under the assumption that the K-DOAs are  $\delta$ -level nondegenerate with respect to 1 and  $N_1 + 1$ , we immediately know from Lemma 8 that when  $m \neq n$ ,

$$\begin{split} |[\mathbf{A}^{H}\mathbf{A}]_{m,n}| &\leq 2\delta^{-1}, \\ |[\dot{\mathbf{A}}^{H}\mathbf{A}]_{m,n}| &\leq \frac{M}{2}\delta^{-2}, \\ |[\dot{\mathbf{A}}^{H}\dot{\mathbf{A}}]_{m,n}| &\leq \frac{2\mu^{2} - 2\mu + 1}{2}M^{2}\delta^{-2} + \frac{M}{2}\delta^{-3}. \end{split}$$

Here we use the fact that  $N_1 + N_2 = M$  and that  $\mu = N_1/M$ . When m = n, we know from (39)–(41) that,

$$\begin{split} & [\mathbf{A}^{H}\mathbf{A}]_{m,m} = M, \\ & [\dot{\mathbf{A}}^{H}\mathbf{A}]_{m,m} \approx -j\frac{1}{2}\mu(1-\mu)^{2}M^{3} \\ & [\dot{\mathbf{A}}^{H}\dot{\mathbf{A}}]_{m,m} \approx \frac{1}{3}\mu^{2}(1-\mu)^{3}M^{5}. \end{split}$$

Because  $\mu$  and  $\delta$  are fixed, for a sufficiently large M, the off-diagonal elements in  $A^H A$ ,  $\dot{A}^H A$ , and  $\dot{A}^H \dot{A}$  are indeed negligible compared with the their corresponding diagonal elements, leading to

$$\begin{split} \mathbf{A}^{H}\mathbf{A} &\approx M\mathbf{I}, \\ \dot{\mathbf{A}}^{H}\mathbf{A} &\approx -j\frac{1}{2}\mu(1-\mu)^{2}M^{3}\mathbf{I}, \\ \dot{\mathbf{A}}^{H}\dot{\mathbf{A}} &\approx \frac{1}{3}\mu^{2}(1-\mu)^{3}M^{5}\mathbf{I}. \end{split}$$

Following the derivations of (30)–(32), we obtain that, in the nested array case,

$$\boldsymbol{A}^{H}\boldsymbol{R}^{-1}\boldsymbol{A}\approx\boldsymbol{P}^{-1},\tag{49}$$

$$\dot{\boldsymbol{A}}^{H}\boldsymbol{R}^{-1}\boldsymbol{A}\approx -j\frac{\mu(1-\mu)^{2}}{2}M^{2}\boldsymbol{P}^{-1},$$
(50)

$$\dot{A}^{H}R^{-1}\dot{A} \approx \frac{1}{12}\mu^{2}(1-\mu)^{3}(1+3\mu)M^{5}\sigma^{-1}I.$$
 (51)

Following the derivations of (33)-(35), we also obtain that

$$\operatorname{tr}(\boldsymbol{R}^{-2}) \approx \sigma^{-2}(M-K),$$
$$\boldsymbol{A}^{H}\boldsymbol{R}^{-2}\boldsymbol{A} \approx \frac{\boldsymbol{P}^{-2}}{M},$$
$$\dot{\boldsymbol{A}}^{H}\boldsymbol{R}^{-2}\boldsymbol{A} \approx -j\frac{1}{2}\mu(1-\mu)^{2}M\boldsymbol{P}^{-2}$$

We can observe that the approximations of  $A^{H}R^{-1}A$ , tr( $R^{-2}$ ), and  $A^{H}R^{-2}A$  are the same as those in the ULA case. Additionally, both  $\dot{A}^{H}R^{-1}A$  and  $\dot{A}^{H}R^{-2}A$  are approximately imaginary. Following the same reasoning as in Appendix C, we only need to compute  $J_{\omega\omega}$  to approximate  $B_{(\text{sto-uc})}(\omega)$ . Combining (49)–(51) with the definition of  $J_{\omega\omega}$ , we obtain that

$$\boldsymbol{J}_{\boldsymbol{\omega}\boldsymbol{\omega}} \approx \frac{1}{6}\mu^2(1-\mu)^3(1+3\mu)M^5\sigma^{-1}\boldsymbol{P}.$$

