On the Performance of One-Bit DoA Estimation via Sparse Linear Arrays

Saeid Sedighi, *Student Member, IEEE,* M. R. Bhavani Shankar, *Senior Member, IEEE,* Mojtaba Soltanalian, *Senior Member, IEEE,* and Björn Ottersten, *Fellow, IEEE*

Abstract—Direction of Arrival (DoA) estimation using Sparse Linear Arrays (SLAs) has recently gained considerable attention in array processing thanks to their capability to provide enhanced degrees of freedom in resolving uncorrelated source signals. Additionally, deployment of one-bit Analog-to-Digital Converters (ADCs) has emerged as an important topic in array processing, as it offers both a low-cost and a low-complexity implementation. In this paper, we study the problem of DoA estimation from one-bit measurements received by an SLA. Specifically, we first investigate the identifiability conditions for the DoA estimation problem from one-bit SLA data and establish an equivalency with the case when DoAs are estimated from infinite-bit unquantized measurements. Towards determining the performance limits of DoA estimation from one-bit quantized data, we derive a pessimistic approximation of the corresponding Cramér-Rao Bound (CRB). This pessimistic CRB is then used as a benchmark for assessing the performance of one-bit DoA estimators. We also propose a new algorithm for estimating DoAs from one-bit quantized data. We investigate the analytical performance of the proposed method through deriving a closed-form expression for the covariance matrix of the asymptotic distribution of the DoA estimation errors and show that it outperforms the existing algorithms in the literature. Numerical simulations are provided to validate the analytical derivations and corroborate the resulting performance improvement.

Index Terms—Sparse linear arrays, direction of arrival (DoA) estimation, Cramér-Rao bound (CRB), one-bit quantization

I. INTRODUCTION

The problem of Direction of Arrival (DoA) estimation is of central importance in the field of array processing with many applications in radar, sonar, and wireless communications [1–3]. Estimating DoAs using Uniform Linear Arrays (ULAs) is well-investigated in the literature; a number of algorithms such as the Maximum Likelihood (ML) estimator, MUSIC, ESPRIT and subspace fitting were presented and their performance thoroughly analyzed [4–9]. However, it is widely known that ULAs are not capable of identifying more sources than the number of physical elements in the array [2, 7].

To transcend this limitation, exploitation of Sparse Linear Arrays (SLAs) with particular geometries, such as Minimum Redundancy Arrays (MRAs) [10], co-prime arrays [11] and nested arrays [12] has been proposed. These architectures can dramatically boost the degrees of freedom of the array for uncorrelated source signals such that a significantly larger number of sources than the number of physical elements in the array can be identified. In addition, the enhanced degrees of freedom provided by these SLAs can improve the resolution performance appreciably compared to ULAs [12]. These features have spurred further research on DoA estimation using SLAs in recent years. A detailed study on DoA estimation via SLAs through an analysis of the Cramér-Rao Bound (CRB) was conducted in [13]. Further, a number of approaches to estimating DoAs from SLA measurements were proposed in the literature. In general, the proposed approaches can be classified under two main groups: 1) Sparsity-Based Methods (SBMs); 2) Augmented Covariance-Based Methods (ACBMs). SBMs estimate DoAs by imposing sparsity constraints on source profiles and exploiting the compressive sensing recovery techniques [14–20]. However, in ACBMs, DoAs are estimated by applying conventional subspace methods such as MUSIC and ESPRIT on an Augmented Sample Covariance Matrix (ASCM) developed from the original sample covariance matrix by exploiting the difference co-array structure [12, 21, 22]. In addition, the authors of this paper recently proposed a Weighted Least Squares (WLS) estimator capable of asymptotically achieving the corresponding CRB for DoA estimation from SLA data [23, 24].

The aforementioned techniques for DoA estimation from SLA data rest on the assumption that the analog array measurements are digitally represented by a significantly large number of bits per sample such that the resulting quantization errors can be disregarded. However, the production costs and energy consumption of Analog-to-Digital Converters (ADCs) escalate dramatically as the number of quantization bits and sampling rate increase [25]. In consequence, deployment of high-resolution ADCs in many modern applications, e.g. cognitive radio [26], cognitive radars [27], automotive radars [28], radio astronomy [29] and massive multiple-input multiple-output (MIMO) systems [30], is not economically viable owing to their very high bandwidth. In order to reduce energy consumption and production cost in such applications, researchers and system designers have recently proposed using low-resolution ADCs. As an extreme case of low-resolution ADCs, one-bit ADCs, which convert an analog signal into digital data using a single bit per sample, has received significant attention in the literature.

One-bit ADCs offer an extremely high sampling rate at a low cost and very low energy consumption [25]. Additionally, they enjoy the benefits of relatively easy implementation due to their simple architecture [31]. In the past few years, numerous studies were conducted to investigate the impact of using one-bit sampling on various applications such as massive MIMO systems [32–36], dictionary learning [37], radar [38–42], and array processing [43, 44].

A. Relevant Works

The problem of DoA estimation from one-bit quantized data has been studied in the literature presuming both the deterministic signal model [6] and the stochastic signal model [7]. The studies in [45–49] presuppose the deterministic signal model. The authors in [45] developed an algorithm for reconstruction of the unquantized array measurements from one-bit samples followed by MUSIC to determine DOAs. The ML estimation was deployed in [46] for finding DoAs from one-bit data. In [49], the authors utilized a sparse Bayesian learning algorithm to solve the DoA estimation problem from one-bit samples. Two sparsity-based approaches were also proposed in [47, 48]. Further, DoA estimation from one-bit data assuming the stochastic signal model has been discussed in [43, 44, 50, 51]. In the special case of a two-sensor array, the exact CRB expression for the DoA estimation problem from one-bit quantized data was derived in [43]. Moreover, an approach for estimating DoAs from one-bit ULA samples was proposed in [43] which is based on reconstruction of the covariance matrix of unquantized data using the arcsine law [52]. In contrast to the approach employed in [53] which relies on the covaraince matrix reconstruction of unquantized data, the DoA estimation was performed in [51] by directly applying MUSIC on the sample covariance matrix of one-bit ULA data. The numerical simulations demonstrated that the approach proposed in [51] performs similar to the algorithm proposed in [43] in the low Signal-to-Noise Ratio (SNR) regime. An upper bound on the CRB of estimating a single source DoA from one-bit ULA measurements was derived in [44].

The aforementioned research works considered using ULAs for one-bit DoA estimation. Exploitation of SLAs for one-bit DoA estimation has been studied in [53–56]. The authors in [53] deployed the arcsine law [52] to reconstruct the ASCM from one-bit SLA data. Then, they applied MUSIC on the reconstructed ASCM to estimate DoAs. It was shown in [53] that the performance degradation due to one-bit quantization can, to some extent, be compensated using SLAs. An array interpolation-based algorithm was employed in [56] to estimate DoAs from one-bit data received by co-prime arrays. Cross-dipoles sparse arrays were deployed in [55] to develop a method for one-bit DoA estimation which is robust against polarization states. In [54], the authors proposed an approach to jointly estimate DoAs and array calibration errors from one-bit data.

Nonetheless, the analytical performance of DoA estimation from one-bit SLA measurements has not yet been studied in the literature and performance analysis in the literature has been limited to simulations studies. Therefore, fundamental performance limitations of DoA estimation form one-bit SLA measurements have not well understood.

B. Our Contributions

It is of great importance to analytically investigate the performance of DoA estimation from one-bit SLA measurements. Such a performance analysis not only provides us with valuable insights into the performance of DoA estimation from one-bit SLA data but also enables us to compare its performance with that of DoA estimation using infinite-bit (unquantized) SLA data. Hence, as one of the contributions of this paper, we conduct a rigorous study on the performance of estimating source DoAs from one-bit SLA samples. Furthermore, we propose a new algorithm for estimating source DoAs from one-bit SLA measurements and analyze its asymptotic performance. Specifically, the contributions of this paper are described as follows:

- **Identifiability Analysis:** We study the identifiability conditions for the DoA estimation from one-bit SLA data. We first show that the identifiability condition for estimating DoAs from one-bit SLA data is equivalent to the case when DoAs are estimated from infinite-bit (unquantized) SLA data. Then, we determine a sufficient condition for global identifiablity of DoAs from one-bit data based on the relationship between the number of source and array elements.

- **CRB Derivation and Analysis:** We derive a pessimistic approximation of the CRB of DoA estimation using one-bit data received by an SLA. This pessimistic CRB approximation provides a benchmark for the performance of DoA estimation algorithms from one-bit data. Additionally, it helps us to spell out the condition under which the Fisher Information Matrix (FIM) of one-bit data is invertible, and thus, the CRB is a valid bound for one-bit DoA estimators. Further, we derive the performance limits of one-bit DoA estimation using SLAs at different conditions.

- **Novel One-bit DoA Estimator:** We propose a new MUSIC-based algorithm for estimating DoAs from one-bit SLA measurements. In this regard, we first construct an enhanced estimate of the normalized covariance matrix of infinite-bit (unquantized) data by exploiting the structure of the normalized covariance matrix efficiently. Then, we apply MUSIC to an augmented version of the enhanced normalized covariance matrix estimate to determine the DoAs.

- **Performance Analysis of the Proposed Estimator:** We derive a closed-form expression for the second-order statistics of the asymptotic distribution (for the large number of snapshots) of the proposed algorithm. Our asymptotic performance analysis shows that the proposed estimator outperform its counterparts in the literature and that its performance is very close to the proposed pessimistic approximation of the CRB. Moreover, the asymptotic performance analysis of the proposed DoA estimator enables us to provide valuable insights on its performance. For examples, we observe that the Mean Square Error (MSE) depends on both the physical array geometry and the co-array geometry. In addition, we observe that the MSE does not drop to zero even if the SNR approaches infinity.

- **Wider Applicability of the derived performance Analysis:** We provide a closed-form expression for the large sample performance of the one-bit DoA estimator in [53] as a byproduct of the performance analysis of our proposed DoA estimator.

Organization: Section II describes the system model. In Section III, the identifiability condition for DoA estimation problem from one-bit quantized data is discussed. Section IV presents the pessimistic approximation of the CRB and related discussions. In Section V, the proposed algorithm for DoA estimation from one-bit measurements is given and its performance is analyzed. The simulation results and related discussions are included in Section VI. Finally, Section VII concludes the paper.

Notation: Vectors and matrices are referred to by lower- and upper-case bold-face, respectively. The superscripts $*$, T, H denote the conjugate, transpose and Hermitian (conjugate transpose) operations, respectively. $[\mathbf{A}]_{i,j}$ and $[\mathbf{a}]_i$ indicate the $(i,j)^{\text{th}}$ and i^{th} entry of $\mathbf{A}$ and $\mathbf{a}$, respectively. $\|\mathbf{a}\|_2$ stands for the ℓ_2-norm of $\mathbf{a}$. $|\mathbb{A}|$ represents the cardinality of the set $\mathbb{A}$. $|a|$, $\lceil a \rceil$ and $\lfloor a \rfloor$ represent the absolute value of, the least integer greater than or equal to and greatest integer less than or equal to the scalar a, respectively. $\text{diag}(\mathbf{a})$ and $\text{diag}(\mathbf{A})$ are diagonal matrices whose diagonal entries are equal to the elements of $\mathbf{a}$ and to the diagonal elements of $\mathbf{A}$, respectively. The $M \times M$ identity matrix is denoted by $\mathbf{I}_M$. $\text{sgn}(x)$ denotes the sign function with $\text{sgn}(x) = 1$ for $x \geq 0$ and $\text{sgn}(x) = -1$ otherwise. The real and image part of a are denoted by $\Re\{a\}$ and $\Im\{a\}$, respectively. $\mathbb{E}\{.\}$ stands for the statistical expectation. $\otimes$ and $\odot$ represent Kronecker and Khatri-Rao products, respectively. $\text{tr}(\mathbf{A})$, $\det(\mathbf{A})$ and $\text{rank}(\mathbf{A})$ denote the trace, determinant and rank, respectively. $\text{vec}(\mathbf{A}) = \begin{bmatrix} \mathbf{a}_1^T & \mathbf{a}_2^T & \cdots & \mathbf{a}_n^T \end{bmatrix}^T$ represents the vectorization operation and $\text{mat}_{m,n}(.)$ is its inverse operation. $\mathbf{A}^\dagger$ and $\Pi_{\mathbf{A}}^\perp$ indicate the pseudoinverse and the projection matrix onto the null space of the full column rank matrix $\mathbf{A}^H$, respectively. $\mathcal{CN}(\mathbf{a}, \mathbf{A})$ denote the circular complex Gaussian distribution with mean $\mathbf{a}$ and covariance matrix $\mathbf{A}$.

II. SYSTEM MODEL

We consider an SLA with M elements located at positions $m_1\frac{\lambda}{2}, m_2\frac{\lambda}{2}, \cdots, m_M\frac{\lambda}{2}$ with $m_i \in \mathbb{M}$. Here $\mathbb{M}$ is a set of integers with cardinality $|\mathbb{M}| = M$, and λ denotes the wavelength of the incoming signals. It is assumed that K narrowband signals with distinct DoAs $\boldsymbol{\theta} = [\theta_1, \theta_2, \cdots, \theta_K]^T \in [-\pi/2, \pi/2]^{K \times 1}$ impinge on the SLA from far field. The signal received at the array at time instance t can be modeled as

$$\mathbf{y}(t) = \mathbf{A}(\boldsymbol{\theta})\mathbf{s}(t) + \mathbf{n}(t) \in \mathbb{C}^{M \times 1}, \quad t = 0, \cdots, N-1, \quad (1)$$

where $\mathbf{s}(t) \in \mathbb{C}^{K \times 1}$ denotes the vector of source signals, $\mathbf{n}(t) \in \mathbb{C}^{M \times 1}$ is additive noise, and $\mathbf{A}(\boldsymbol{\theta}) = [\mathbf{a}(\theta_1), \mathbf{a}(\theta_2), \cdots, \mathbf{a}(\theta_K)] \in \mathbb{C}^{M \times K}$ represents the SLA steering matrix with

$$\mathbf{a}(\theta_k) = [e^{j\pi \sin\theta_k m_1}, e^{j\pi \sin\theta_k m_2}, \cdots, e^{j\pi \sin\theta_k m_M}]^T, \quad (2)$$

being the SLA manifold vector for the ith signal. Further, the following assumptions are made on source signals and noise:

A1 $\mathbf{n}(t)$ follows a zero-mean circular complex Gaussian distribution with the covariance matrix $\mathbb{E}\{\mathbf{n}(t)\mathbf{n}^H(t)\} = \sigma^2 \mathbf{I}_M$.
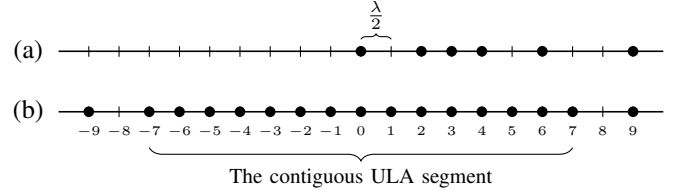

Figure 1. Array geometry of a co-prime array with $M = 6$ elements: (a) physical array with $\mathbb{M} = \{0, 2, 3, 4, 6, 9\}$; (b) difference co-array with $\mathbb{D} = \{0, 1, 2, 3, 4, 5, 6, 7, 9\}$ and $v = 8$.

A2 The source signals are modeled as zero-mean *uncorrelated* circular complex Gaussian random variables with covariance matrix $\mathbb{E}\{\mathbf{s}(t)\mathbf{s}^H(t)\} = \text{diag}(\mathbf{p})$ where $\mathbf{p} = [p_1, p_2, \cdots, p_K]^T \in \mathbb{R}_{>0}^{K \times 1}$ (i.e., $p_k > 0, \forall k$).

A3 Source and noise vectors are mutually independent.

A4 There is no temporal correlation between the snapshots, i.e., $\mathbb{E}\{\mathbf{n}(t_1)\mathbf{n}^H(t_2)\} = \mathbb{E}\{\mathbf{s}(t_1)\mathbf{s}^H(t_2)\} = \mathbf{0}$ if $t_1 \neq t_2$.

A5 An exact knowledge of the number of sources is available.

Given **A1** - **A4**, the covariance matrix of $\mathbf{y}(t)$ is expressed as

$$\mathbf{R} = \mathbf{A}(\boldsymbol{\theta})\text{diag}(\mathbf{p})\mathbf{A}^H(\boldsymbol{\theta}) + \sigma^2 \mathbf{I}_M \in \mathbb{C}^{M \times M}. \quad (3)$$

Vectorizing $\overline{\mathbf{R}}$ leads to [13, 21, 23]

$$\mathbf{r} \doteq \text{vec}(\mathbf{R}) = (\mathbf{A}^*(\boldsymbol{\theta}) \odot \mathbf{A}(\boldsymbol{\theta}))\mathbf{p} + \sigma^2 \text{vec}(\mathbf{I}_M),$$
$$= \mathbf{J}\mathbf{A}_d(\boldsymbol{\theta})\mathbf{p} + \sigma^2 \mathbf{J}\mathbf{e} \in \mathbb{C}^{M^2 \times 1}, \quad (4)$$

where $\mathbf{A}_d(\boldsymbol{\theta}) \in \mathbb{C}^{(2D-1) \times K}$ corresponds to the steering matrix of the difference co-array of the SLA whose elements are located at $(-\ell_{D-1}\frac{\lambda}{2}, \cdots, 0, \cdots, \ell_{D-1}\frac{\lambda}{2})$ with $\ell_i \in \mathbb{D} = \{|m_p - m_q| : m_p, m_q \in \mathbb{M}\}$ and $D = |\mathbb{D}|$. Moreover, $\mathbf{e} \in \{0, 1\}^{(2D-1) \times 1}$ is a column vector with $[\mathbf{e}]_i = \delta[i - D]$, and the selection matrix $\mathbf{J} \in \{0, 1\}^{M^2 \times (2D-1)}$ is represented as follows [13]:

$$\mathbf{J} = \begin{bmatrix} \text{vec}(\mathbf{L}_{D-1}^T), & \cdots, & \text{vec}(\mathbf{L}_0), & \cdots, & \text{vec}(\mathbf{L}_{D-1}) \end{bmatrix}, \quad (5)$$

where $[\mathbf{L}_n]_{p,q} = \begin{cases} 1, & \text{if } m_p - m_q = \ell_n, \\ 0, & \text{otherwise,} \end{cases}$ with $1 \leq p, q \leq M$ and $0 \leq n \leq D - 1$. The steering matrix of the difference co-array includes a contiguous ULA segment around the origin with the size of $2v - 1$ where v is the largest integer such that $\{0, 1, \cdots, v-1\} \subseteq \mathbb{D}$. The size of the contiguous ULA segment of the difference co-array plays a crucial role in the number of identifiable sources such that K distinct sources are identifiable if $K \leq v - 1$. Hence, in case the SLA is designed properly such that $v > M$, we are able to identify more sources than the number of physical elements in the SLA; exploiting the resulting structure of $\mathbf{R}$ efficiently [11–13, 23]. An illustrative example of an SLA, the corresponding difference co-array, and its contiguous ULA segment is presented in Fig. 1.