Therefore

$$\boldsymbol{B}_{(\text{sto-uc})}(\boldsymbol{\omega}) \approx \frac{1}{N} \boldsymbol{J}_{\boldsymbol{\omega}\boldsymbol{\omega}} \approx \frac{1}{h_{\text{ne}}(\mu)} \frac{1}{N} \frac{1}{M^5} \sigma \boldsymbol{P}^{-1}.$$
 (52)

## B. The Fully Degenerate Case

For brevity, we denote  $\omega_m - \omega_n$  by  $\omega_{mn}$ . In the following discussion,  $D_1$ ,  $D_2$ ,  $A_1$ , and  $A_2$  follow the same definitions as in (38) and (47). In the fully degenerate case, we have  $(N_1 + 1)\omega_{mn} = 2k\pi$  for some non-zero integer k whenever  $m \neq n$ . Therefore  $A_2^H A_2 = N_2 \mathbf{1}_K \mathbf{1}_K$ , where  $\mathbf{1}_K \mathbf{1}_K$  is  $K \times K$  matrix of ones. For a fixed  $\mu$ , when M is sufficiently large,  $N_1$  is also large and  $A_1^H A_1 \approx N_1 \mathbf{I}$  [26]. Hence, we have

$$\boldsymbol{A}^{H}\boldsymbol{A} \approx N_{1}\boldsymbol{I} + N_{2}\boldsymbol{1}_{K}\boldsymbol{1}_{K}^{T}$$
$$= \mu M\boldsymbol{I} + (1-\mu)M\boldsymbol{1}_{K}\boldsymbol{1}_{K}^{T}.$$
(53)

We next evaluate  $A^H D A = A_1^H D_1 A_1 + A_2^H D_2 A_2$ , where

$$[\boldsymbol{A}_{1}^{H}\boldsymbol{D}_{1}\boldsymbol{A}_{1}]_{m,n} = \sum_{q=1}^{N_{1}} q e^{jq\omega_{nm}},$$
(54)

$$[\boldsymbol{A}_{2}^{H}\boldsymbol{D}_{2}\boldsymbol{A}_{2}]_{m,n} = \sum_{q=1}^{N_{2}} q(N_{1}+1)e^{jq(N_{1}+1)\omega_{nm}}.$$
 (55)

Combining Lemma 5 with the fact that  $(N_1 + 1)\omega_{mn} = 2k\pi$ , we have

$$\begin{split} & [\boldsymbol{A}_{1}^{H}\boldsymbol{D}_{1}\boldsymbol{A}_{1}]_{m,n} \\ = & \frac{N_{1}\cos((N_{1}+1)\omega_{nm}) - (N_{1}+1)\cos(N_{1}\omega_{nm}) + 1}{2(\cos\omega_{nm}-1)} \\ & + j\frac{N_{1}\sin((N_{1}+1)\omega_{nm}) - (N_{1}+1)\sin(N_{1}\omega_{nm})}{2(\cos\omega_{nm}-1)} \\ = & \frac{N_{1}+1}{2(\cos\omega_{nm}-1)} \left[1 - \cos(N_{1}\omega_{nm}) - j\sin(N_{1}\omega_{nm})\right], \end{split}$$

when  $m \neq n$ ,

$$[\boldsymbol{A}_{1}^{H}\boldsymbol{D}_{1}\boldsymbol{A}_{1}]_{m,n} = \frac{1}{2}N_{1}(N_{1}+1),$$

when m = n, and

$$[\boldsymbol{A}_2^H \boldsymbol{D}_2 \boldsymbol{A}_2]_{m,n} = \sum_{q=1}^{N_2} q(N_1 + 1) = \frac{1}{2} N_2 (N_2 + 1)(N_1 + 1).$$

When M is sufficiently large,  $A_1^H D_1 A_1$  becomes negligible. Substituting  $N_1$  with  $\mu M$  and  $N_2$  with  $(1 - \mu)M$ , we obtain that

$$\dot{\boldsymbol{A}}^{H}\boldsymbol{A}\approx -j\boldsymbol{A}_{2}^{H}\boldsymbol{D}_{2}\boldsymbol{A}_{2}\approx -j\frac{\mu(1-\mu)^{2}}{2}M^{3}\boldsymbol{1}_{K}\boldsymbol{1}_{K}^{T}.$$
 (56)

Similar to the non-degenerate case, we can still show that  $B_{(\text{sto-uc})} \approx J_{\omega\omega}^{-1}/N$ . However, because  $A^H A$ ,  $\dot{A}^H A$  can no longer be approximated as diagonal matrices, we need an alternative way of approximating  $J_{\omega\omega}$ .

According to (31),

$$\dot{\boldsymbol{A}}^{H}\boldsymbol{R}^{-1}\boldsymbol{A} = \dot{\boldsymbol{A}}^{H}\boldsymbol{A}(\sigma\boldsymbol{P}^{-1} + \boldsymbol{A}^{H}\boldsymbol{A})^{-1}\boldsymbol{P}^{-1}$$

Combined with (53) and (56), we know that  $\dot{A}^{H}R^{-1}A$  is  $O(M^{2})$ . Hence,  $(\dot{A}^{H}R^{-1}A)^{*} \circ (PA^{H}R^{-1}\dot{A}P)$  is  $O(M^{4})$ . One the other hand, following the steps in (30), we have

$$(\dot{A}^{H}R^{-1}\dot{A})^{*} \circ (PA^{H}R^{-1}AP) \approx (\dot{A}^{H}R^{-1}\dot{A})^{*} \circ P,$$
(57)

which is approximately diagonal. Therefore, we need to only evaluate the diagonals of  $\dot{A}^H R^{-1} \dot{A}$ . According (32), we have

$$\sigma[\dot{A}^{H}R^{-1}\dot{A}]_{k,k}\approx[\dot{A}^{H}\dot{A}]_{k,k}-[\dot{A}^{H}A(A^{H}A)^{-1}A^{H}\dot{A}]_{k,k}$$

Following (41), the first term is given by

$$[\dot{A}^{H}\dot{A}]_{k,k} \approx \frac{\mu^{2}(1-\mu)^{3}}{3}M^{5}.$$
 (58)

To evaluate the second term, we need to first evaluate  $(A^H A)^{-1}$ . By Lemma 2,

$$(\boldsymbol{A}^{H}\boldsymbol{A})^{-1} = -\frac{1}{\mu M} \left[ \frac{1}{K + \frac{\mu}{1-\mu}} \boldsymbol{1}_{K} \boldsymbol{1}_{K}^{T} - \boldsymbol{I} \right].$$

Combined with (56), we obtain that

$$\begin{split} & [\dot{A}^{H} A (A^{H} A)^{-1} A^{H} \dot{A}]_{k,k} \\ = & - \frac{1}{\mu M} \left[ \frac{1}{K + \frac{\mu}{1-\mu}} \left| \sum_{j=1}^{K} [\dot{A}^{H} A]_{k,j} \right|^{2} - \sum_{j=1}^{K} \left| [\dot{A}^{H} A]_{k,j} \right|^{2} \right] \\ &\approx & \frac{K}{4} \frac{\mu^{2} (1-\mu)^{4}}{\mu + (1-\mu)K} M^{5}. \end{split}$$

Combining the above result with (58), we have

$$\sigma[\dot{\mathbf{A}}^{H}\mathbf{R}^{-1}\dot{\mathbf{A}}]_{k,k} \approx \frac{\mu^{2}(1-\mu)^{3}}{12} \frac{4\mu + (1-\mu)K}{\mu + (1-\mu)K} M^{5}.$$