Here it is assumed that each array element is connected to a one-bit ADC which directly converts the received analog signal into binary data by comparing the real and imaginary parts of the received signal individually with zero. In such a case, the one-bit measurements at the mth array element are given by

$$[\mathbf{x}(t)]_m = \frac{1}{\sqrt{2}}\text{sgn}(\Re\{[\mathbf{y}(t)]_m\}) + \frac{j}{\sqrt{2}}\text{sgn}(\Im\{[\mathbf{y}(t)]_m\}). \quad (6)$$

The problem under consideration is the estimation of source DoAs, i.e., $\boldsymbol{\theta}$, from one-bit quantized measurements, i.e., $\mathbf{X} = [\mathbf{x}(0), \mathbf{x}(1), \cdots, \mathbf{x}(N-1)]$, collected by the SLA.

III. IDENTIFIABILITY CONDITIONS

Note that there is a significant information loss expected when going from infinite-bit (unquantized) data, i.e., $\mathbf{Y} = [\mathbf{y}(0), \mathbf{y}(1), \cdots, \mathbf{y}(M)]$, to one-bit data, i.e., $\mathbf{X}$. This information loss may affect the attractive capability of SLAs to identify a larger number of uncorrelated sources than the number of array elements. To address this concern, we will consider the identifiability conditions for DoA estimation from one-bit SLA measurements in this section. Before proceeding further, we first need to give a clear definition of identifiability for this problem.

Definition 1 (Identifiability). Let $f(\mathbf{X} \,|\, \boldsymbol{\theta}, \mathbf{p}, \sigma^2)$ denote the Probability Density Function (PDF) of $\mathbf{X}$ parameterized by $\boldsymbol{\theta}, \mathbf{p}$ and σ^2. Then, the source DoAs are said to be identifiable from $\mathbf{X}$ at point $\boldsymbol{\theta}_0 \in [-\pi/2, \pi/2]^{K \times 1}$ if there exist no $\breve{\boldsymbol{\theta}} \neq \boldsymbol{\theta}_0 \in [-\pi/2, \pi/2]^{K \times 1}$ such that $f(\mathbf{X} \,|\, \boldsymbol{\theta}_0, \mathbf{p}, \sigma^2) = f(\mathbf{X} \,|\, \breve{\boldsymbol{\theta}}, \breve{\mathbf{p}}, \breve{\sigma}^2)$ for any arbitrary values of $\mathbf{p} \in \mathbb{R}_{>0}^{K \times 1}$, $\breve{\mathbf{p}} \in \mathbb{R}_{>0}^{K \times 1}$, σ^2 and $\breve{\sigma}^2$ [57, Ch. 1, Definition 5.2] [58, pp. 62].

Remark 1. The above definition can be used for identifiabilty of $\boldsymbol{\theta}_0$ from $\mathbf{Y}$ by replacing $f(\mathbf{X} \,|\, \boldsymbol{\theta}, \mathbf{p}, \sigma^2)$ with $f(\mathbf{Y} \,|\, \boldsymbol{\theta}, \mathbf{p}, \sigma^2)$.

Based on the above definition, the necessary and sufficient condition for a particular DoA point to be identifiable from one-bit SLA data is given in the following Theorem.

Theorem 1. *The source DoAs are identifiable from $\mathbf{X}$ at $\boldsymbol{\theta}_0 \in [-\pi/2, \pi/2]^{K \times 1}$ if and only if they are identifiable from $\mathbf{Y}$ at $\boldsymbol{\theta}_0$.*

Proof. See Appendix A. □

The above Theorem shows that the identifiability condition for the DoA estimation problem from one-bit SLA measurements is equivalent to that for the DoA estimation problem from infinite-bit (unquantized) SLA measurements. Hence, the information loss arises from one-bit quantization does not influence the number of identifiable sources. However, Theorem 1 simply spells out the identifiability condition of a single DoA point. A sufficient condition for global identifiablity of source DoAs from one-bit data is given in the following theorem.

Definition 2 (Global identifiability). The source DoAs are said to be globally identifiable from $\mathbf{X}$ if there exists no distinct $\boldsymbol{\theta} \in [-\pi/2, \pi/2]^{K \times 1}$ and $\breve{\boldsymbol{\theta}} \in [-\pi/2, \pi/2]^{K \times 1}$ such that $f(\mathbf{X} \,|\, \boldsymbol{\theta}, \mathbf{p}, \sigma^2) = f(\mathbf{X} \,|\, \breve{\boldsymbol{\theta}}, \breve{\mathbf{p}}, \breve{\sigma}^2)$ for any arbitrary values of $\mathbf{p} \in \mathbb{R}_{>0}^{K \times 1}$, $\breve{\mathbf{p}} \in \mathbb{R}_{>0}^{K \times 1}$, σ^2 and $\breve{\sigma}^2$.

Theorem 2. *The sufficient conditions for global indentifiability and global non-indentifiability of source DoAs from one-bit SLA data are given as follows:*

S1 *The source DoAs are globally identifiable (with probability one) from $\mathbf{X}$ for any value of $\boldsymbol{\theta} \in [-\pi/2, \pi/2]^{K \times 1}$ if $K \leq v - 1$.*

S2 *The source DoAs are globally unidentifiable from $\mathbf{X}$ for any value of $\boldsymbol{\theta} \in [-\pi/2, \pi/2]^{K \times 1}$ if $K \geq D$.*

Proof. See Appendix B. □

Having revealed that one-bit quantization does not affect the indentifiability conditions of source DoAs, we will investigate

the performance of DoA estimation from one-bit SLA data through a CRB analysis in the next section.

IV. CRAMÉR-RAO BOUND ANALYSIS

It is well-known that the CRB offers a lower bound on the covariance of any unbiased estimator [59]. Hence, it is considered as a standard metric for evaluating the performance of estimators. In particular, the CRB can provide valuable insights into the fundamental limits of estimation for specific problems as well as the dependence of the estimation performance on various system parameters. Deriving a closed-form expression for the CRB requires knowledge of the data distribution. However, the data distribution may not be known for some problems. In such cases, the Gaussian assumption is a natural choice which leads to the largest (most pessimistic) CRB in a general class of data distributions [60].

In the problem of DoA estimation from one-bit SLA measurements, the true PDF of one-bit data is obtained from the orthant probabilities [61] of Gaussian distribution, for which a closed-form expression is not available in general. Motivated by this fact, in what follows, we derive a pessimistic closed-form approximation for the CRB of the DoA estimation problem from one-bit SLA data through considering a Gaussian distribution for $\mathbf{x}(t)$. This pessimistic closed-form approximation is used for benchmarking the performance of one-bit DoA estimators as well as for investigating the performance limits of the DoA estimation problem from one-bit data. Making use of assumptions **A1**-**A4**, it is readily confirmed that $\mathbb{E}\{\mathbf{x}(t)\} = \mathbf{0}$. Further, the arcsine law [52] establishes the following relationship between $\mathbf{R}$ and $\mathbf{R_x}$:

$$\mathbf{R_x} = \mathbb{E}\{\mathbf{x}(t)\mathbf{x}^H(t)\} = \frac{2}{\pi}\mathrm{arcsine}(\overline{\mathbf{R}}), \tag{7}$$

where $[\mathrm{arcsine}(\overline{\mathbf{R}})]_{m,n} = \arcsin(\Re\{[\overline{\mathbf{R}}]_{m,n}\}) + j\arcsin(\Im\{[\overline{\mathbf{R}}]_{m,n}\})$ and

$$\overline{\mathbf{R}} = \frac{\mathbf{R}}{\sigma^2 + \sum_{k=1}^{K} p_k} = \mathbf{A}(\boldsymbol{\theta})\mathrm{diag}(\overline{\mathbf{p}})\mathbf{A}^H(\boldsymbol{\theta}) + (1 - \sum_{k=1}^{K} \overline{p}_k)\mathbf{I}_M, \tag{8}$$

is the normalized covariance matrix of $\mathbf{y}(t)$ with $\overline{\mathbf{p}} = [\overline{p}_1, \overline{p}_2, \cdots, \overline{p}_K]^T$ and $\overline{p}_k = \frac{p_k}{\sigma^2 + \sum_{k=1}^{K} p_k}$. It follows from (7) and (8) that $\mathbf{R_x}$ is a function of the parameters $\boldsymbol{\theta}$ and $\overline{\mathbf{p}}$. Let $\boldsymbol{\varrho} = [\boldsymbol{\theta}, \overline{\mathbf{p}}]^T$ denote the vector of unknown parameters. Then, considering the Gaussian assumption, the worst-case Fisher Information Matrix (FIM) $\mathcal{I}_w(\boldsymbol{\varrho})$ is given by [59]

$$[\mathcal{I}_w(\boldsymbol{\varrho})]_{m,n} = N\mathrm{tr}(\mathbf{R_x}^{-1}\frac{\partial \mathbf{R_x}}{\partial[\boldsymbol{\varrho}]_m}\mathbf{R_x}^{-1}\frac{\partial \mathbf{R_x}}{\partial[\boldsymbol{\varrho}]_n})$$
$$= N\frac{\partial \mathbf{r_x}^H}{\partial[\boldsymbol{\varrho}]_m}(\mathbf{R_x}^{-T} \otimes \mathbf{R_x}^{-1})\frac{\partial \mathbf{r_x}}{\partial[\boldsymbol{\varrho}]_n}, \tag{9}$$

where $\mathbf{r_x} = \mathrm{vec}(\mathbf{R_x})$ and the last equality is obtained by using the relation $\mathrm{tr}(\mathbf{C}_1\mathbf{C}_2\mathbf{C}_3\mathbf{C}_4) = \mathrm{vec}^H(\mathbf{C}_2^H)(\mathbf{C}_1^T \otimes \mathbf{C}_3)\mathrm{vec}(\mathbf{C}_4)$. From (4), (7) and (9), we obtain

$$\mathbf{r_x} = \frac{2}{\pi}\mathrm{arcsine}(\mathrm{vec}(\overline{\mathbf{R}})) = \frac{2}{\pi}\mathbf{J}\mathrm{arcsine}\left(\mathbf{A}_d(\boldsymbol{\theta})\overline{\mathbf{p}} + (1 - \sum_{k=1}^{K} \overline{p}_k)\mathbf{e}\right). \tag{10}$$

Computing the derivative of $\mathbf{r_x}$ with respect to θ_k and $\overline{p}_k$ yields

$$\frac{\partial \mathbf{r_x}}{\partial \theta_k} = j\pi\cos(\theta_k)\overline{p}_k\mathbf{J}\mathrm{diag}(\mathbf{d}) \tag{11}$$
$$\times [\mathrm{diag}(\overline{\mathbf{h}})\Re\{\mathbf{a}_d(\theta_k)\} + j\mathrm{diag}(\mathbf{h})\Im\{\mathbf{a}_d(\theta_k)\}],$$

$$\frac{\partial \mathbf{r_x}}{\partial \overline{p}_k} = \mathbf{J}[\mathrm{diag}(\mathbf{h})\Re\{\mathbf{a}_d(\theta_k)\} + j\mathrm{diag}(\overline{\mathbf{h}})\Im\{\mathbf{a}_d(\theta_k)\}], \tag{12}$$

$$\mathbf{h} = \left[\frac{1}{\sqrt{1-|\Re\{\sum_{k=1}^{K} \overline{p}_k e^{-j\pi\sin\theta_k \ell_{D-1}}\}|^2}} \quad \cdots \quad 0 \quad \cdots \quad \frac{1}{\sqrt{1-|\Re\{\sum_{k=1}^{K} \overline{p}_k e^{j\pi\sin\theta_k \ell_{D-1}}\}|^2}} \right]^T, \tag{13}$$

$$\overline{\mathbf{h}} = \left[\frac{1}{\sqrt{1-|\Im\{\sum_{k=1}^{K} \overline{p}_k e^{-j\pi\sin\theta_k \ell_{D-1}}\}|^2}} \quad \cdots \quad 0 \quad \cdots \quad \frac{1}{\sqrt{1-|\Im\{\sum_{k=1}^{K} \overline{p}_k e^{j\pi\sin\theta_k \ell_{D-1}}\}|^2}} \right]^T, \tag{14}$$

where $\mathbf{h}$ and $\overline{\mathbf{h}}$ are given in (13) and (14) at the top of the next page, $\mathbf{a}_d(\theta_k)$ denotes the kth column of $\mathbf{A}_d(\boldsymbol{\theta})$ and $\mathbf{d} = [-\ell_{D-1}, \cdots, \ell_0, \cdots, \ell_{D-1}]^T$. It follows from (9), (11) and (12) that

$$\mathcal{I}_w(\boldsymbol{\varrho}) = N \begin{bmatrix} \mathbf{G}^H \\ \mathbf{V}^H \end{bmatrix} \mathbf{J}^H (\mathbf{R}_{\mathbf{x}}^{-T} \otimes \mathbf{R}_{\mathbf{x}}^{-1}) \mathbf{J} \begin{bmatrix} \mathbf{G} & \mathbf{V} \end{bmatrix}, \tag{15}$$

where

$$\mathbf{G} = j\pi \text{diag}(\mathbf{d}) \big[\text{diag}(\overline{\mathbf{h}}) \Re\{\mathbf{A}_d(\boldsymbol{\theta})\} \tag{16}$$
$$+ j\text{diag}(\mathbf{h}) \Im\{\mathbf{A}_d(\boldsymbol{\theta})\} \big] \boldsymbol{\Phi}(\boldsymbol{\theta}) \text{diag}(\overline{\mathbf{p}}),$$

$$\mathbf{V} = \text{diag}(\mathbf{h}) \Re\{\mathbf{A}_d(\boldsymbol{\theta})\} + j\text{diag}(\overline{\mathbf{h}}) \Im\{\mathbf{A}_d(\boldsymbol{\theta})\}, \tag{17}$$

with $\boldsymbol{\Phi}(\boldsymbol{\theta}) = \text{diag}([\cos\theta_1, \cos\theta_2, \cdots, \cos\theta_K]^T)$. If $\mathcal{I}_w(\boldsymbol{\varrho})$ is non-singular, a pessimistic approximation for the CRB of estimating DoAs from one-bit SLA data can be obtained through inverting $\mathcal{I}_w(\boldsymbol{\varrho})$. Hence, we need to first establish the non-singularity of $\mathcal{I}_w(\boldsymbol{\varrho})$.

Lemma 1. *Define* $\boldsymbol{\Upsilon} = \begin{bmatrix} \boldsymbol{\Delta} & F \end{bmatrix} \in \mathbb{C}^{(2D-1) \times 2K}$, *where*

$$\boldsymbol{\Delta} = \text{diag}(\mathbf{d}) \big[\text{diag}(\overline{\mathbf{h}}) \Re\{\mathbf{A}_d(\boldsymbol{\theta})\} + j\text{diag}(\mathbf{h}) \Im\{\mathbf{A}_d(\boldsymbol{\theta})\} \big], \tag{18}$$

$$F = \text{diag}(\mathbf{h}) \Re\{\mathbf{A}_d(\boldsymbol{\theta})\} + j\text{diag}(\overline{\mathbf{h}}) \Im\{\mathbf{A}_d(\boldsymbol{\theta})\}. \tag{19}$$

Then, $\mathcal{I}_w(\boldsymbol{\varrho})$ *is non-singular if and only if* $\boldsymbol{\Upsilon}$ *is full-column rank.*

Proof. See Appendix C $\qquad \square$

Remark 2. Assuming $\mathcal{I}(\boldsymbol{\varrho})$ to be the true FIM, it follows from $\mathcal{I}(\boldsymbol{\varrho}) \succeq \mathcal{I}_w(\boldsymbol{\varrho})$ that $\boldsymbol{\Upsilon}$ being full-column rank is also a sufficient condition for the non-sigularity of $\mathcal{I}(\boldsymbol{\varrho})$.

Theorem 3. *Let* $CRB(\boldsymbol{\theta})$ *denote the CRB for source DoAs* $\boldsymbol{\theta}$ *from* $\mathbf{X}$. *If* $\mathcal{I}_w(\boldsymbol{\varrho})$ *is non-singular, then a pessimistic approximation of* $CRB(\boldsymbol{\theta})$, *denoted by* $CRB_w(\boldsymbol{\theta})$, *is given by*

$$CRB(\boldsymbol{\theta}) \preceq CRB_w(\boldsymbol{\theta}) = \frac{1}{4N\pi^2} (\mathbf{Q}^H \Pi_{\mathbf{M}^{\frac{1}{2}}\mathbf{V}}^{\perp} \mathbf{Q})^{-1}, \tag{20}$$

where $\boldsymbol{\Omega} = \frac{1}{\pi} \mathbf{G}$,

$$\mathbf{M} = \mathbf{J}^H \left(\text{arcsine}(\overline{\mathbf{R}}^T) \otimes \text{arcsine}(\overline{\mathbf{R}}) \right)^{-1} \mathbf{J}, \tag{21}$$

$$\mathbf{Q} = \mathbf{M}^{\frac{1}{2}} \text{diag}(\mathbf{d}) \boldsymbol{\Omega} \boldsymbol{\Phi}(\boldsymbol{\theta}) \text{diag}(\overline{\mathbf{p}}), \tag{22}$$

with $\mathbf{G}$ *and* $\mathbf{V}$ *being given in* (16) *and* (17), *respectively.*

Proof. See Appendix D $\qquad \square$

Remark 3. We note that $CRB_w(\boldsymbol{\theta})$ bears a superficial resemblance to the CRB expression for DoA estimation from unquantized data, given by [13, Theorem 2]

$$CRB_I(\boldsymbol{\theta}) = \frac{1}{4N\pi^2} (\widetilde{\mathbf{Q}}^H \Pi_{\widetilde{\mathbf{M}}^{\frac{1}{2}}\widetilde{\mathbf{V}}}^{\perp} \widetilde{\mathbf{Q}})^{-1}, \tag{23}$$

where

$$\widetilde{\mathbf{M}} = \mathbf{J}^H \left(\overline{\mathbf{R}}^T \otimes \overline{\mathbf{R}} \right)^{-1} \mathbf{J}, \tag{24}$$

$$\widetilde{\mathbf{Q}} = \widetilde{\mathbf{M}}^{\frac{1}{2}} \text{diag}(\mathbf{d}) \mathbf{A}_d(\boldsymbol{\theta}) \boldsymbol{\Phi}(\boldsymbol{\theta}) \text{diag}(\overline{\mathbf{p}}), \tag{25}$$

$$\widetilde{\mathbf{V}} = \begin{bmatrix} \mathbf{A}_d(\boldsymbol{\theta}) & \mathbf{e} \end{bmatrix}. \tag{26}$$

Theorem 4. *Assume all sources have equal power* p *and* $SNR = p/\sigma^2$. *Then, we have*

$$\lim_{SNR \to \infty} CRB_w(\boldsymbol{\theta}) \succ \mathbf{0}. \tag{27}$$

Proof. See Appendix E $\qquad \square$

Remark 4. Theorem 4 implies that the $CRB_w(\boldsymbol{\theta})$ does not go to zero as the SNR increases. As a consequence, in the one-bit DoA estimation problem, we may not be able to render estimation errors arbitrarily small by increasing the SNR.