Therefore, the term  $(\dot{A}^{H}R^{-1}\dot{A})^{*} \circ (PA^{H}R^{-1}AP)$  is  $O(M^{5})$  and the term  $(\dot{A}^{H}R^{-1}A)^{*} \circ (PA^{H}R^{-1}\dot{A}P)$  will be negligible when M is sufficiently large, leading to

$$J_{\omega\omega} \approx 2\Re[(\dot{A}^{H}R^{-1}\dot{A})^{*} \circ (PA^{H}R^{-1}AP)]$$
$$\approx \frac{\mu^{2}(1-\mu)^{3}}{6} \frac{4\mu + (1-\mu)K}{\mu + (1-\mu)K} M^{5}\sigma^{-1}P.$$

Finally, because  $B_{(\text{sto-uc})} \approx J_{\omega\omega}^{-1}/N$ , we obtain (17).

#### REFERENCES

- H. Krim and M. Viberg, "Two decades of array signal processing research: the parametric approach," *IEEE Signal Processing Magazine*, vol. 13, no. 4, pp. 67–94, Jul. 1996.
- [2] H. L. Van Trees, *Optimum Array Processing*, ser. Detection, Estimation, and Modulation Theory / Harry L. Van Trees. New York: Wiley, 2002, no. 4.
- [3] P. Stoica and R. L. Moses, *Spectral Analysis of Signals*. Upper Saddle River, N.J: Pearson/Prentice Hall, 2005.
- [4] R. Schmidt, "Multiple emitter location and signal parameter estimation," *IEEE Transactions on Antennas and Propagation*, vol. 34, no. 3, pp. 276–280, Mar. 1986.
- [5] B. Friedlander, "The root-MUSIC algorithm for direction finding with interpolated arrays," *Signal Processing*, vol. 30, no. 1, pp. 15–29, Jan. 1993.
- [6] M. Pesavento, A. Gershman, and M. Haardt, "Unitary root-MUSIC with a real-valued eigendecomposition: a theoretical and experimental performance study," *IEEE Transactions on Signal Processing*, vol. 48, no. 5, pp. 1306–1314, May 2000.
- [7] A. Moffet, "Minimum-redundancy linear arrays," *IEEE Transactions on Antennas and Propagation*, vol. 16, no. 2, pp. 172–175, Mar. 1968.
- [8] M. Ishiguro, "Minimum redundancy linear arrays for a large number of antennas," *Radio Science*, vol. 15, no. 6, pp. 1163–1170, Nov. 1980.
- [9] D. Linebarger, I. Sudborough, and I. Tollis, "Difference bases and sparse sensor arrays," *IEEE Transactions on Information Theory*, vol. 39, no. 2, pp. 716–721, Mar. 1993.
- [10] P. Vaidyanathan and P. Pal, "Sparse sensing with co-prime samplers and arrays," *IEEE Transactions on Signal Processing*, vol. 59, no. 2, pp. 573–586, Feb. 2011.
- [11] P. Pal and P. P. Vaidyanathan, "Coprime sampling and the music algorithm." IEEE, Jan. 2011, pp. 289–294.
- [12] Y. Zhang, M. Amin, and B. Himed, "Sparsity-based DOA estimation using co-prime arrays," in 2013 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), May 2013, pp. 3967–3971.

- [13] Z. Tan and A. Nehorai, "Sparse direction of arrival estimation using co-prime arrays with off-grid targets," *IEEE Signal Processing Letters*, vol. 21, no. 1, pp. 26–29, Jan. 2014.
- [14] Z. Tan, Y. C. Eldar, and A. Nehorai, "Direction of arrival estimation using co-prime arrays: A super resolution viewpoint," *IEEE Transactions* on Signal Processing, vol. 62, no. 21, pp. 5565–5576, Nov. 2014.
- [15] S. Qin, Y. Zhang, and M. Amin, "Generalized coprime array configurations for direction-of-arrival estimation," *IEEE Transactions on Signal Processing*, vol. 63, no. 6, pp. 1377–1390, Mar. 2015.
- [16] S. Qin, Y. D. Zhang, M. G. Amin, and F. Gini, "Frequency diverse coprime arrays with coprime frequency offsets for multitarget localization," *IEEE Journal of Selected Topics in Signal Processing*, vol. 11, no. 2, pp. 321–335, March 2017.
- [17] J. Ramirez and J. L. Krolik, "Synthetic aperture processing for passive co-prime linear sensor arrays," *Digital Signal Processing*, vol. 61, pp. 62–75, Feb. 2017.
- [18] P. Pal and P. Vaidyanathan, "Nested arrays: A novel approach to array processing with enhanced degrees of freedom," *IEEE Transactions on Signal Processing*, vol. 58, no. 8, pp. 4167–4181, Aug. 2010.
- [19] K. Han and A. Nehorai, "Wideband gaussian source processing using a linear nested array," *IEEE Signal Processing Letters*, vol. 20, no. 11, pp. 1110–1113, Nov. 2013.
- [20] —, "Nested vector-sensor array processing via tensor modeling," *IEEE Transactions on Signal Processing*, vol. 62, no. 10, pp. 2542– 2553, May 2014.
- [21] K. Han, P. Yang, and A. Nehorai, "Calibrating nested sensor arrays with model errors," *IEEE Transactions on Antennas and Propagation*, vol. 63, no. 11, pp. 4739–4748, Nov. 2015.
- [22] C. L. Liu and P. P. Vaidyanathan, "Super nested arrays: Linear sparse arrays with reduced mutual coupling — Part I: Fundamentals," *IEEE Transactions on Signal Processing*, vol. 64, no. 15, pp. 3997–4012, Aug. 2016.
- [23] —, "Super nested arrays: Linear sparse arrays with reduced mutual coupling — Part II: High-order extensions," *IEEE Transactions on Signal Processing*, vol. 64, no. 16, pp. 4203–4217, Aug. 2016.
- [24] M. Wang, Z. Zhang, and A. Nehorai, "Direction finding using sparse linear arrays with missing data," in 2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Mar. 2017, pp. 3066–3070.
- [25] H. Qiao and P. Pal, "On maximum-likelihood methods for localizing more sources than sensors," *IEEE Signal Processing Letters*, vol. 24, no. 5, pp. 703–706, May 2017.
- [26] P. Stoica and A. Nehorai, "MUSIC, maximum likelihood, and Cramer-Rao bound," *IEEE Transactions on Acoustics, Speech and Signal Processing*, vol. 37, no. 5, pp. 720–741, May 1989.
- [27] —, "MUSIC, maximum likelihood, and Cramer-Rao bound: further results and comparisons," *IEEE Transactions on Acoustics, Speech and Signal Processing*, vol. 38, no. 12, pp. 2140–2150, Dec. 1990.
- [28] —, "Performance study of conditional and unconditional direction-ofarrival estimation," *IEEE Transactions on Acoustics, Speech and Signal Processing*, vol. 38, no. 10, pp. 1783–1795, Oct. 1990.
- [29] F. Li, H. Liu, and R. Vaccaro, "Performance analysis for DOA estimation algorithms: unification, simplification, and observations," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 29, no. 4, pp. 1170–1184, Oct. 1993.
- [30] R. Roy and T. Kailath, "ESPRIT-estimation of signal parameters via rotational invariance techniques," *IEEE Transactions on Acoustics, Speech* and Signal Processing, vol. 37, no. 7, pp. 984–995, Jul. 1989.
- [31] C. Chambers, T. C. Tozer, K. C. Sharman, and T. S. Durrani, "Temporal and spatial sampling influence on the estimates of superimposed narrowband signals: when less can mean more," *IEEE Transactions on Signal Processing*, vol. 44, no. 12, pp. 3085–3098, Dec. 1996.
- [32] M. Wang and A. Nehorai, "Coarrays, MUSIC, and the Cramér Rao bound," *IEEE Transactions on Signal Processing*, vol. 65, no. 4, pp. 933–946, Feb. 2017.
- [33] C.-L. Liu and P. Vaidyanathan, "Remarks on the spatial smoothing step in coarray MUSIC," *IEEE Signal Processing Letters*, vol. 22, no. 9, Sep. 2015.
- [34] C.-L. Liu and P. P. Vaidyanathan, "CramrRao bounds for coprime and other sparse arrays, which find more sources than sensors," *Digital Signal Processing*, vol. 61, pp. 43 – 61, 2017.
- [35] A. Koochakzadeh and P. Pal, "Cramér Rao bounds for underdetermined source localization," *IEEE Signal Process. Lett.*, vol. 23, no. 7, pp. 919– 923, Jul. 2016.
- [36] P. Stoica, E. G. Larsson, and A. B. Gershman, "The stochastic CRB for array processing: a textbook derivation," *IEEE Signal Processing Letters*, vol. 8, no. 5, pp. 148–150, May 2001.