V. PROPOSED ONE-BIT DOA ESTIMATOR

In this section, we first derive an enhanced estimate of the normalized covariance matrix of $\mathbf{y}(t)$, i.e., $\overline{\mathbf{R}}$, from one-bit SLA measurements through exploiting the structure of $\overline{\mathbf{R}}$. Then, we obtain DoA estimates by applying Co-Array-Based MUSIC (CAB-MUSIC) [21, 62] to the enhanced estimate of $\overline{\mathbf{R}}$. Further, we investigate the analytical performance of the proposed method for estimating DoAs from one-bit measurements.

A. Enhanced One-Bit Co-Array-Based MUSIC

It is deduced from the strong law of large numbers [63, ch. 8] that the sample covariance matrix of one-bit data provides a consistent estimate of $\mathbf{R}_{\mathbf{x}}$ with probability 1, i.e., $\Pr\left(\lim_{N\to\infty} \widehat{\mathbf{R}}_{\mathbf{x}} = \mathbf{R}_{\mathbf{x}}\right) = 1$, where $\widehat{\mathbf{R}}_{\mathbf{x}} = \frac{1}{N} \mathbf{X} \mathbf{X}^H$. In addition, reformulating (7) gives $\overline{\mathbf{R}}$ based on the covariance matrix of one-bit data as follows:

$$\overline{\mathbf{R}} = \text{sine}(\frac{\pi}{2} \mathbf{R}_{\mathbf{x}}), \tag{28}$$

where $[\text{sine}(\frac{\pi}{2}\mathbf{R}_{\mathbf{x}})]_{m,n} = \sin(\frac{\pi}{2}\Re\{[\overline{\mathbf{R}}]_{m,n}\}) + j\sin(\frac{\pi}{2}\Im\{[\overline{\mathbf{R}}]_{m,n}\})$. Accordingly, a consistent estimate of $\overline{\mathbf{R}}$ is obtained as

$$\widetilde{\overline{\mathbf{R}}} = \text{sine}(\frac{\pi}{2} \widehat{\mathbf{R}}_{\mathbf{x}}). \tag{29}$$

Most of the algorithms in the literature employ $\widetilde{\overline{\mathbf{R}}}$ for estimating DoAs from one-bit measurements [43, 53]. However, an enhanced estimate of $\overline{\mathbf{R}}$ compared to $\widetilde{\overline{\mathbf{R}}}$ can be found if the structure of $\overline{\mathbf{R}}$ is taken into account. This enhanced estimate could in turn yield a better DoA estimation performance. In what follows, we introduce such an enhanced estimate of $\overline{\mathbf{R}}$ by exploiting its structure. Then, we use this enhanced estimate to improve the DoA estimation performance from one-bit data.

It is readily known from (8) that $\overline{\mathbf{R}}$ has the following structure

$$\overline{\mathbf{R}} = \mathbf{I}_M + \sum_{n=1}^{D-1} u_n \mathbf{L}_n + \sum_{n=1}^{D-1} u_n^* \mathbf{L}_n^T, \tag{30}$$

where $u_n = \sum_{k=1}^{K} \overline{p}_k e^{j\pi\sin\theta_k \ell_n}$ and $\mathbf{L}_n$ is given after eq. (5) for $1 \leq n \leq D-1$. It can be observed from (30) that the diagonal elements of $\overline{\mathbf{R}}$ are all one while the off-diagonal elements are parameterized by the vector $\mathbf{u} = [u_1, \cdots, u_{D-1}]^T \in \mathbb{C}^{(D-1) \times 1}$. This means that there exist

$$f(\widetilde{\ddot{\mathbf{r}}} \mid \boldsymbol{\phi}) = \left(\frac{N^{M^2-M}}{(2\pi)^{M^2-M}\det(\boldsymbol{\Sigma}(\boldsymbol{\phi}))}\right)\frac{\exp\{-N[\arcsin(\widetilde{\ddot{\mathbf{r}}}) - \overline{\mathbf{J}}\boldsymbol{\Psi}\arcsin(\boldsymbol{\phi})]^H\boldsymbol{\Sigma}^{-1}(\boldsymbol{\phi})[\arcsin(\widetilde{\ddot{\mathbf{r}}}) - \overline{\mathbf{J}}\boldsymbol{\Psi}\arcsin(\boldsymbol{\phi})]\}}{\Pi_{n=1}^{D-1}(1-[\boldsymbol{\phi}]_n^2)^{\nu_n}(1-[\boldsymbol{\phi}]_{n+D-1}^2)^{\nu_n}}. \tag{33}$$

$$L(\boldsymbol{\phi}) = \ln\det(\boldsymbol{\Sigma}(\boldsymbol{\phi})) - \sum_{n=1}^{D-1}\nu_n\ln(1-[\boldsymbol{\phi}]_n^2)(1-[\boldsymbol{\phi}]_{n+D-1}^2) + N[\arcsin(\widetilde{\ddot{\mathbf{r}}}) - \overline{\mathbf{J}}\boldsymbol{\Psi}\arcsin(\boldsymbol{\phi})]^H\boldsymbol{\Sigma}^{-1}(\boldsymbol{\phi})[\arcsin(\widetilde{\ddot{\mathbf{r}}}) - \overline{\mathbf{J}}\boldsymbol{\Psi}\arcsin(\boldsymbol{\phi})]. \tag{35}$$

only $2D-2$ free real parameters in $\overline{\mathbf{R}}$. Let $\ddot{\mathbf{r}} \in \mathbb{C}^{(M^2-M)\times 1}$ be the vector containing the off-diagonal elements of $\overline{\mathbf{R}}$, obtained by removing the diagonal elements of $\overline{\mathbf{R}}$ from $\text{vec}(\overline{\mathbf{R}})$. Evidently, $\ddot{\mathbf{r}}$ is given by

$$\ddot{\mathbf{r}} = \overline{\mathbf{J}}[\mathbf{u}^* \quad \mathbf{u}]^T = \overline{\mathbf{J}}\boldsymbol{\Psi}\boldsymbol{\phi}, \tag{31}$$

where $\boldsymbol{\phi} = [\Re\{\mathbf{u}\}^T, \Im\{\mathbf{u}\}^T]^T \in \mathbb{R}^{(2D-1\times 1)}$,

$$\boldsymbol{\Psi} = \begin{bmatrix} \mathbf{I}_{D-1} & -j\mathbf{I}_{D-1} \\ \mathbf{I}_{D-1} & j\mathbf{I}_{D-1} \end{bmatrix}. \tag{32}$$

and $\overline{\mathbf{J}} \in \{0,1\}^{(M^2-M)\times(2D-2)}$ is obtained by removing the D-th column as well as the rows with indices $(i-1)M+1$ for all $1 \leq i \leq M$ from $\mathbf{J}$. It follows from (31) that $\overline{\mathbf{R}}$ is parameterized by the real-valued vector $\boldsymbol{\phi}$. We wish to find $\boldsymbol{\phi} \in \mathbb{E}_{\boldsymbol{\phi}} = \{\boldsymbol{\phi} \mid \overline{\mathbf{R}}(\boldsymbol{\phi}) \succeq \mathbf{0}\}$ from $\widehat{\mathbf{R}}_\mathbf{x}$. To this end, let $\widehat{\ddot{\mathbf{r}}}_\mathbf{x} \in \mathbb{R}^{(M^2-M)\times 1}$ denote the vector containing the off-diagonal elements of $\widehat{\mathbf{R}}_\mathbf{x}$, obtained by removing the diagonal entries of $\widehat{\mathbf{R}}_\mathbf{x}$ from $\text{vec}(\widehat{\mathbf{R}}_\mathbf{x})$. For large N, it follows from the Central Limit Theorem (CLT) [63, ch. 8] that the distribution of $\widehat{\ddot{\mathbf{r}}}_\mathbf{x}$ asymptotically approaches a complex proper Gaussian distribution, i.e., $\widehat{\ddot{\mathbf{r}}}_\mathbf{x} \xrightarrow{D} \mathcal{CN}(\ddot{\mathbf{r}}_\mathbf{x}, \frac{4}{\pi^2 N}\boldsymbol{\Sigma})$, where $\ddot{\mathbf{r}}_\mathbf{x}$ is the vector obtained from stacking the off-diagonal elements of $\mathbf{R}_\mathbf{x}$ and $\boldsymbol{\Sigma} = \frac{\pi^2 N}{4}\mathbb{E}\{(\widehat{\ddot{\mathbf{r}}}_\mathbf{x} - \ddot{\mathbf{r}}_\mathbf{x})(\widehat{\ddot{\mathbf{r}}}_\mathbf{x} - \ddot{\mathbf{r}}_\mathbf{x})^H\} \in \mathbb{C}^{(M^2-M)\times(M^2-M)}$. The closed-form expressions for the elements of $\boldsymbol{\Sigma}$ are provided in Appendix K (kindly refer to the supplementary document). It is observed that the elements of $\boldsymbol{\Sigma}$ are functions of $\ddot{\mathbf{r}}$, thereby parameterized by $\boldsymbol{\phi}$ as well. Considering the transformation (29), the asymptotic distribution of the off-diagonal elements of $\overline{\mathbf{R}}$, denoted by $\widetilde{\ddot{\mathbf{r}}} \in \mathbb{C}^{(M^2-M)\times 1}$, is given by (33) at top of this page. Hence, the asymptotic ML estimation of $\boldsymbol{\phi}$ from $\widetilde{\ddot{\mathbf{r}}}$ is derived as follows;

$$\widehat{\boldsymbol{\phi}} = \underset{\boldsymbol{\phi} \in \mathbb{E}_{\boldsymbol{\phi}}}{\arg\min} \, L(\boldsymbol{\phi}), \tag{34}$$

where the cost function $L(\boldsymbol{\phi})$ is given in (35) at the top of the next page in which $\nu_n = \|\text{vec}(\mathbf{L}_n)\|^2$. However, the minimization of (35) with respect to $\boldsymbol{\phi}$ is very complicated owing to the nonlinearity of the cost function as well as the constraint $\boldsymbol{\phi} \in \mathbb{E}_{\boldsymbol{\phi}}$. To make the problem computationally tractable, we first find an asymptotic equivalent approximation of $L(\boldsymbol{\phi})$ which is much simpler to minimize. Let $\boldsymbol{\gamma} \in \mathbb{E}_{\boldsymbol{\gamma}} \subset \mathbb{R}^{(M^2-M)\times 1}$ be the $(M^2-M)\times 1$ vector containing the real and imaginary parts of the elements of $\overline{\mathbf{R}}$ above its main diagonal elements. Obviously, there is the following relationship between $\boldsymbol{\phi}$ and $\boldsymbol{\gamma}$:

$$\boldsymbol{\gamma} = \mathbf{F}\overline{\mathbf{J}}\boldsymbol{\Psi}\boldsymbol{\phi}, \forall \boldsymbol{\phi} \in \mathbb{E}_{\boldsymbol{\phi}}, \tag{36}$$

where $\mathbf{F} = \frac{1}{2}\left[\ddot{\mathbf{F}}^T \quad j\widetilde{\mathbf{F}}^T\right]^T \in \{0,1\}^{(M^2-M)\times(M^2-M)}$ such that for all $1 \leq p < q \leq M$:

1) the $\left((p-1)M+q-\frac{p(p+1)}{2}\right)$-th rows of $\ddot{\mathbf{F}} \in \{0,1\}^{\frac{(M^2-M)}{2}\times(M^2-M)}$ is obtained by removing the elements with indices $(i-1)M+1$ for all $1 \leq i \leq M$

from $\overline{\mathbf{e}}_p^T \otimes \overline{\mathbf{e}}_q^T + \overline{\mathbf{e}}_q^T \otimes \overline{\mathbf{e}}_p^T$ with $[\overline{\mathbf{e}}_p]_n = \delta[p-n]$ for $1 \leq n \leq M$.

2) the $\left((p-1)M+q-\frac{p(p+1)}{2}\right)$-th rows of $\widetilde{\mathbf{F}} \in \{0,1\}^{\frac{(M^2-M)}{2}\times(M^2-M)}$ is obtained by removing the elements with indices $(i-1)M+1$ for all $1 \leq i \leq M$ from $\overline{\mathbf{e}}_p^T \otimes \overline{\mathbf{e}}_q^T - \overline{\mathbf{e}}_q^T \otimes \overline{\mathbf{e}}_p^T$ with $[\overline{\mathbf{e}}_p]_n = \delta[p-n]$ for $1 \leq n \leq M$.

Lemma 2. *The matrices $\mathbf{F}$, $\boldsymbol{\Psi}$ and $\overline{\mathbf{J}}$ are full rank.*

Proof. See Appendix F. $\square$

The mapping from $\boldsymbol{\phi} \in \mathbb{E}_{\boldsymbol{\phi}}$ to $\boldsymbol{\gamma} \in \mathbb{E}_{\boldsymbol{\gamma}}$ is one-to-one due to the full rankness of $\mathbf{F}$, $\boldsymbol{\Psi}$ and $\overline{\mathbf{J}}$. Hence, it is possible to equivalently reparameterize (35) in terms of $\boldsymbol{\gamma}$ instead of $\boldsymbol{\phi}$. This can be done by simply replacing $\boldsymbol{\phi}$ with $\boldsymbol{\Psi}^{-1}\overline{\mathbf{J}}^\dagger\mathbf{F}^{-1}\boldsymbol{\gamma}$. To achieve computational simplification, we make use of the fact that a consistent estimate of $\boldsymbol{\gamma}$ can be obtained as $\widetilde{\boldsymbol{\gamma}} = \mathbf{F}\widetilde{\ddot{\mathbf{r}}}$. We see that $\widetilde{\boldsymbol{\gamma}} \in \mathbb{R}^{(M^2-M)\times 1} \notin \mathbb{E}_{\boldsymbol{\gamma}}$ with probability one, since the $\mathbb{E}_{\boldsymbol{\gamma}}$ is a zero-measure subset of $\mathbb{R}^{(M^2-M)\times 1}$. Now, considering the Taylor series expansion of $L(\boldsymbol{\gamma})$ around $\widetilde{\boldsymbol{\gamma}}$, we obtain

$$L(\boldsymbol{\gamma}) = L(\widetilde{\boldsymbol{\gamma}}) + (\boldsymbol{\gamma} - \widetilde{\boldsymbol{\gamma}})^H\nabla_{\boldsymbol{\gamma}}L(\widetilde{\boldsymbol{\gamma}}) + \frac{1}{2}(\widetilde{\boldsymbol{\gamma}} - \boldsymbol{\gamma})^H\nabla_{\boldsymbol{\gamma}}^2 L(\widetilde{\boldsymbol{\gamma}})(\widetilde{\boldsymbol{\gamma}} - \boldsymbol{\gamma}) + \cdots, \tag{37}$$

where $\nabla_{\boldsymbol{\gamma}}L(\widetilde{\boldsymbol{\gamma}})$ and $\nabla_{\boldsymbol{\gamma}}^2 L(\widetilde{\boldsymbol{\gamma}})$ denote the gradient vector and the Hessian matrix of $L(\boldsymbol{\gamma})$ with respect to $\boldsymbol{\gamma}$, computed at $\widetilde{\boldsymbol{\gamma}}$, respectively. The first term in (37) is constant and, moreover, the higher other term can be neglected for large N considering the fact that $\widetilde{\boldsymbol{\gamma}}$ is a consistent estimate of $\boldsymbol{\gamma}$. Consequently, making use of (36) and the fact that $\widetilde{\boldsymbol{\gamma}} = \mathbf{F}\widetilde{\ddot{\mathbf{r}}}$, we have

$$\widehat{\boldsymbol{\phi}} \simeq \underset{\boldsymbol{\phi} \in \mathbb{E}_{\boldsymbol{\phi}}}{\arg\min} \, (\overline{\mathbf{J}}\boldsymbol{\Psi}\boldsymbol{\phi} - \widetilde{\ddot{\mathbf{r}}})^H\mathbf{F}^H\nabla_{\boldsymbol{\gamma}}L(\widetilde{\boldsymbol{\gamma}}) + (\widetilde{\ddot{\mathbf{r}}} - \overline{\mathbf{J}}\boldsymbol{\Psi}\boldsymbol{\phi})^H\mathbf{F}^H\nabla_{\boldsymbol{\gamma}}^2 L(\widetilde{\boldsymbol{\gamma}})\mathbf{F}(\widetilde{\ddot{\mathbf{r}}} - \overline{\mathbf{J}}\boldsymbol{\Psi}\boldsymbol{\phi}). \tag{38}$$

The above quadratic optimization problem is asymptotically equivalent to (34) but is much more convenient to work with. Relaxing the constraint $\boldsymbol{\phi} \in \mathbb{E}_{\boldsymbol{\phi}}$ with $\boldsymbol{\phi} \in \mathbb{R}^{2D-2}$ yields the following closed-form solution for $\widehat{\boldsymbol{\phi}}$

$$\widehat{\boldsymbol{\phi}} \simeq \boldsymbol{\Psi}^{-1}\left(\overline{\mathbf{J}}^H\mathbf{F}^H\nabla_{\boldsymbol{\gamma}}^2 L(\widetilde{\boldsymbol{\gamma}})\mathbf{F}\overline{\mathbf{J}}\right)^{-1}\overline{\mathbf{J}}^H \times \left[\mathbf{F}^H\nabla_{\boldsymbol{\gamma}}^2 L(\widetilde{\boldsymbol{\gamma}})\mathbf{F}\widetilde{\ddot{\mathbf{r}}} - \mathbf{F}^H\nabla_{\boldsymbol{\gamma}}L(\widetilde{\boldsymbol{\gamma}})\right]. \tag{39}$$