- [37] M. Wang, Z. Zhang, and A. Nehorai, "Performance analysis of coarraybased MUSIC and the Cramér-Rao bound," in 2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Mar. 2017, pp. 3061–3065.
- [38] P. P. Vaidyanathan and P. Pal, "Why does direct-MUSIC on sparsearrays work?" in 2013 Asilomar Conference on Signals, Systems and Computers, Nov 2013, pp. 2007–2011.
- [39] G. A. F. Seber, A Matrix Handbook for Statisticians. Hoboken, N.J.: Wiley-Interscience, 2008, oCLC: 191879555.



**Mianzhi Wang** (S'15) received the B.Sc. degree from Fudan University, Shanghai, China, in 2013. He received the M.Sc. and Ph.D. degrees from Washington University in St. Louis, St. Louis, MO, in 2018, under the guidance of Dr. A. Nehorai. His research interests are mainly in the areas of statistical signal processing for sensor arrays, optimization, and machine learning.



**Zhen Zhang** Zhen Zhang (S'17) received the B.Sc. degree from University of Science and Technology of China, in 2014. He is a Ph.D. candidate in the Preston M. Green Department of Electrical and Systems Engineering, Washington University in St. Louis, St. Louis, MO, USA under the guidance of Dr. A. Nehorai. His research interests include the areas of machine learning and computer vision.



**Arye Nehorai** Arye Nehorai (S'80, M'83, SM'90, F'94, LF'17) is the Eugene and Martha Lohman Professor of Electrical Engineering in the Preston M. Green Department of Electrical and Systems Engineering (ESE) at Washington University in St. Louis (WUSTL). He served as chair of this department from 2006 to 2016. Under his leadership, the undergraduate enrollment has more than tripled and the masters enrollment has grown seven-fold. He is also Professor in the Division of Biology and Biomedical Sciences (DBBS), the Division of

Biostatistics, the Department of Biomedical Engineering, and Department of Computer Science and Engineering, and Director of the Center for Sensor Signal and Information Processing at WUSTL. Prior to serving at WUSTL, he was a faculty member at Yale University and the University of Illinois at Chicago. He received the B.Sc. and M.Sc. degrees from the Technion, Israel and the Ph.D. from Stanford University, California.

Dr. Nehorai served as Editor-in-Chief of the IEEE Transactions on Signal Processing from 2000 to 2002. From 2003 to 2005 he was the Vice President (Publications) of the IEEE Signal Processing Society (SPS), the Chair of the Publications Board, and a member of the Executive Committee of this Society. He was the founding editor of the special columns on Leadership Reflections in IEEE Signal Processing Magazine from 2003 to 2006.

Dr. Nehorai received the 2006 IEEE SPS Technical Achievement Award and the 2010 IEEE SPS Meritorious Service Award. He was elected Distinguished Lecturer of the IEEE SPS for a term lasting from 2004 to 2005. He received several best paper awards in IEEE journals and conferences. In 2001 he was named University Scholar of the University of Illinois. Dr. Nehorai was the Principal Investigator of the Multidisciplinary University Research Initiative (MURI) project titled Adaptive Waveform Diversity for Full Spectral Dominance from 2005 to 2010. He is a Fellow of the Royal Statistical Society since 1996, and Fellow of AAAS since 2012, and Life Fellow of the IEEE since 2017.