To derive the final expression for $\widehat{\boldsymbol{\phi}}$, we need to calculate $\nabla_{\boldsymbol{\gamma}}L(\widetilde{\boldsymbol{\gamma}})$ and $\nabla_{\boldsymbol{\gamma}}^2 L(\widetilde{\boldsymbol{\gamma}})$. It is straightforward to derive $L(\boldsymbol{\gamma})$ by making use of (36). It follows that

$$\nabla_{\boldsymbol{\gamma}}L(\widetilde{\boldsymbol{\gamma}}) = \mathbf{g}(\widetilde{\boldsymbol{\gamma}}), \tag{40}$$

where $[\mathbf{g}(\boldsymbol{\gamma})]_n = \frac{4[\boldsymbol{\gamma}]_n}{1-|[\boldsymbol{\gamma}]_n|^2} + \frac{\partial \ln\det(\boldsymbol{\Sigma}(\boldsymbol{\gamma}))}{\partial[\boldsymbol{\gamma}]_n}$, for $1 \leq n \leq M^2 - M$. Additionally, the Hessian matrix at $\widetilde{\boldsymbol{\gamma}}$ is obtained as

$$\nabla_{\boldsymbol{\gamma}}^2 L(\widetilde{\boldsymbol{\gamma}}) = N\text{diag}(\widehat{\mathbf{b}})\mathbf{F}^{-H}\widehat{\boldsymbol{\Sigma}}^{-1}\mathbf{F}^{-1}\text{diag}(\widehat{\mathbf{b}}) + \mathbf{E}(\widetilde{\boldsymbol{\gamma}}), \tag{41}$$

where $\widehat{\boldsymbol{\Sigma}} = \boldsymbol{\Sigma}(\widetilde{\boldsymbol{\gamma}})$, $[\widehat{\mathbf{b}}]_n = \frac{1}{\sqrt{1-|[\widetilde{\boldsymbol{\gamma}}]_n|^2}}$, for $1 \leq n \leq M^2 - M$ and $[\mathbf{E}(\boldsymbol{\phi})]_{n,l} = \frac{2\nu_n(1+|[\boldsymbol{\phi}]_n|^2)}{(1-|[\boldsymbol{\phi}]_n|^2)^2} + \frac{\partial^2 \ln \det(\boldsymbol{\Sigma}(\boldsymbol{\gamma}))}{\partial[\boldsymbol{\gamma}]_n \partial[\boldsymbol{\gamma}]_m}$. Inserting (40) and (41) into (39) leads to

$$\widehat{\boldsymbol{\phi}} \simeq \boldsymbol{\Psi}^{-1} \left(\overline{\mathbf{J}}^H \mathbf{F}^H \mathrm{diag}(\widehat{\mathbf{b}}) \mathbf{F}^{-H} \widehat{\boldsymbol{\Sigma}}^{-1} \mathbf{F}^{-1} \mathrm{diag}(\widehat{\mathbf{b}}) \mathbf{F} \overline{\mathbf{J}} + \overbrace{\frac{\mathbf{E}(\widetilde{\boldsymbol{\gamma}})}{N}}^{\hbar} \right)^{-1} \times \quad (42)$$

$$\overline{\mathbf{J}}^H \left(\mathbf{F}^H \mathrm{diag}(\widehat{\mathbf{b}}) \mathbf{F}^{-H} \widehat{\boldsymbol{\Sigma}}^{-1} \mathbf{F}^{-1} \mathrm{diag}(\widehat{\mathbf{b}}) \mathbf{F} \widetilde{\mathbf{r}} + \underbrace{\frac{\mathbf{F}^H \mathbf{E}(\widetilde{\boldsymbol{\gamma}}) \mathbf{F} \widetilde{\mathbf{r}} - \mathbf{F}^H \mathbf{g}(\widetilde{\boldsymbol{\gamma}})}{N}}_{\aleph} \right).$$

In the above equation, the terms $\hbar$ and $\aleph$ can be neglected for large N, thus (42) may be simplified as

$$\widehat{\boldsymbol{\phi}} \simeq \boldsymbol{\Psi}^{-1} \left(\overline{\mathbf{J}}^H \mathbf{F}^H \mathrm{diag}(\widehat{\mathbf{b}}) \mathbf{F}^{-H} \widehat{\boldsymbol{\Sigma}}^{-1} \mathbf{F}^{-1} \mathrm{diag}(\widehat{\mathbf{b}}) \mathbf{F} \overline{\mathbf{J}} \right)^{-1}$$

$$\times \overline{\mathbf{J}}^H \mathbf{F}^H \mathrm{diag}(\widehat{\mathbf{b}}) \mathbf{F}^{-H} \widehat{\boldsymbol{\Sigma}}^{-1} \mathbf{F}^{-1} \mathrm{diag}(\widehat{\mathbf{b}}) \mathbf{F} \widetilde{\mathbf{r}}. \quad (43)$$

Hence, from (31), an enhanced consistent estimate of $\overline{\mathbf{r}} = \mathrm{vec}(\overline{\mathbf{R}})$ is derived as follows

$$\widehat{\overline{\mathbf{r}}} = \mathbf{J} \begin{bmatrix} \mathbf{0} & \mathbf{I}_{D-1} & -j\mathbf{I}_{D-1} \\ 1 & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_{D-1} & j\mathbf{I}_{D-1} \end{bmatrix} \begin{bmatrix} 1 \\ \widehat{\boldsymbol{\phi}} \end{bmatrix}. \quad (44)$$

Remark 5. Considering $\lim_{N \to \infty} \widetilde{\mathbf{r}} = \ddot{\mathbf{r}}$, it is readily observed from (31) and (43) that $\widehat{\boldsymbol{\phi}}$ is a consistent estimate of $\boldsymbol{\phi}$. This in turn implies that $\widehat{\overline{\mathbf{r}}}$ is also a consistent estimate of $\overline{\mathbf{r}}$.

To estimate DoAs using $\widehat{\overline{\mathbf{r}}}$, we resort to CAB-MUSIC [53]. Specifically, we first construct the normalized augmented covariance matrix as

$$\widehat{\overline{\mathbf{R}}}_v = \begin{bmatrix} \mathbf{T}_v \mathbf{J}^\dagger \widehat{\overline{\mathbf{r}}} & \mathbf{T}_{v-1} \mathbf{J}^\dagger \widehat{\overline{\mathbf{r}}} & \cdots & \mathbf{T}_1 \mathbf{J}^\dagger \widehat{\overline{\mathbf{r}}} \end{bmatrix} \in \mathbb{C}^{v \times v}, \quad (45)$$

where $\mathbf{T}_i$ is a selection matrix, defined as

$$\mathbf{T}_i = \begin{bmatrix} \mathbf{0}_{v \times (i+D-v-1)} & \mathbf{I}_v & \mathbf{0}_{v \times (D-i)} \end{bmatrix} \in \{0,1\}^{v \times (2D-1)}. \quad (46)$$

It follows from the consistency of $\widehat{\overline{\mathbf{r}}}$ that

$$\lim_{N \to \infty} \widehat{\overline{\mathbf{R}}}_v = \begin{bmatrix} \mathbf{T}_v \mathbf{J}^\dagger \overline{\mathbf{r}} & \mathbf{T}_{v-1} \mathbf{J}^\dagger \overline{\mathbf{r}} & \cdots & \mathbf{T}_1 \mathbf{J}^\dagger \overline{\mathbf{r}} \end{bmatrix} \in \mathbb{C}^{v \times v}$$

$$= \mathbf{A}_v(\boldsymbol{\theta}) \mathrm{diag}(\overline{\mathbf{p}}) \mathbf{A}_v^H(\boldsymbol{\theta}) + \overline{\sigma}^2 \mathbf{I}_v, \quad (47)$$

where $\mathbf{A}_v(\boldsymbol{\theta}) = [\mathbf{a}_v(\theta_1), \mathbf{a}_v(\theta_2), \cdots, \mathbf{a}_v(\theta_K)] \in \mathbb{C}^{v \times K}$ denotes the steering matrix of a contiguous ULA with v elements located at $(0, \frac{\lambda}{2}, \cdots, (v-1)\frac{\lambda}{2})$. Hence, we can apply MUSIC to $\widehat{\overline{\mathbf{R}}}_v$ to estimate the DoAs. We call the proposed method Enhanced One-bit CAB-MUSIC (EOCAB-MUSIC). Algorithm 1 summarizes the steps of EOCAB-MUSIC.

Remark 6. The computational complexity of each step of Algorithm 1 is separately specified in Table I where $\mathcal{G}(n)$, $\mathcal{K}(n)$ and $\mathcal{Z}$ denote the complexity of the chosen algorithm for multiplication of two n-digit numbers, the complexity of integration in (108) and the number of grid point of the MUSIC algorithm, respectively. Considering that D and v are typically in the order of M^2 and, moreover, n and M are normally very smaller than $\mathcal{Z}$, it follows from Table I that the complexity of EOCAB-MUSIC is in the order of $\mathcal{O}(MN + M^2(\mathcal{G}(n)(\mathcal{Z} + M^4) + \mathcal{K}(n)M^2))$. On the other hand implementation of OCAB-MUSIC needs only steps 1, 2, 9 and 10 in algorithm 1. Hence, its complexity is given by $\mathcal{O}(MN + M^2\mathcal{G}(n)(\mathcal{Z} + M^4))$. Typically, we have $\mathcal{G}(n)(\mathcal{Z} + M^4) \gg \mathcal{K}(n)M^2$, implying that the complexity of EOCAB-MUSIC is almost in the same order as that of OCAB-MUSIC.

Algorithm 1 EOCAB-MUSIC

Input: SLA one-bit observations, i.e., $\mathbf{X}$.
Output: The estimates of source DoAs.
1: Compute the sample covariance matrix of one-bit data as $\widehat{\mathbf{R}}_{\mathbf{x}} = \frac{1}{N} \mathbf{X} \mathbf{X}^H$.
2: Compute $\widetilde{\overline{\mathbf{R}}}$ from (29).
3: Form $\widetilde{\mathbf{r}}$ by removing the diagonal elements of $\widetilde{\overline{\mathbf{R}}}$ from $\mathrm{vec}(\widetilde{\overline{\mathbf{R}}})$.
4: Compute $\widetilde{\boldsymbol{\gamma}}$ from $\widetilde{\boldsymbol{\gamma}} = \mathbf{F} \widetilde{\mathbf{r}}$.
5: Compute $\widehat{\mathbf{b}}$ using $[\widehat{\mathbf{b}}]_n = \frac{1}{\sqrt{1-|[\widetilde{\boldsymbol{\gamma}}]_n|^2}}$, for $1 \leq n \leq M^2 - M$.
6: Compute $\widehat{\boldsymbol{\Sigma}}$ by using (125) and replacing $\overline{\mathbf{R}}$ with $\widetilde{\overline{\mathbf{R}}}$ in (129), (130), (133) - (137), (139)-(148), (150) and (152)-(161) given in Appendix K.
7: Compute $\widehat{\boldsymbol{\phi}}$ from (43).
8: Compute $\widehat{\overline{\mathbf{r}}}$ from (44).
9: Compute $\widehat{\overline{\mathbf{R}}}_v$ from (45).
10: Apply MUSIC to $\widehat{\overline{\mathbf{R}}}_v$ to estimate DoAs.

Table I
COMPLEXITY OF THE STEPS OF ALGORITHM 1

Step order	Complexity
1	$\mathcal{O}(MN)$
2	$\mathcal{O}(\mathcal{G}(n)\sqrt{n}M^2)$
3	$\mathcal{O}(M^2)$
4	$\mathcal{O}(\mathcal{G}(n)M^4)$
5	$\mathcal{O}(\mathcal{G}(n)M^2)$
6	$\mathcal{O}(\mathcal{K}(n)M^4)$
7	$\mathcal{O}(\mathcal{G}(n)(DM^4 + M^6 + D^3))$
8	$\mathcal{O}(\mathcal{G}(n)(D^2M^2))$
9	$\mathcal{O}(\mathcal{G}(n)(M^2v(2D-1+v)))$
10	$\mathcal{O}(\mathcal{G}(n)(\mathcal{Z}M^2 + M^3))$

B. Asymptotic Performance Analysis

In this section, we investigate the asymptotic performance of the proposed estimator through the derivation of a closed-form expression for the second-order statistics of the asymptotic distribution (as $N \to \infty$) of the DoA estimation errors. Our main results are summarized in Theorem 5, Corollary 1 and Theorem 6.

Lemma 3. $\widehat{\boldsymbol{\theta}}$ obtained by EOCAB-MUSIC is a consistent estimate of $\boldsymbol{\theta}$ if $K \leq v - 1$.

Proof. See Appendix G $\qquad \square$

Theorem 5. *The closed-form expression for the covariance of the asymptotic distribution (as $N \to \infty$) of the DoA estimation errors obtained by EOCAB-MUSIC is given by*

$$\mathcal{E}_{\theta_{k_1}, \theta_{k_2}} = \mathbb{E}\{(\theta_{k_1} - \widehat{\theta}_{k_1})(\theta_{k_2} - \widehat{\theta}_{k_2})^*\} \quad (48)$$

$$= \frac{(\sigma^2 + \sum_{k=1}^K p_k)^2}{N\pi^2 p_{k_1} p_{k_2} q_{k_1} q_{k_2} \cos\theta_{k_1} \cos\theta_{k_2}}$$

$$\times \Re\{\mathbf{z}_{k_1}^T \overline{\mathbf{T}}(\overline{\mathbf{J}}^H \mathbf{W} \overline{\mathbf{J}})^{-1} \overline{\mathbf{J}}^H \mathbf{W} \boldsymbol{\Gamma} \mathbf{W} \overline{\mathbf{J}} (\overline{\mathbf{J}}^H \mathbf{W} \overline{\mathbf{J}})^{-1} \overline{\mathbf{T}}^H \mathbf{z}_{k_2}^*\},$$

where

$$\mathbf{z}_k = \boldsymbol{\beta}_k \otimes \boldsymbol{\alpha}_k, \quad (49)$$

$$\boldsymbol{\beta}_k = \Pi_{\mathbf{A}_v(\boldsymbol{\theta})}^\perp \mathrm{diag}(\mathbf{v}) \mathbf{a}_v(\theta_k), \quad (50)$$

$$\boldsymbol{\alpha}_k = \mathbf{A}_v^{\dagger T}(\boldsymbol{\theta}) \boldsymbol{\imath}_k, \quad (51)$$

$$q_k = \mathbf{a}_v^H(\theta_k)\mathrm{diag}(\mathbf{v})\Pi_{\mathbf{A}_v}^{\perp}\mathrm{diag}(\mathbf{v})\mathbf{a}_v(\theta_k), \tag{52}$$

$$\mathbf{W} = \mathbf{F}^H\mathrm{diag}(\mathbf{b})\mathbf{F}^{-H}\mathbf{\Sigma}^{-1}\mathbf{F}^{-1}\mathrm{diag}(\mathbf{b})\mathbf{F}, \tag{53}$$

$$[\mathbf{\Gamma}]_{p,q} = \frac{1}{2}\Bigg(\sqrt{1-[\Re\{\ddot{\mathbf{r}}_p\}]^2}\times\sqrt{1-[\Re\{\ddot{\mathbf{r}}_q\}]^2} \tag{54}$$

$$+\sqrt{1-[\Im\{\ddot{\mathbf{r}}_p\}]^2}\times\sqrt{1-[\Im\{\ddot{\mathbf{r}}_q\}]^2}\Bigg)\Re\{[\mathbf{\Sigma}]_{p,q}\}$$

$$+\frac{j}{2}\Bigg(\sqrt{1-[\Im\{\ddot{\mathbf{r}}_p\}]^2}\times\sqrt{1-[\Re\{\ddot{\mathbf{r}}_q\}]^2}$$

$$+\sqrt{1-[\Re\{\ddot{\mathbf{r}}_p\}]^2}\times\sqrt{1-[\Im\{\ddot{\mathbf{r}}_q\}]^2}\Bigg)\Im\{[\mathbf{\Sigma}]_{p,q}\},$$

with $\mathbf{v} = [0,1,2,\cdots,v-1]^T$, $[\mathbf{b}]_n = \frac{1}{\sqrt{1-|[\boldsymbol{\gamma}]_n|^2}}$ for $1 \le n \le M^2-M$, $\mathbf{\Sigma} \in \mathbb{C}^{(M^2-M)\times(M^2-M)}$ as given in Appendix K (kindly refer to the supplementary document), $\overline{\mathbf{T}} \in \mathbb{C}^{v^2\times(2D-2)}$ as defined in (105) in Appendix H, and $\boldsymbol{\imath}_k$ being the kth column of $\mathbf{I}_K$.

Proof. See Appendix H $\qquad\qquad\square$

Corollary 1. *The asymptotic MSE expression (as $N \to \infty$) for the DoA estimates obtained by EOCAB-MUSIC is given by*

$$\mathcal{E}_{\theta_k} = \mathbb{E}\{(\theta_{k_1}-\widehat{\theta}_k)^2\} = \frac{(\sigma^2+\sum_{k'=1}^K p_{k'})^2}{N\pi^2 p_k^2 q_k^2\cos^2\theta_k} \tag{55}$$

$$\times\Re\{\mathbf{z}_k^T\overline{\mathbf{T}}(\overline{\mathbf{J}}^H\mathbf{W}\overline{\mathbf{J}})^{-1}\overline{\mathbf{J}}^H\mathbf{W}\mathbf{\Gamma}\mathbf{W}\overline{\mathbf{J}}(\overline{\mathbf{J}}^H\mathbf{W}\overline{\mathbf{J}})^{-1}\overline{\mathbf{T}}^H\mathbf{z}_k^*\}.$$

Corollary 2. *The covariance of the asymptotic distribution (as $N \to \infty$) of the DoA estimation errors and the asymptotic MSE expression (as $N \to \infty$) for the one-bit DoA estimator given in [53], named as One-bit CAB-MUSIC (OCAB-MUSIC), is easily obtained by replacing $\mathbf{W}$ with $\mathbf{I}_{M^2-M}$ in (48) and (55), respectively.*

Proof. See Appendix I $\qquad\qquad\square$

Remark 7. It is concluded from Corollary 1 and Corollary 2 that, similar to Infinite-bit Co-Array-Based MUSIC (ICAB-MUSIC) [21], the MSEs of EOCAB-MUSIC and OCAB-MUSIC depend on both the physical and the virtual array geometries through $\mathbf{A}_v(\theta)$ and $\overline{\mathbf{R}}$, respectively.

Remark 8. Another interesting implication of Corollary 1 is that the MSEs of EOCAB-MUSIC and OCAB-MUSIC reduce at the same rate as that of ICAB-MUSIC [21] with respect to N; i.e. $\mathcal{E}_{\theta_k} \propto \frac{1}{N}$ for both.

Remark 9. It is readily clear from the definition that $\overline{\mathbf{r}}$ is a function of the SNR, and not $\mathbf{p}$ and σ^2. This indicates that $\mathbf{W}$ and $\mathbf{\Gamma}$ are also functions of the SNR instead of $\mathbf{p}$ and σ^2. Further, multiplying the numerator and denominator of $(\sigma^2+\sum_{k'=1}^K p_{k'})^2/p_k^2$ by $1/\sigma^4$ reformulates it as a function of the SNR. These observations imply that the MSEs of EOCAB-MUSIC and OCAB-MUSIC are functions of the SNR instead of $\mathbf{p}$ and σ^2. This fact can also be deduced directly from system model where we have

$$[\mathbf{x}(t)]_m = \frac{1}{\sqrt{2}}\mathrm{sgn}\left(\Re\{[\mathbf{y}(t)]_m\}\right)+\frac{j}{\sqrt{2}}\,\mathrm{sgn}\left(\Im\{[\mathbf{y}(t)]_m\}\right)$$

$$= \frac{1}{\sqrt{2}}\mathrm{sgn}\left(\Re\{\frac{[\mathbf{y}(t)]_m}{\sigma}\}\right)+\frac{j}{\sqrt{2}}\,\mathrm{sgn}\left(\Im\{\frac{[\mathbf{y}(t)]_m}{\sigma}\}\right). \tag{56}$$

for $\sigma > 0$. This implies that, without loss of generality, we can consider the power of each source equal to the SNR for that source and the noise variance equal to 1.

Theorem 6. *Assume all sources have equal power p and $SNR = p/\sigma^2$. Then, for a sufficiently large SNR, the MSE of EOCAB-MUSIC converges to the following constant value:*

$$\lim_{SNR\to\infty}\mathcal{E}_{\theta_k} = \frac{K^2}{N\pi^2 q_k^2\cos^2\theta_k}\times \tag{57}$$

$$\Re\{\mathbf{z}_k^T\overline{\mathbf{T}}(\overline{\mathbf{J}}^H\mathbf{W}_\infty\overline{\mathbf{J}})^{-1}\overline{\mathbf{J}}^H\mathbf{W}_\infty\mathbf{\Gamma}_\infty\mathbf{W}_\infty\overline{\mathbf{J}}(\overline{\mathbf{J}}^H\mathbf{W}_\infty\overline{\mathbf{J}})^{-1}\overline{\mathbf{T}}^H\mathbf{z}_k^*\} > 0,$$

where $\mathbf{W}_\infty$ and $\mathbf{\Gamma}_\infty$ are obtained by replacing $\overline{\mathbf{R}}$, $\ddot{\mathbf{r}}$ and $\boldsymbol{\gamma}$ in the definitions of $\mathbf{W}$ and $\mathbf{\Gamma}$ (kindly refer to Theorem 5) with $\overline{\mathbf{R}}_\infty$, $\boldsymbol{\gamma}_\infty$ and $\ddot{\mathbf{r}}_\infty$, respectively, where

$$\overline{\mathbf{R}}_\infty = \frac{1}{K}\mathbf{A}(\boldsymbol{\theta})\mathbf{A}^H(\boldsymbol{\theta}) + (1-\frac{1}{K})\mathbf{I}_M, \tag{58}$$

$\boldsymbol{\gamma}_\infty$ is the $(M^2-M)\times 1$ vector containing the real and imaginary parts of the elements of $\overline{\mathbf{R}}_\infty$ above its main diagonal elements and $\ddot{\mathbf{r}}_\infty = \mathbf{\Psi}^{-1}\overline{\mathbf{J}}^{\dagger}\mathbf{F}^{-1}\boldsymbol{\gamma}_\infty$.

Proof. See Appendix J. $\qquad\qquad\square$

Remark 10. It follows from Theorem 6 that it is not possible to make the MSEs of EOCAB-MUSIC and OCAB-MUSIC arbitrarily small by increasing the SNR.

VI. SIMULATION RESULTS

In this section, we provide some numerical results to validate the analytical results obtained in previous sections as well as to assess the performance of the proposed DoA estimator. Specifically, we will show that the proposed estimator yields better performance in terms of estimation accuracy and resolution compared to the approach given in [53]. In the rest of this section, we will refer to: 1) the CRB for DoA estimation from infinite-bit measurements as Infinite-bit CRB (I-CRB), whose expression is given in Remark 3; 2) the pessimistic approximation of the CRB for DoA estimation from one-bit measurements as One-bit CRB (O-CRB); 3) CAB-MUSIC using infinite-bit measurements as Infinite-bit CAB-MUSIC (ICAB-MUSIC); 4) the DoA estimator given in [53] as one-bit CAB-MUSIC (OCAB-MUSIC); 5) the proposed estimator in this paper as Enhanced One-bit CAB-MUSIC (EOCAB-MUSIC).

A. General Set-up

In all experiments, each simulated point has been computed by 5000 Monte Carlo repetitions. Unless the source locations are specified for a particular result, it is assumed that the K independent sources are equally spaced in the angular domain $[-60°, 60°]$ such that $\theta = -60°$ when $K = 1$. Further, all sources are assumed to have equal powers, i.e., $p_k = p$ for all k, and the SNR is defined as $10\log\frac{p}{\sigma^2}$. For our numerical investigation, we use four different types of arrays with $M = 10$ physical elements and the following geometries:

$$\mathbb{M}_{\text{nested}} : \{1,2,3,4,5,6,12,18,24,30\}, \tag{59}$$

$$\mathbb{M}_{\text{co-prime}} : \{0,3,5,6,9,10,12,15,20,25\}, \tag{60}$$

$$\mathbb{M}_{\text{MRA}} : \{0,1,3,6,13,20,27,31,35,36\}, \tag{61}$$

$$\mathbb{M}_{\text{ULA}} : \{0,1,2,\cdots,9\}. \tag{62}$$

These arrays generate the difference co-arrays:

$$\mathbb{D}_{\text{nested}} : \{0,1,2,\cdots,29\}, \tag{63}$$

$$\mathbb{D}_{\text{co-prime}} : \{0,1,2,\cdots,22,25\}, \tag{64}$$

$$\mathbb{D}_{\text{MRA}} : \{0,1,2,\cdots,36\}, \tag{65}$$

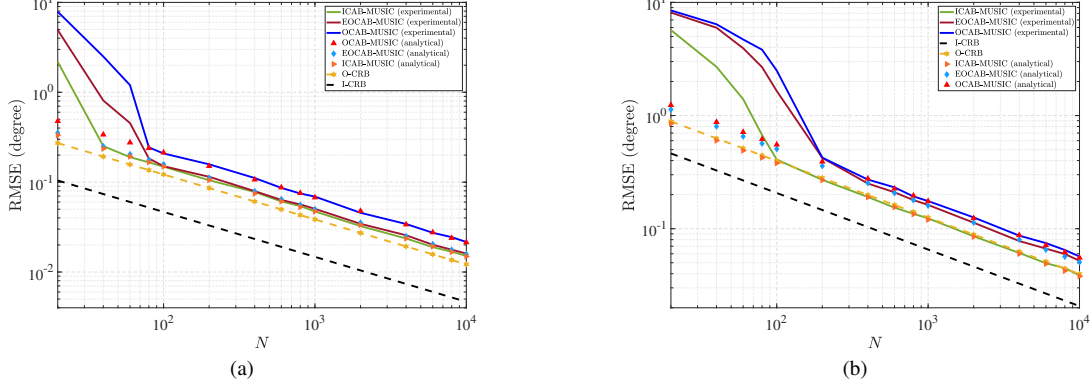$$\mathbb{D}_{\text{ULA}} : \{0,1,2,\cdots,9\}. \tag{66}$$

Figure 2. RMSE in degrees for θ_2 versus N for a nested array with $M = 10$ elements and configuration given in (59), SNR = 3 dB, and: (a) $K = 5 < M$; (b) $K = 12 > M$.
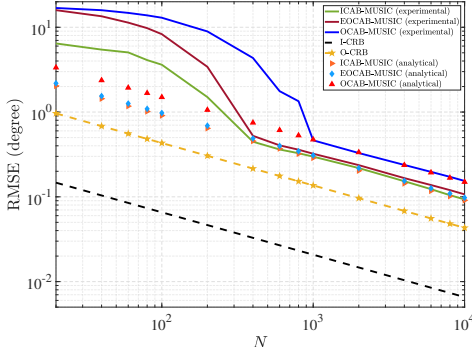

Figure 3. RMSE in degrees for θ_2 versus N for a nested array with $M = 10$ elements and configuration given in (59) when $K = 3$, $\theta_1 = 2°$, $\theta_2 = 3°$, $\theta_3 = 75°$, $\mathrm{SNR}_1 = 20$ dB, $\mathrm{SNR}_2 = 8$ dB and $\mathrm{SNR}_3 = 22$ dB.

Further, we generate the grid from $-90°$ to $90°$ with step size $0.001°$ to implement MUSIC. To avoid griding, alternatively, it is also possible to use root-MUSIC.

B. MSE vs. the Number of Snapshots

Fig. 2 depicts the Root-Mean-Squares-Error (RMSE) for θ_2 in degree versus the number of snapshots when the nested array in (59) is used. The SNR is assumed to be 3 dB. In addition, noting $M = 10$, two different scenarios are considered: (a) $K = 5 < M$, and (b) $K = 12 > M$. Fig. 2 illustrates a close agreement between the numerical simulations and analytical expression derived for RMSEs of OCAB-MUSIC and EOCAB-MUSIC when about 200 or more snapshots are available. Further, a considerable gap is observed between the performance of OCAB-MUSIC and that of the EOCAB-MUSIC. For instance, at $N = 400$, Figs. 2a and 2b show a performance gain of roughly 3 dB and 1 dB, respectively, in terms of the RMSE when the EOCAB-MUSIC is used. It is also observed that EOCAB-MUSIC performs as well as ICAB-MUSIC when $K = 5 < M$. Further, it is observed that the RMSE of EOCAB-MUSIC is very close to O-CRB when $K = 5 < M$ but we see a gap between them when $K = 12 > M$.

Fig. 2 also shows that when a small number of snapshots is available, e.g. less than 1000, all estimators are confronted with substantial performance degradation. This performance loss

is justified by the subspace swap arising from the inaccurate estimate of the normalized covariance matrix of $\mathbf{y}(t)$, i.e. $\overline{\mathbf{R}}$, in this case. However, it is seen that the proposed estimator still has superior performance compared to OCAB-MUSIC, even in the low snapshot paradigm.

Fig. 3 depicts the RMSE θ_2 in degree versus the number of snapshots when $K = 3$ and the sources powers are unequal. Specifically, It is assumed that $\theta_1 = 2°$, $\theta_2 = 3°$, $\theta_3 = 75°$, $\mathrm{SNR}_1 = 20$ dB, $\mathrm{SNR}_2 = 8$ dB and $\mathrm{SNR}_3 = 22$ dB. Comparing Fig. 2 with Fig. 3 reveals that a high difference between the SNRs of the closely-spaced source signals do not have a meaningful impact on the relative asymptotic performance of ICAB-MUSIC, OCAB-MUSIC and EOCAB-MUSIC, however, by increasing the difference between SNRs, OCAB-MUSIC needs more number of snapshots to achieve its asymptotic performance compared to EOCAB-MUSIC and ICAB-MUSIC.

C. MSE vs. SNR

Fig. 4 shows the RMSE for θ_2 in degrees versus SNR for the same setup used for Fig. 2. The number of snapshots is considered to be $N = 500$. It is seen in Figs. 4a and Fig. 4b that the RMSEs of OCAB-MUSIC and EOCAB-MUSIC perfectly match with their asymptotic analytical RMSEs given in Corollary 1 and Corollary 2.

Fig. 4 demonstrates that the I-CRB tends to decay to zero as the SNR increases when $K = 5 < M$ while it gets saturated as the SNR increases when $K = 12 > M$. However, as opposed to the I-CRB, O-CRB tends to converge to a constant non-zero value at the high SNR regime for both the cases $K = 5 < M$ and $K = 12 > M$. This behavior of O-CRB was already predicted by Theorem 4. In addition, as shown in Theorem 6, the RMSEs of OCAB-MUSIC and EOCAB-MUSIC also converge to a constant non-zero value as the SNR increases for both $K = 5 < M$ and $K = 12 > M$.

We observe from Fig. 4 that EOCAB-MUSIC preforms better than OCAB-MUSIC in both scenarios $K = 5 < M$ and $K = 12 > M$. For example, at $\mathrm{SNR} = 5$, EOCAB-MUSIC leads to performance gains of about 3.7 dB and 1.15 dB in terms of RMSE compared to OCAB-MUSIC. Further, it is seen that EOCAB-MUSIC even outperforms ICAB-MUSIC at high SNR regime when $K = 5 < M$. Another interesting
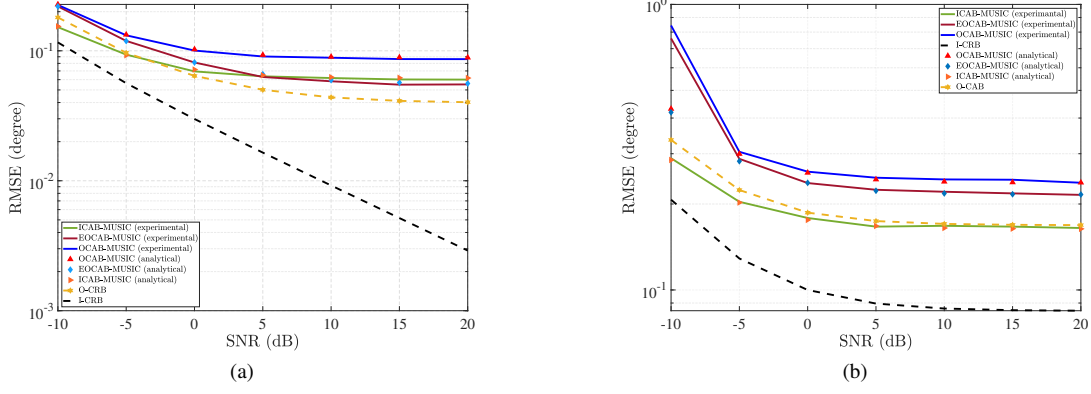
Figure 4. RMSE in degrees for θ_2 versus SNR wen the source powers are equal for a nested array with $M = 10$ elements and configuration given in (59), $N = 500$, and: (a) $K = 5 < M$; (b) $K = 12 > M$.
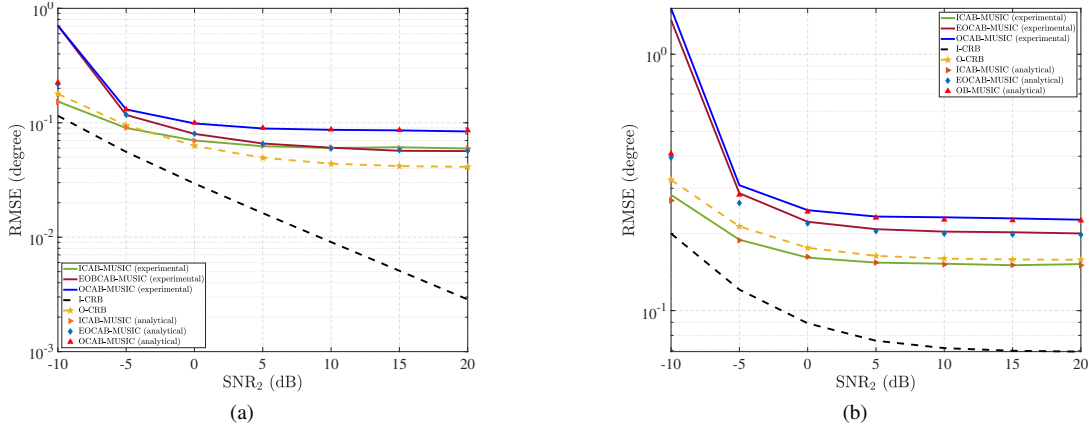

Figure 5. RMSE in degrees for θ_2 versus SNR when the source powers are unequal for a nested array with $M = 10$ elements and configuration given in (59), $N = 500$, and: (a) $K = 5 < M$; (b) $K = 12 > M$.

observation is that the RMSE of O-CRB is either better or equal to that of ICAB-MUSIC.

Fig. 5 shows the RMSE for θ_2 in degrees versus SNR when the sources powers are unequal and DoAs are not exactly on the grid as opposed to Fig. 4. The number of snapshots is considered to be $N = 500$. In case of $K = 5 < M$, the sources are located at $\theta_1 = -49.4551°, \theta_2 = -30.1443°, \theta_3 = -2.4525°, \theta_4 = 26.8293°$ and $\theta_5 = 56.5149°$. Further, the source SNRs are assumed to be $\text{SNR}_1 = 0.75 \times \text{SNR}_2, \text{SNR}_3 = 1.22 \times \text{SNR}_2, \text{SNR}_4 = 0.92 \times \text{SNR}_2$ and $\text{SNR}_5 = 0.66 \times \text{SNR}_2$ while SNR_2 varies from 10 dB to 20 dB as shown in Fig. 5a. Further, in case of $K = 12 > M$, the sources are located at $\theta_1 = -56.3351°, \theta_2 = -36.2628°, \theta_3 = -19.9004°, \theta_4 = -2.4093°, \theta_5 = 0.0027°, \theta_6 = 13.1840°, \theta_7 = 23.8495°, \theta_8 = 25.8044°, \theta_9 = 29.2889°, \theta_10 = 40.9107°, \theta_11 = 48.4465°$ and $\theta_12 = 48.5667°$. The source SNRs are assumed to be $\text{SNR}_1 = 1.34 \times \text{SNR}_2, \text{SNR}_3 = 0.84 \times \text{SNR}_2, \text{SNR}_4 = 0.83 \times \text{SNR}_2, \text{SNR}_5 = 0.67 \times \text{SNR}_2, \text{SNR}_6 = 0.69 \times \text{SNR}_2, \text{SNR}_7 = 0.95 \times \text{SNR}_2, \text{SNR}_8 = 0.61 \times \text{SNR}_2, \text{SNR}_9 = 0.79 \times \text{SNR}_2, \text{SNR}_{10} = 0.56 \times \text{SNR}_2, \text{SNR}_10 = 0.82 \times \text{SNR}_2$ and $\text{SNR}_12 = 0.88 \times \text{SNR}_2$ while SNR_2 varies from 10 dB to 20 dB as shown in Fig. 5b. Comparing Fig. 5 with Fig. 4 reveals that unequal source powers do not have remarkable impact on the estimation accuracy particularly in high-SNR regime.

D. CRB vs. the Number of Source Signals

Fig. 6 plots the I-CRB and the O-CRB for θ_2 in degree versus the number of source signals for $\text{SNR} = 3$ dB and $N = 500$ and different types of arrays given in (59), (60), (61) and (62). The values of D and v for the different types of arrays are as: 1) MRA: $D = 37$ and $v = 37$; 2) nested array: $D = 30$ and $v = 30$; 3) co-prime array: $D = 26$ and $v = 23$; ULA: $D = 10$ and $v = 10$. Fig. 6 indicates that both the I-CRB and the O-CRB increase as the number of source signals increases. Moreover, it is observed that the I-CRB and the O-CRB are quite small for all the SLAs as long as $1 \le K \le v - 1$, but they escalate dramatically when K approaches values that are equal to or larger than D. This observation is in compliance with Theorem 2 which indicates that the DoA estimation problem is globally identifiable when $1 \le K \le v - 1$ and is globally non-identifiable when $K \ge D$.

E. Resolution Probability

Fig. 7 depicts the probability of resolution versus the source separation for ICAB-MUSIC, EOCAB-MUSIC and OCAB-MUSIC when the nested array given in (59) is employed. The number of snapshots and the SNR are considered to be $N = 500$ and 0 dB, respectively. In addition, we consider two sources with equal powers, located at $\theta_1 = 20° - \frac{\Delta\theta}{2}$ and $\theta_2 = 20° + \frac{\Delta\theta}{2}$. We define the two sources as being resolvable if $\max_{i \in \{1,2\}} |\hat{\theta}_i - \theta_i| < \frac{\Delta\theta}{2}$ [64]. According to this definition and
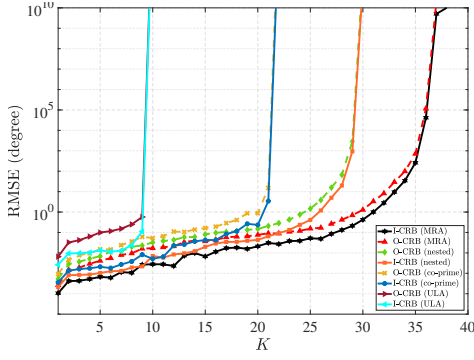
Figure 6. The CRB versus K for various array configurations given from (59) to (61), $N = 500$ and SNR = 3 dB.
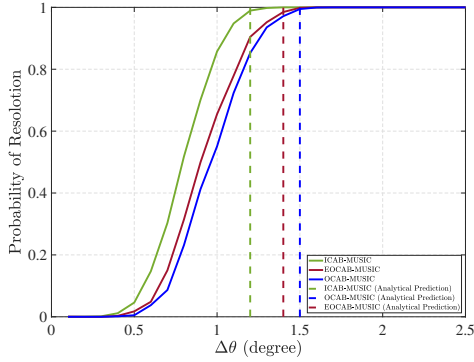

Figure 7. Probability of resolution versus source separation in degree for a nested array with $M = 10$ elements and configuration given in (59), $N = 500$ and SNR = 0 dB.

making use of two-dimensional Chebychev's bound [65], the probability of resolution can be lower bounded as

$$\mathbb{P}(\max_{i \in \{1,2\}} |\hat{\theta}_i - \theta_i| < \frac{\Delta\theta}{2}) \tag{67}$$

$$= \mathbb{P}(|\hat{\theta}_1 - \theta_1| < \frac{\Delta\theta}{2}, |\hat{\theta}_2 - \theta_2| < \frac{\Delta\theta}{2}) \geq 1 - \frac{2[\mathcal{E}(\theta_1) + \mathcal{E}(\theta_2)]}{\Delta\theta^2}$$

$$+ \frac{2\sqrt{\mathcal{E}_{\theta_1}^2 + \mathcal{E}_{\theta_2}^2 + 2\mathcal{E}_{\theta_1}\mathcal{E}_{\theta_2} - 4\mathcal{E}_{\theta_1,\theta_2}^2}}{\Delta\theta^2},$$

where $\mathcal{E}(\theta_1)$, $\mathcal{E}(\theta_2)$ and $\mathcal{E}(\theta_1, \theta_2)$ are given in (55) and (48). The analytical expression on the right-hand side of (67) enables us to predict the minimum source separation required for achieving a particular probability of resolution. For example, Fig. 6 shows the predicted values for the minimum source separation to achieve a probability of resolution greater than 0.9, obtained from (67), for ICAB-MUSIC, OCAB-MUSIC and EOCAB-MUSIC. It is observed that the predicted values of the minimum source separation for ICAB-MUSIC, EOCAB-MUSIC and OCAB-MUSIC, which are respectively $\Delta\theta = 1.2°$, $\Delta\theta = 1.4°$ and $\Delta\theta = 1.5°$, are in a good agreement with the values obtained from the numerical simulations, which are respectively $\Delta\theta = 1.1°$, $\Delta\theta = 1.2°$ and $\Delta\theta = 1.3°$. Additionally, Fig. 7 demonstrates the resolution performance of EOCAB-MUSIC is superior to that of OCAB-MUSIC while ICAB-MUSIC outperforms both of them.

VII. CONCLUSION

In this paper, we considered the problem of DoA estimation from one-bit measurements received by an SLA. We showed that the idetifiability condition for the DoA estimation problem from one-bit SLA data is equivalent to that for the case when DoAs are estimated from infinite-bit unquantized measurements. Then, we derived a pessimistic approximation of the corresponding CRB. This pessimistic CRB was used as a benchmark for assessing the performance of one-bit DoA estimators. Further, it provides us with valuable insights on the performance limits of DoA estimation from one-bit quantized data. For example, it was shown that the DoA estimation errors in one-bit scenario reduces at the same rate as that of infinite-bit case with respect to the number of samples and, moreover, that the DoA estimation errors in one-bit scenario converges to a constant value by increasing the SNR. We also proposed a new algorithm for estimating DoAs from one-bit quantized data. We investigated the analytical performance of the proposed method through deriving a closed-form expression for the second-order statistics of its asymptotic distribution (for the large number of snapshots) and show that it outperforms the existing algorithms in the literature. Numerical simulations were provided to validate the analytical derivations and corroborate the improvement in estimation performance.

APPENDIX A
PROOF OF THEOREM 1

We first prove the sufficiency. Assume that $\boldsymbol{\theta}_0 \in [-\pi/2, \pi/2]^{K \times 1}$ is identifiable from $\mathbf{Y}$. This implies that $f(\mathbf{Y} \mid \boldsymbol{\theta}_0, \mathbf{p}, \sigma^2) \neq f(\mathbf{Y} \mid \breve{\boldsymbol{\theta}}, \breve{\mathbf{p}}, \breve{\sigma}^2)$ for any arbitrary values of $\breve{\boldsymbol{\theta}} \neq \boldsymbol{\theta}_0 \in [-\pi/2, \pi/2]^{K \times 1}$, $\mathbf{p} \in \mathbb{R}_{>0}^{K \times 1}$, $\breve{\mathbf{p}} \in \mathbb{R}_{>0}^{K \times 1}$, σ^2 and $\breve{\sigma}^2$. Hence, considering $\mathbf{y}(0), \mathbf{y}(1), \cdots, \mathbf{y}(N-1)$ are independent and identically distributed with $\mathbf{y}(t) \sim \mathcal{CN}(\mathbf{0}, \mathbf{R})$, we have

$$\mathbf{A}(\boldsymbol{\theta}_0)\mathrm{diag}(\mathbf{p})\mathbf{A}^H(\boldsymbol{\theta}_0) + \sigma^2\mathbf{I}_M \neq \mathbf{A}(\breve{\boldsymbol{\theta}})\mathrm{diag}(\breve{\mathbf{p}})\mathbf{A}^H(\breve{\boldsymbol{\theta}}) + \breve{\sigma}^2\mathbf{I}_M, \tag{68}$$

for all $\breve{\boldsymbol{\theta}} \neq \boldsymbol{\theta}_0 \in [-\pi/2, \pi/2]^{K \times 1}$, $\mathbf{p} \in \mathbb{R}_{>0}^{K \times 1}$, $\breve{\mathbf{p}} \in \mathbb{R}_{>0}^{K \times 1}$, σ^2 and $\breve{\sigma}^2$.

In what follows, we employ the method of proof by contradiction to prove the sufficiency. In particular, we assume that $\boldsymbol{\theta}_0 \in [-\pi/2, \pi/2]^{K \times 1}$ is non-identifiable from $\mathbf{X}$. Hence, there exists a $\breve{\boldsymbol{\theta}} \neq \boldsymbol{\theta}_0 \in [-\pi/2, \pi/2]^{K \times 1}$ at which $f(\mathbf{X} \mid \boldsymbol{\theta}_0, \widetilde{\mathbf{p}}, \widetilde{\sigma}^2) = f(\mathbf{X} \mid \breve{\boldsymbol{\theta}}, \dot{\mathbf{p}}, \dot{\sigma}^2)$ for some values of $\widetilde{\mathbf{p}} \in \mathbb{R}_{>0}^{K \times 1}$, $\dot{\mathbf{p}} \in \mathbb{R}_{>0}^{K \times 1}$, $\widetilde{\sigma}^2$ and $\dot{\sigma}^2$. It is readily clear from assumption **A4** and (6) that $\mathbb{E}\{\mathbf{x}(t_1)\mathbf{x}^H(t_2)\} = \mathbf{0}$ when $t_1 \neq t_2$. Accordingly, we have

$$\mathbb{E}\left\{\mathbf{X}\mathbf{X}^H \mid \boldsymbol{\theta}_0, \widetilde{\mathbf{p}}, \widetilde{\sigma}^2\right\} = \mathbb{E}\left\{\mathbf{X}\mathbf{X}^H \mid \breve{\boldsymbol{\theta}}, \dot{\mathbf{p}}, \dot{\sigma}^2\right\}, \tag{69}$$

$$\Rightarrow \sum_{t=0}^{N-1} \mathbb{E}\{\mathbf{x}(t)\mathbf{x}^H(t) \mid \boldsymbol{\theta}_0, \widetilde{\mathbf{p}}, \widetilde{\sigma}^2\} = \sum_{t=0}^{N-1} \mathbb{E}\{\mathbf{x}(t)\mathbf{x}^H(t) \mid \breve{\boldsymbol{\theta}}, \dot{\mathbf{p}}, \dot{\sigma}^2\}.$$

From (69), (7), (3) and the fact that the arcsine function is one-to-one when its argument is between -1 and 1, it follows that

$$\frac{1}{\widetilde{\sigma}^2 + \sum_{k=1}^K \widetilde{p}_k}\left[\mathbf{A}(\boldsymbol{\theta}_0)\mathrm{diag}(\widetilde{\mathbf{p}})\mathbf{A}^H(\boldsymbol{\theta}_0) + \widetilde{\sigma}^2\mathbf{I}_M\right] =$$

$$\frac{1}{\dot{\sigma}^2 + \sum_{k=1}^K \dot{p}_k}\left[\mathbf{A}(\breve{\boldsymbol{\theta}})\mathrm{diag}(\dot{\mathbf{p}})\mathbf{A}^H(\breve{\boldsymbol{\theta}}) + \dot{\sigma}^2\mathbf{I}_M\right]. \tag{70}$$

Considering $\mathbf{p} = \frac{\widetilde{\mathbf{p}}}{\widetilde{\sigma}^2 + \sum_{k=1}^K \widetilde{p}_k}$, $\sigma^2 = \frac{\widetilde{\sigma}^2}{\widetilde{\sigma}^2 + \sum_{k=1}^K \widetilde{p}_k}$, $\breve{\mathbf{p}} = \frac{\dot{\mathbf{p}}}{\dot{\sigma}^2 + \sum_{k=1}^K \dot{p}_k}$ and $\breve{\sigma}^2 = \frac{\dot{\sigma}^2}{\dot{\sigma}^2 + \sum_{k=1}^K \dot{p}_k}$, we obtain

$$\mathbf{A}(\boldsymbol{\theta}_0)\mathrm{diag}(\mathbf{p})\mathbf{A}^H(\boldsymbol{\theta}_0) + \sigma^2\mathbf{I}_M = \mathbf{A}(\breve{\boldsymbol{\theta}})\mathrm{diag}(\breve{\mathbf{p}})\mathbf{A}^H(\breve{\boldsymbol{\theta}}) + \breve{\sigma}^2\mathbf{I}_M, \tag{71}$$

which is in contradiction with (68). Hence, the initial assumption that $\boldsymbol{\theta}_0 \in [-\pi/2, \pi/2]^{K \times 1}$ is non-identifiable from $\mathbf{X}$ cannot be true. This proves the sufficiency.

To show the necessity, let assume that $\boldsymbol{\theta}_0 \in [-\pi/2, \pi/2]^{K \times 1}$ is non-identifiable from $\mathbf{Y}$. This implies that there exist some $\breve{\boldsymbol{\theta}} \in [-\pi, \pi]^{q \times 1} \neq \boldsymbol{\theta}_0$, $\mathbf{p}$, $\breve{\mathbf{p}}$, σ^2 and $\breve{\sigma}^2$ for which $f(\mathbf{Y} \mid \boldsymbol{\theta}_0, \mathbf{p}, \sigma^2) = f(\mathbf{Y} \mid \breve{\boldsymbol{\theta}}, \breve{\mathbf{p}}, \breve{\sigma}^2)$. Since the true PDF of $\mathbf{X}$ is obtained from the orthant probabilities of $\mathbf{Y}$, it is readily deduced that $f(\mathbf{X} \mid \boldsymbol{\theta}_0, \mathbf{p}, \sigma^2) = f(\mathbf{X} \mid \breve{\boldsymbol{\theta}}, \breve{\mathbf{p}}, \breve{\sigma}^2)$ as well. This proves that identifiability of $\boldsymbol{\theta}_0 \in [-\pi/2, \pi/2]^{K \times 1}$ from $\mathbf{Y}$ is a necessary condition for identifiability of $\boldsymbol{\theta}_0 \in [-\pi/2, \pi/2]^{K \times 1}$ from $\mathbf{X}$.

APPENDIX B
PROOF OF THEOREM 2

We first prove **S1**. Consider arbitrary $\boldsymbol{\theta} \neq\, \in [-\pi/2, \pi/2]^{K \times 1}$ and $\breve{\boldsymbol{\theta}} \in [-\pi/2, \pi/2]^{K \times 1}$ such that $\boldsymbol{\theta}_k \neq \breve{\boldsymbol{\theta}}$. Moreover, let $\mathbf{A}_v(\boldsymbol{\theta})$ be the steering matrix of a contiguous ULA with v elements located at $(0, \frac{\lambda}{2}, \cdots, (v-1)\frac{\lambda}{2})$. Considering the fact that $\mathbf{A}_v(\boldsymbol{\theta})$ is a Vandermonde matrix, if $K \leq v-1$, it follows from Caratheodory-Fejer-Pisarenko decomposition [66] that

$$\mathbf{A}_v(\boldsymbol{\theta})\mathrm{diag}(\mathbf{p})\mathbf{A}_v^H(\boldsymbol{\theta}) + \sigma^2 \mathbf{I}_v \neq \mathbf{A}_v(\breve{\boldsymbol{\theta}})\mathrm{diag}(\breve{\mathbf{p}})\mathbf{A}_v^H(\breve{\boldsymbol{\theta}}) + \breve{\sigma}^2 \mathbf{I}_v, \quad (72)$$

for any arbitrary values of $\mathbf{p} \in \mathbb{R}_{>0}^{K \times 1}$, $\breve{\mathbf{p}} \in \mathbb{R}_{>0}^{K \times 1}$, σ^2 and $\breve{\sigma}^2$. From [23, Eq. (113)], vectorizing both sides of (72) leads to

$$\mathbf{T}'\mathbf{A}_\vartheta(\boldsymbol{\theta})\mathbf{p} + \sigma^2 \mathbf{T}'\mathbf{e}' \neq \mathbf{T}'\mathbf{A}_\vartheta(\breve{\boldsymbol{\theta}})\breve{\mathbf{p}} + \breve{\sigma}^2 \mathbf{T}'\mathbf{e}', \quad (73)$$

where $\mathbf{A}_\vartheta(\boldsymbol{\theta}) \in \mathbb{C}^{(2v-1) \times K}$ denotes the steering matrix corresponding to the contiguous ULA segment of the difference co-array, $\mathbf{T}' \in \{0,1\}^{v^2 \times (2v-1)}$ is a selection matrix defined in [23, Eq. (114)] and $\mathbf{e}' \in \{0,1\}^{(2v-1) \times 1}$ is a column vector with $[\mathbf{e}']_i = \delta[i-v]$. Considering $\mathbf{T}'$ is full-column rank [23], multiplying both sides of (73) by $\mathbf{T}'^\dagger$ and then moving all the terms to one side of the equation yields

$$\mathbf{A}_\vartheta(\boldsymbol{\theta})\mathbf{p} - \mathbf{A}_\vartheta(\breve{\boldsymbol{\theta}})\breve{\mathbf{p}} + (\sigma^2 - \breve{\sigma}^2)\mathbf{e}' \neq \mathbf{0}. \quad (74)$$

It follows from $\breve{\boldsymbol{\theta}} \neq \boldsymbol{\theta}_0$ that $\breve{\boldsymbol{\theta}}$ could differs from $\boldsymbol{\theta}_0$ at q DoAs for some integer $q \in [1, K]$. Noting this fact, (74) is simplified to

$$\begin{bmatrix} \mathbf{A}_\vartheta(\boldsymbol{\theta}) & \mathbf{A}_\vartheta(\ddot{\boldsymbol{\theta}}) & \mathbf{e}' \end{bmatrix} \begin{bmatrix} \mathbf{p} - \breve{\mathbf{p}} \odot \boldsymbol{\varepsilon} \\ -\ddot{\mathbf{p}} \\ \sigma^2 - \breve{\sigma}^2 \end{bmatrix} \neq \mathbf{0}, \quad (75)$$

where $\ddot{\boldsymbol{\theta}} \in [-\pi, \pi]^{q \times 1}$ consists of those elements of $\breve{\boldsymbol{\theta}}$ which do not intersect with those in $\boldsymbol{\theta}$, $\ddot{\mathbf{p}} \in \mathbb{R}_{>0}^{q \times 1}$ contains those elements of $\breve{\mathbf{p}}$ corresponding to $\ddot{\boldsymbol{\theta}}$ and

$$[\boldsymbol{\varepsilon}]_i = \begin{cases} 1, & [\boldsymbol{\theta}]_i = [\breve{\boldsymbol{\theta}}]_i, \\ 0, & \text{otherwise.} \end{cases} \quad (76)$$

Considering that $\begin{bmatrix} \mathbf{A}_\vartheta(\boldsymbol{\theta}) & \mathbf{A}_\vartheta(\ddot{\boldsymbol{\theta}}) & \mathbf{e}' \end{bmatrix} \in \mathbb{C}^{(2v-1) \times (2K+1)}$ is a sub-matrix of $\begin{bmatrix} \mathbf{A}_d(\boldsymbol{\theta}) & \mathbf{A}_d(\ddot{\boldsymbol{\theta}}) & \mathbf{e} \end{bmatrix} \in \mathbb{C}^{(2D-1) \times (2K+1)}$, obtained from $2v-1$ rows of $\begin{bmatrix} \mathbf{A}_d(\boldsymbol{\theta}) & \mathbf{A}_d(\ddot{\boldsymbol{\theta}}) & \mathbf{e} \end{bmatrix}$, it follows from (75) that

$$\begin{bmatrix} \mathbf{A}_d(\boldsymbol{\theta}) & \mathbf{A}_d(\ddot{\boldsymbol{\theta}}) & \mathbf{e} \end{bmatrix} \begin{bmatrix} \mathbf{p} - \breve{\mathbf{p}} \odot \boldsymbol{\varepsilon} \\ -\ddot{\mathbf{p}} \\ \sigma^2 - \breve{\sigma}^2 \end{bmatrix} \neq \mathbf{0}, \quad (77)$$

$$\Rightarrow \mathbf{A}_d(\boldsymbol{\theta})\mathbf{p} - \mathbf{A}_d(\breve{\boldsymbol{\theta}})\breve{\mathbf{p}} + (\sigma^2 - \breve{\sigma}^2)\mathbf{e} \neq \mathbf{0}. \quad (78)$$

Multiplying (78) by $\mathbf{J}$ and exploiting (3) and (4), after some algebraic manipulations, we obtain

$$\mathrm{vec}(\mathbf{A}(\boldsymbol{\theta})\mathrm{diag}(\mathbf{p})\mathbf{A}^H(\boldsymbol{\theta}) + \sigma^2 \mathbf{I}_M)$$
$$\neq \mathrm{vec}(\mathbf{A}(\breve{\boldsymbol{\theta}})\mathrm{diag}(\breve{\mathbf{p}})\mathbf{A}^H(\breve{\boldsymbol{\theta}}) + \breve{\sigma}^2 \mathbf{I}_M), \quad (79)$$

which in turn implies that

$$\mathbf{A}(\boldsymbol{\theta})\mathrm{diag}(\mathbf{p})\mathbf{A}^H(\boldsymbol{\theta}) + \sigma^2 \mathbf{I}_M \neq \mathbf{A}(\breve{\boldsymbol{\theta}})\mathrm{diag}(\breve{\mathbf{p}})\mathbf{A}^H(\breve{\boldsymbol{\theta}}) + \breve{\sigma}^2 \mathbf{I}_M, \quad (80)$$

for all $\boldsymbol{\theta} \neq \breve{\boldsymbol{\theta}} \in [-\pi/2, \pi/2]^{K \times 1}$, $\mathbf{p} \in \mathbb{R}_{>0}^{K \times 1}$, $\breve{\mathbf{p}} \in \mathbb{R}_{>0}^{K \times 1}$, σ^2 and $\breve{\sigma}^2$. Considering $\mathbf{y}(0), \mathbf{y}(1), \cdots, \mathbf{y}(N-1)$ are independent and identically distributed with $\mathbf{y}(t) \sim \mathcal{CN}(\mathbf{0}, \mathbf{R})$, it follows from (80) that $f(\mathbf{Y} \mid \boldsymbol{\theta}_0, \mathbf{p}, \sigma^2) \neq f(\mathbf{Y} \mid \breve{\boldsymbol{\theta}}, \breve{\mathbf{p}}, \breve{\sigma}^2)$ for any arbitrary values of $\boldsymbol{\theta} \neq \breve{\boldsymbol{\theta}} \in [-\pi/2, \pi/2]^{K \times 1}$, $\mathbf{p} \in \mathbb{R}_{>0}^{K \times 1}$, $\breve{\mathbf{p}} \in \mathbb{R}_{>0}^{K \times 1}$, σ^2 and $\breve{\sigma}^2$ if $K \leq v-1$. Now, from Theorem 1, we conclude that $f(\mathbf{X} \mid \boldsymbol{\theta}_0, \mathbf{p}, \sigma^2) \neq f(\mathbf{X} \mid \breve{\boldsymbol{\theta}}, \breve{\mathbf{p}}, \breve{\sigma}^2)$. This completes the proof of **S1**.

We now prove **S2**. We know from Lemma 1 that the FIM is singular for any value of $\boldsymbol{\theta} \in [-\pi/2, \pi/2]^{K \times 1}$ if $K \geq D$. This means that the problem is not even locally indentifiable at any $\boldsymbol{\theta}$ [67]. Since the local identifiablity is a necessary condition for the identifiablity any particular point, the problem is not identifiable for any $\boldsymbol{\theta}$.

APPENDIX C
PROOF OF LEMMA 1

Let $\mathbf{R}_{\mathbf{x}}^r$ and $\overline{\mathbf{R}}^r$ denote the equivalent real representation for $\mathbf{R}_{\mathbf{x}}$ and $\overline{\mathbf{R}}$, respectively, given as

$$\mathbf{R}_{\mathbf{x}}^r = \begin{bmatrix} \Re\{\mathbf{R}_{\mathbf{x}}\} & -\Im\{\mathbf{R}_{\mathbf{x}}\} \\ \Im\{\mathbf{R}_{\mathbf{x}}\} & \Re\{\mathbf{R}_{\mathbf{x}}\} \end{bmatrix}, \quad \overline{\mathbf{R}}^r = \begin{bmatrix} \Re\{\overline{\mathbf{R}}\} & -\Im\{\overline{\mathbf{R}}\} \\ \Im\{\overline{\mathbf{R}}\} & \Re\{\overline{\mathbf{R}}\}. \end{bmatrix} \quad (81)$$

Making use of (7) and Taylor expansion of arcsine function, we have

$$\mathbf{R}_{\mathbf{x}}^r = \frac{2}{\pi}\arcsin(\overline{\mathbf{R}}^r) = \overline{\mathbf{R}}^r + \frac{1}{6}\overline{\mathbf{R}}^r \odot \overline{\mathbf{R}}^r \odot \overline{\mathbf{R}}^r$$
$$+ \frac{3}{40}\overline{\mathbf{R}}^r \odot \overline{\mathbf{R}}^r \odot \overline{\mathbf{R}}^r \odot \overline{\mathbf{R}}^r \odot \overline{\mathbf{R}}^r + \cdots$$
$$= \sum_{n=0}^{\infty} \frac{(2n)!}{(2^n n!)^2 (2n+1)} \underbrace{\overline{\mathbf{R}}^r \odot \overline{\mathbf{R}}^r \odot \cdots \odot \overline{\mathbf{R}}^r}_{2n+1 \text{ times}}. \quad (82)$$

It is clear from (8) that $\overline{\mathbf{R}}$ is positive definite, and so is $\overline{\mathbf{R}}^r$. Further, it follows from the Schur product theorem [68, Theorem 3.1], which establishes that the Hadamard product of two positive-definite matrices is also a positive-definite matrix, that the $2n+1$ times Hadamard products of $\overline{\mathbf{R}}^r$ by itself is also positive definite for any integer $n \in [0, \infty)$. Hence, it follows from (82) that $\mathbf{R}_{\mathbf{x}}^r$ is obtained from a weighted sum of positive definite matrices, and thus it is positive definite. Evidently, $\mathbf{R}_{\mathbf{x}}$ is also positive definite. This in turn indicates non-singularity of $(\mathbf{R}_{\mathbf{x}}^{-T} \otimes \mathbf{R}_{\mathbf{x}}^{-1})$. Hence, since $\mathbf{J}$ is also full-column rank [13], we easily conclude that $\mathbf{J}^H(\mathbf{R}_{\mathbf{x}}^{-T} \otimes \mathbf{R}_{\mathbf{x}}^{-1})\mathbf{J}$ is full rank. This implies that $\mathcal{I}_w(\boldsymbol{\varrho})$ is non-singular if and only if $\begin{bmatrix} \mathbf{G} & \mathbf{V} \end{bmatrix} \in \mathbb{R}^{(2D-1) \times 2K}$ is full-column rank. In other words, $\mathcal{I}_w(\boldsymbol{\varrho})$ is non-singular if and only if

$$\begin{bmatrix} \mathbf{G} & \mathbf{V} \end{bmatrix} \begin{bmatrix} \mathbf{c}_1 \\ \mathbf{c}_2 \end{bmatrix} \neq \mathbf{0}, \quad (83)$$

for any arbitrary non-zero $\mathbf{c} = [\mathbf{c}_1^T, \mathbf{c}_2^T]^T \in \mathbb{C}^{2K \times 1}$. Inserting (16) and (17) into (83) leads to

$$\begin{bmatrix} \boldsymbol{\Delta} & F \end{bmatrix} \begin{bmatrix} \widetilde{\mathbf{c}}_1 \\ \mathbf{c}_2 \end{bmatrix} \neq \mathbf{0}, \quad (84)$$

where $\widetilde{\mathbf{c}}_1 = j\pi\boldsymbol{\Phi}(\boldsymbol{\theta})\mathrm{diag}(\overline{\mathbf{p}})\mathbf{c}_1$. This completes the proof.

APPENDIX D
PROOF OF THEOREM 3

We know from Appendix C that $\mathbf{M} = \mathbf{J}^H(\mathbf{R_x}^{-T} \otimes \mathbf{R_x}^{-1})\mathbf{J}$ is positive-definite. Hence, (15) can be rewritten as

$$\mathcal{I}_w(\boldsymbol{\varrho}) = N \begin{bmatrix} \mathbf{G}^H\mathbf{M}^{\frac{1}{2}} \\ \mathbf{V}^H\mathbf{M}^{\frac{1}{2}} \end{bmatrix} \begin{bmatrix} \mathbf{M}^{\frac{1}{2}}\mathbf{G} & \mathbf{M}^{\frac{1}{2}}\mathbf{V} \end{bmatrix}$$

$$= N \begin{bmatrix} \mathbf{G}^H\mathbf{M}\mathbf{G} & \mathbf{G}^H\mathbf{M}\mathbf{V} \\ \mathbf{V}^H\mathbf{M}\mathbf{G} & \mathbf{V}^H\mathbf{M}\mathbf{V} \end{bmatrix}. \tag{85}$$

The $CRB_w(\boldsymbol{\theta})$ is then obtained by block-wise inversion as follows:

$$CRB_w(\boldsymbol{\theta}) = \frac{1}{N}\left(\mathbf{G}^H\mathbf{M}\mathbf{G} - \mathbf{G}^H\mathbf{M}\mathbf{V}\left(\mathbf{V}^H\mathbf{M}\mathbf{V}\right)^{-1}\mathbf{V}^H\mathbf{M}\mathbf{G}\right)^{-1}$$

$$= \frac{1}{N}\left(\mathbf{G}^H\mathbf{M}^{\frac{1}{2}}\Pi^\perp_{\mathbf{M}^{\frac{1}{2}}\mathbf{V}}\mathbf{M}^{\frac{1}{2}}\mathbf{G}\right)^{-1}. \tag{86}$$

The facts that $\mathbf{G} = j\pi\mathrm{diag}(\mathbf{d})\boldsymbol{\Omega}\boldsymbol{\Phi}(\boldsymbol{\theta})\mathrm{diag}(\overline{\mathbf{p}})$ and $\mathbf{R_x} = \frac{2}{\pi}\mathrm{arcsine}(\overline{\mathbf{R}})$ will lead to (20). In addition, It follows from $\mathcal{I}(\varrho) \succeq \mathcal{I}_w(\boldsymbol{\varrho})$ that $CRB(\boldsymbol{\theta}) \preceq CRB_w(\boldsymbol{\theta})$.

APPENDIX E
PROOF OF THEOREM 4

Recalling $\overline{p}_k = \frac{p_k}{\sigma^2 + \sum_{k=1}^K p_k}$ and assuming that all sources have equal power p, we have

$$\lim_{SNR\to\infty} \overline{p}_k = \lim_{SNR\to\infty} \frac{SNR}{K \times SNR + 1} = \frac{1}{K}. \tag{87}$$

Making use of (87), it can be readily shown that

$$\lim_{SNR\to\infty} \overline{\mathbf{R}} = \frac{1}{K}\mathbf{A}(\boldsymbol{\theta})\mathbf{A}^H(\boldsymbol{\theta}) + (1 - \frac{1}{K})\mathbf{I}_M. \tag{88}$$

The above equation implies that $\lim_{SNR\to\infty}\overline{\mathbf{R}}$ is a positive-definite matrix independent of the SNR. Further, it follows from (87) that

$$\lim_{SNR\to\infty}\mathrm{diag}(\overline{\mathbf{p}}) = \frac{1}{K}\mathbf{I}_K, \tag{89}$$

$$\lim_{SNR\to\infty}\mathbf{h} = \left[\frac{1}{\sqrt{1 - \frac{|\Re\{\sum_{k=1}^K e^{-j\pi\sin\theta_k\ell_{D-1}}\}|^2}{K^2}}}, \cdots, 0,\right.$$

$$\left.\cdots, \frac{1}{\sqrt{1 - \frac{|\Re\{\sum_{k=1}^K e^{j\pi\sin\theta_k\ell_{D-1}}\}|^2}{K^2}}}\right]^T, \tag{90}$$

$$\lim_{SNR\to\infty}\overline{\mathbf{h}} = \left[\frac{1}{\sqrt{1 - \frac{|\Im\{\sum_{k=1}^K e^{-j\pi\sin\theta_k\ell_{D-1}}\}|^2}{K^2}}}, \cdots, 0,\right.$$

$$\left.\cdots, \frac{1}{\sqrt{1 - \frac{|\Im\{\sum_{k=1}^K e^{j\pi\sin\theta_k\ell_{D-1}}\}|^2}{K^2}}}\right]^T. \tag{91}$$

Substituting (89), (90) and (91) back into (16) and (17) indicates that $\lim_{SNR\to\infty}\begin{bmatrix}\mathbf{G} & \mathbf{V}\end{bmatrix}$ is a full-column rank matrix independent of the SNR. Hence, recalling (15), we can conclude that $\lim_{SNR\to\infty}\mathcal{I}_w(\varrho)$ is positive-definite and independent of the SNR. This in turn implies that $\lim_{SNR\to\infty}CRB_w(\boldsymbol{\theta})$, which is Schur complement of $\lim_{SNR\to\infty}\mathcal{I}_w(\varrho)$, is also positive-definite and independent of the SNR. This completes the proof.

APPENDIX F
PROOF OF LEMMA 2

We start with showing that $\boldsymbol{\Psi}$ is full rank. Making use of relations $\det(\begin{bmatrix}\mathbf{C}_1 & \mathbf{C}_2 \\ \mathbf{C}_3 & \mathbf{C}_4\end{bmatrix}) = \det(\mathbf{C}_1)\det(\mathbf{C}_4 - \mathbf{C}_3\mathbf{C}_1^{-1}\mathbf{C}_2)$, we obtain

$$\det(\boldsymbol{\Psi}) = \det(\mathbf{I}_{D-1})\det(2j\mathbf{I}_{D-1}) = (2j)^{D-1} \neq 0, \tag{92}$$

which implies full rankness of $\boldsymbol{\Psi}$.

Next, we proceed with proving that $\overline{\mathbf{J}}$ is full rank. Let $\ddot{\mathbf{J}}$ denote the matrix obtained after removing the D-th column from $\mathbf{J}$. $\ddot{\mathbf{J}}$ is full column rank since its columns are a sub-set of the columns of the full-column-rank matrix $\mathbf{J}$ [13]. Further, for $1 \leq i \leq M$, it is readily confirmed that the $((i-1)M+1)$-th row of $\mathrm{vec}(\mathbf{L}_n)$ as well as $\mathrm{vec}(\mathbf{L}_n^T)$ equals the i-th diagonal element of $\mathbf{L}_n$, which is obviously zero for $n \neq 0$ according to the definition given after (5). Given (5), this in turn implies that the rows of $\ddot{\mathbf{J}}$ with indices $(i-1)M+1$, for all $1 \leq i \leq M$, are zero vectors. As a result, the matrix obtained by removing these rows from $\ddot{\mathbf{J}}$, i.e., $\overline{\mathbf{J}}$, has the same column rank as $\ddot{\mathbf{J}}$. This completes the proof.

Finally, we show that $\mathbf{F}$ is full rank. It follows from the fact that $\overline{\mathbf{e}}_p^T\overline{\mathbf{e}}_q = 0$ for $p \neq q$ that

$$(\overline{\mathbf{e}}_i^T \otimes \overline{\mathbf{e}}_j^T \pm \overline{\mathbf{e}}_j^T \otimes \overline{\mathbf{e}}_i^T)(\overline{\mathbf{e}}_p \otimes \overline{\mathbf{e}}_q \pm \overline{\mathbf{e}}_q \otimes \overline{\mathbf{e}}_p) = \overline{\mathbf{e}}_i^T\overline{\mathbf{e}}_p \otimes \overline{\mathbf{e}}_j^T\overline{\mathbf{e}}_q$$

$$\pm \overline{\mathbf{e}}_i^T\overline{\mathbf{e}}_q \otimes \overline{\mathbf{e}}_j^T\overline{\mathbf{e}}_p \pm \overline{\mathbf{e}}_j^T\overline{\mathbf{e}}_p \otimes \overline{\mathbf{e}}_i^T\overline{\mathbf{e}}_q \pm \overline{\mathbf{e}}_j^T\overline{\mathbf{e}}_q \otimes \overline{\mathbf{e}}_i^T\overline{\mathbf{e}}_p = 0. \tag{93}$$

for $1 \leq i < j \leq M$ and $1 \leq p < q \leq M$ when either p or q differs from i and j. In addition, in case $i = p$ and $j = q$, we have

$$(\overline{\mathbf{e}}_i^T \otimes \overline{\mathbf{e}}_j^T + \overline{\mathbf{e}}_j^T \otimes \overline{\mathbf{e}}_i^T)(\overline{\mathbf{e}}_i \otimes \overline{\mathbf{e}}_j - \overline{\mathbf{e}}_j \otimes \overline{\mathbf{e}}_i) = \overline{\mathbf{e}}_i^T\overline{\mathbf{e}}_i \otimes \overline{\mathbf{e}}_j^T\overline{\mathbf{e}}_j$$

$$- \overline{\mathbf{e}}_i^T\overline{\mathbf{e}}_j \otimes \overline{\mathbf{e}}_j^T\overline{\mathbf{e}}_i + \overline{\mathbf{e}}_j^T\overline{\mathbf{e}}_i \otimes \overline{\mathbf{e}}_i^T\overline{\mathbf{e}}_j - \overline{\mathbf{e}}_j^T\overline{\mathbf{e}}_j \otimes \overline{\mathbf{e}}_i^T\overline{\mathbf{e}}_i = 0. \tag{94}$$

It is also observed that, for $1 \leq i \leq M$ and $1 \leq p < q \leq M$, the $((i-1)M+1)$-th element of $\overline{\mathbf{e}}_p^T \otimes \overline{\mathbf{e}}_q^T \pm \overline{\mathbf{e}}_p^T \otimes \overline{\mathbf{e}}_q^T$ is equal to the i-th diagonal element of $\overline{\mathbf{e}}_p\overline{\mathbf{e}}_q^T \pm \overline{\mathbf{e}}_p\overline{\mathbf{e}}_q^T$, which is obviously zero for $p \neq q$. Consequently, the row vectors obtained by removing the elements with indices $(i-1)M+1$ for all $1 \leq i \leq M$ from $\overline{\mathbf{e}}_p^T \otimes \overline{\mathbf{e}}_q^T + \overline{\mathbf{e}}_q^T \otimes \overline{\mathbf{e}}_p^T$ and $\overline{\mathbf{e}}_p^T \otimes \overline{\mathbf{e}}_q^T - \overline{\mathbf{e}}_q^T \otimes \overline{\mathbf{e}}_p^T$ will be still orthogonal with each other. Hence, it is deduced that the square matrix $\mathbf{F}$ has orthogonal rows, thereby being full rank.

APPENDIX G
PROOF OF LEMMA 3

Let define $E(\theta) = \mathbf{a}_v^H(\theta)\widehat{\mathbf{U}}_n\widehat{\mathbf{U}}_n^H\mathbf{a}_v(\theta)$ and $\breve{E}(\theta) = \mathbf{a}_v^H(\theta)\mathbf{U}_n\mathbf{U}_n^H\mathbf{a}_v(\theta)$ where $\widehat{\mathbf{U}}_n$ and $\mathbf{U}_n$ consist of, respectively, the eigenvectors of $\widehat{\overline{\mathbf{R}}}_v$ and $\mathbf{A}_v(\boldsymbol{\theta})\mathrm{diag}(\overline{\mathbf{p}})\mathbf{A}_v^H(\boldsymbol{\theta}) + \overline{\sigma}^2\mathbf{I}_v$ corresponding to their $v - K$ smallest eigenvalues with $K \leq v - 1$. We know that the elements of $\widehat{\boldsymbol{\theta}}$ are equal to the minimizers of $E(\theta)$. Defining $E_n = \sup_\theta|E(\theta) - \breve{E}(\theta)|$, we have

$$E_n = \sup_\theta\left|\mathbf{a}_v^H(\boldsymbol{\theta})(\widehat{\mathbf{U}}_n\widehat{\mathbf{U}}_n^H - \mathbf{U}_n\mathbf{U}_n^H)\mathbf{a}_v(\boldsymbol{\theta})\right|$$

$$= \sup_\theta\left|\left(\mathbf{a}_v^T(\boldsymbol{\theta}) \otimes \mathbf{a}_v^H(\boldsymbol{\theta})\right)\mathrm{vec}(\widehat{\mathbf{U}}_n\widehat{\mathbf{U}}_n^H - \mathbf{U}_n\mathbf{U}_n^H)\right|$$

$$\leq \|\mathbf{a}_v^T(\boldsymbol{\theta}) \otimes \mathbf{a}_v^H(\boldsymbol{\theta})\|_2\|\mathrm{vec}(\widehat{\mathbf{U}}_n\widehat{\mathbf{U}}_n^H - \mathbf{U}_n\mathbf{U}_n^H)\|_2$$

$$= v^2\|\mathrm{vec}(\widehat{\mathbf{U}}_n\widehat{\mathbf{U}}_n^H - \mathbf{U}_n\mathbf{U}_n^H)\|_2. \tag{95}$$

It follows from (47) that $\lim_{N\to\infty} \widehat{\mathbf{U}}_n \widehat{\mathbf{U}}_n^H = \mathbf{U}_n \mathbf{U}_n^H$. Hence, $E_n \to 0$ as $N \to \infty$. This implies that $E(\theta)$ converges uniformly to $\breve{E}(\theta)$ as $N \to \infty$. Thus, the minimizers of $E(\theta)$, i.e., the elements of $\widehat{\boldsymbol{\theta}}$, converge to the minimizers of $\breve{E}(\theta)$, i.e., $\theta_1, \theta_2, \cdots, \theta_K$, as $N \to \infty$. This completes the proof.

APPENDIX H
PROOF OF THEOREM 5

Considering the consistency of $\widehat{\boldsymbol{\theta}}$ and following the same arguments as in [21, App. B], for sufficiently large N, the asymptotic estimation error expression for EOCAB-MUSIC is given by

$$\hat{\theta}_k - \theta_k = -\frac{\Re\{\mathbf{z}_k^T \mathbf{T} \mathbf{J}^\dagger \Delta\bar{\mathbf{r}}\}}{\pi \bar{p}_k q_k \cos(\theta_k)}, \tag{96}$$

where $\Delta\bar{\mathbf{r}} = \widehat{\bar{\mathbf{r}}} - \bar{\mathbf{r}}$ and $\mathbf{T} = \begin{bmatrix} \mathbf{T}_v^T & \mathbf{T}_{v-1}^T & \cdots & \mathbf{T}_1^T \end{bmatrix}^T \in \mathbb{C}^{v^2 \times (2D-1)}$. From (96), the covariance of the asymptotic distribution (as $N \to \infty$) of the DoA estimation errors is given by

$$\begin{aligned}
\mathcal{E}_{\theta_{k_1},\theta_{k_2}} &= \mathbb{E}\{(\hat{\theta}_{k_1} - \theta_{k_1})(\hat{\theta}_{k_2} - \theta_{k_2})\} \\
&= \frac{\mathbb{E}\{\Re\{\mathbf{z}_{k_1}^T \mathbf{T} \mathbf{J}^\dagger \Delta\bar{\mathbf{r}}\}\Re\{\mathbf{z}_{k_2}^T \mathbf{T} \mathbf{J}^\dagger \Delta\bar{\mathbf{r}}\}\}}{\pi^2 \bar{p}_{k_1} \bar{p}_{k_2} q_{k_1} q_{k_2} \cos(\theta_{k_1}) \cos(\theta_{k_2})}.
\end{aligned} \tag{97}$$

Making use of the identity $\Re\{\mathbf{c}_1^H \mathbf{c}_2\}\Re\{\mathbf{c}_3^H \mathbf{c}_2\} = \frac{1}{2}\Re\{\mathbf{c}_1^H \mathbf{c}_2 \mathbf{c}_2^H \mathbf{c}_3 + \mathbf{c}_1^H \mathbf{c}_2 \mathbf{c}_2^T \mathbf{c}_3^*\}$, we obtain

$$\begin{aligned}
&\mathbb{E}\left\{\Re\{\mathbf{z}_{k_1}^T \mathbf{T} \mathbf{J}^\dagger \Delta\bar{\mathbf{r}}\}\Re\{\mathbf{z}_{k_2}^T \mathbf{T} \mathbf{J}^\dagger \Delta\bar{\mathbf{r}}\}\right\} = \\
&\frac{1}{2}\mathbb{E}\{\Re\{\mathbf{z}_{k_1}^T \mathbf{T} \mathbf{J}^\dagger \Delta\bar{\mathbf{r}}\Delta\bar{\mathbf{r}}^H \mathbf{J}^{\dagger H} \mathbf{T}^H \mathbf{z}_{k_2}^*\} \\
&+ \Re\{\mathbf{z}_{k_1}^T \mathbf{T} \mathbf{J}^\dagger \Delta\bar{\mathbf{r}}\Delta\bar{\mathbf{r}}^T \mathbf{J}^{\dagger H} \mathbf{T}^H \mathbf{z}_{k_2}\}\}.
\end{aligned} \tag{98}$$

The matrix $\mathrm{mat}_{M,M}(\mathbf{J}^{\dagger H} \mathbf{T}^H \mathbf{z}_k)$ is Hermitian [21, Lemma 6], thereby

$$\mathbf{J}^{\dagger H} \mathbf{T}^H \mathbf{z}_k^* = \mathbf{K}_M \mathbf{J}^{\dagger H} \mathbf{T}^H \mathbf{z}_k. \tag{99}$$

where $\mathbf{K}_M \in \{0,1\}^{M^2 \times M^2}$ is the commutation matrix defined as $\mathrm{vec}(\mathbf{C}^T) = \mathbf{K}_M \mathrm{vec}(\mathbf{C})$ for any arbitrary matrix $\mathbf{C}$ [69]. In addition, since $\overline{\mathbf{R}}^H = \overline{\mathbf{R}}$, we have

$$\Delta\bar{\mathbf{r}}^T = \Delta\bar{\mathbf{r}}^H \mathbf{K}_M^H. \tag{100}$$

Inserting (99) and (100) into (98) and using the fact that $\mathbf{K}_M = \mathbf{K}_M^H = \mathbf{K}_M^{-1}$, we obtain

$$\begin{aligned}
&\mathbb{E}\left\{\Re\{\mathbf{z}_{k_1}^T \mathbf{T} \mathbf{J}^\dagger \Delta\bar{\mathbf{r}}\}\Re\{\mathbf{z}_{k_2}^T \mathbf{T} \mathbf{J}^\dagger \Delta\bar{\mathbf{r}}\}\right\} \\
&= \mathbb{E}\{\Re\{\mathbf{z}_{k_1}^T \mathbf{T} \mathbf{J}^\dagger \Delta\bar{\mathbf{r}}\Delta\bar{\mathbf{r}}^H \mathbf{J}^{\dagger H} \mathbf{T}^H \mathbf{z}_{k_2}^*\}\}.
\end{aligned} \tag{101}$$

Recalling (44) and (30), we have

$$\mathbf{T}\mathbf{J}^\dagger\Delta\bar{\mathbf{r}} = \mathbf{T}\begin{bmatrix} \mathbf{0} & \mathbf{I}_{D-1} & -j\mathbf{I}_{D-1} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_{D-1} & j\mathbf{I}_{D-1} \end{bmatrix}\begin{bmatrix} 0 \\ \widehat{\boldsymbol{\phi}} - \boldsymbol{\phi} \end{bmatrix}. \tag{102}$$

Additionally, from [23, Eq. (114) and Eq. (116)], we know

$$\mathbf{T} = \tag{103}$$
$$[\mathbf{0}_{v^2 \times (D-v)}, \mathrm{vec}(\overline{\mathbf{L}}_{v-1}^T), \cdots, \mathrm{vec}(\overline{\mathbf{L}}_0), \cdots, \mathrm{vec}(\overline{\mathbf{L}}_{v-1}), \mathbf{0}_{v^2 \times (D-v)}],$$

where $[\overline{\mathbf{L}}_n]_{p,q} = \begin{cases} 1, & \text{if } p-q = n, \\ 0, & \text{otherwise.} \end{cases}$. Substituting (103) into (102) yields

$$\mathbf{T}\mathbf{J}^\dagger\Delta\bar{\mathbf{r}} = \overline{\mathbf{T}}\boldsymbol{\Psi}(\widehat{\boldsymbol{\phi}} - \boldsymbol{\phi}), \tag{104}$$

where

$$\overline{\mathbf{T}} = \tag{105}$$
$$[\mathbf{0}_{v^2 \times (D-v)}, \mathrm{vec}(\overline{\mathbf{L}}_{v-1}^T), \cdots, \mathrm{vec}(\overline{\mathbf{L}}_{-1}^T), \mathrm{vec}(\overline{\mathbf{L}}_1), \cdots, \mathrm{vec}(\overline{\mathbf{L}}_{v-1}), \mathbf{0}_{v^2 \times (D-v)}].$$

Inserting (104) into (101) gives

$$\begin{aligned}
&\mathbb{E}\left\{\Re\{\mathbf{z}_{k_1}^T \mathbf{T} \mathbf{J}^\dagger \Delta\bar{\mathbf{r}}\}\Re\{\mathbf{z}_{k_2}^T \mathbf{T} \mathbf{J}^\dagger \Delta\bar{\mathbf{r}}\}\right\} \\
&= \Re\{\mathbf{z}_{k_1}^T \overline{\mathbf{T}}\mathbb{E}\{\boldsymbol{\Psi}(\widehat{\boldsymbol{\phi}} - \boldsymbol{\phi})(\widehat{\boldsymbol{\phi}} - \boldsymbol{\phi})^H \boldsymbol{\Psi}^H\}\overline{\mathbf{T}}^H \mathbf{z}_{k_2}^*\}.
\end{aligned} \tag{106}$$

As a result, for sufficiently large N, using a first-order perturbation expansion leads to

$$\begin{aligned}
&\mathbb{E}\{\boldsymbol{\Psi}(\widehat{\boldsymbol{\phi}} - \boldsymbol{\phi})(\widehat{\boldsymbol{\phi}} - \boldsymbol{\phi})^H \boldsymbol{\Psi}^H\} \simeq \\
&\left(\overline{\mathbf{J}}^H \mathbf{F}^H \mathrm{diag}(\mathbf{b})\mathbf{F}^{-H}\boldsymbol{\Sigma}^{-1}\mathbf{F}^{-1}\mathrm{diag}(\mathbf{b})\mathbf{F}\overline{\mathbf{J}}\right)^{-1} \\
&\times \overline{\mathbf{J}}^H \mathbf{F}^H \mathrm{diag}(\mathbf{b})\mathbf{F}^{-H}\boldsymbol{\Sigma}^{-1}\mathbf{F}^{-1}\mathrm{diag}(\mathbf{b})\mathbf{F}\mathbb{E}\{\widetilde{\widetilde{\mathbf{r}}}\widetilde{\widetilde{\mathbf{r}}}^H\} \\
&\times \mathbf{F}^H \mathrm{diag}(\mathbf{b})\mathbf{F}^{-H}\boldsymbol{\Sigma}^{-1}\mathbf{F}^{-1}\mathrm{diag}(\mathbf{b})\mathbf{F}\overline{\mathbf{J}} \\
&\times \left(\overline{\mathbf{J}}^H \mathbf{F}^H \mathrm{diag}(\mathbf{b})\mathbf{F}^{-H}\boldsymbol{\Sigma}^{-1}\mathbf{F}^{-1}\mathrm{diag}(\mathbf{b})\mathbf{F}\overline{\mathbf{J}}\right)^{-1} - \boldsymbol{\Psi}\boldsymbol{\phi}\boldsymbol{\phi}^H \boldsymbol{\Psi}^H,
\end{aligned} \tag{107}$$

where $\boldsymbol{\Sigma} = \boldsymbol{\Sigma}(\boldsymbol{\gamma})$ given in Appendix K (kindly refer to the supplementary document) and $[\mathbf{b}]_n = \frac{1}{\sqrt{1-|[\boldsymbol{\gamma}]_n|^2}}$ for $1 \leq n \leq M^2 - M$. It remains to compute $\mathbb{E}\{\widetilde{\widetilde{\mathbf{r}}}\widetilde{\widetilde{\mathbf{r}}}^H\}$. Making use of the relation $\widetilde{\widetilde{\mathbf{r}}} = \mathrm{sine}(\frac{\pi}{2}\widehat{\widetilde{\mathbf{r}}}_{\mathbf{x}})$, we obtain

$$\begin{aligned}
&\mathbb{E}\{[\widetilde{\widetilde{\mathbf{r}}}]_p[\widetilde{\widetilde{\mathbf{r}}}]_q^*\} = \\
&\frac{1}{4}\mathbb{E}\Bigg\{e^{\frac{j\pi\left(\Re\{[\widehat{\widetilde{\mathbf{r}}}_{\mathbf{x}}]_p\}-\Re\{[\widehat{\widetilde{\mathbf{r}}}_{\mathbf{x}}]_q\}\right)}{2}} + e^{-\frac{j\pi\left(\Re\{[\widehat{\widetilde{\mathbf{r}}}_{\mathbf{x}}]_p\}-\Re\{[\widehat{\widetilde{\mathbf{r}}}_{\mathbf{x}}]_q\}\right)}{2}} \\
&- e^{\frac{j\pi\left(\Re\{[\widehat{\widetilde{\mathbf{r}}}_{\mathbf{x}}]_p\}+\Re\{[\widehat{\widetilde{\mathbf{r}}}_{\mathbf{x}}]_q\}\right)}{2}} - e^{-\frac{j\pi\left(\Re\{[\widehat{\widetilde{\mathbf{r}}}_{\mathbf{x}}]_p\}-\Re\{[\widehat{\widetilde{\mathbf{r}}}_{\mathbf{x}}]_q\}\right)}{2}} \\
&+ e^{\frac{j\pi\left(\Im\{[\widehat{\widetilde{\mathbf{r}}}_{\mathbf{x}}]_p\}-\Im\{[\widehat{\widetilde{\mathbf{r}}}_{\mathbf{x}}]_q\}\right)}{2}} + e^{-\frac{j\pi\left(\Im\{[\widehat{\widetilde{\mathbf{r}}}_{\mathbf{x}}]_p\}-\Im\{[\widehat{\widetilde{\mathbf{r}}}_{\mathbf{x}}]_q\}\right)}{2}} \\
&- e^{\frac{j\pi\left(\Im\{[\widehat{\widetilde{\mathbf{r}}}_{\mathbf{x}}]_p\}+\Im\{[\widehat{\widetilde{\mathbf{r}}}_{\mathbf{x}}]_q\}\right)}{2}} - e^{-\frac{j\pi\left(\Im\{[\widehat{\widetilde{\mathbf{r}}}_{\mathbf{x}}]_p\}+\Im\{[\widehat{\widetilde{\mathbf{r}}}_{\mathbf{x}}]_q\}\right)}{2}}\Bigg\} \\
&+ \frac{j}{4}\mathbb{E}\Bigg\{e^{\frac{j\pi\left(\Im\{[\widehat{\widetilde{\mathbf{r}}}_{\mathbf{x}}]_p\}-\Re\{[\widehat{\widetilde{\mathbf{r}}}_{\mathbf{x}}]_q\}\right)}{2}} + e^{\frac{-j\pi\left(\Im\{[\widehat{\widetilde{\mathbf{r}}}_{\mathbf{x}}]_p\}-\Re\{[\widehat{\widetilde{\mathbf{r}}}_{\mathbf{x}}]_q\}\right)}{2}} \\
&- e^{\frac{j\pi\left(\Im\{[\widehat{\widetilde{\mathbf{r}}}_{\mathbf{x}}]_p\}+\Re\{[\widehat{\widetilde{\mathbf{r}}}_{\mathbf{x}}]_q\}\right)}{2}} - e^{\frac{-j\pi\left(\Im\{[\widehat{\widetilde{\mathbf{r}}}_{\mathbf{x}}]_p\}+\Re\{[\widehat{\widetilde{\mathbf{r}}}_{\mathbf{x}}]_q\}\right)}{2}} \\
&- e^{\frac{j\pi\left(\Re\{[\widehat{\widetilde{\mathbf{r}}}_{\mathbf{x}}]_p\}-\Im\{[\widehat{\widetilde{\mathbf{r}}}_{\mathbf{x}}]_q\}\right)}{2}} - e^{-\frac{j\pi\left(\Re\{[\widehat{\widetilde{\mathbf{r}}}_{\mathbf{x}}]_p\}-\Im\{[\widehat{\widetilde{\mathbf{r}}}_{\mathbf{x}}]_q\}\right)}{2}} \\
&+ e^{\frac{j\pi\left(\Re\{[\widehat{\widetilde{\mathbf{r}}}_{\mathbf{x}}]_p\}+\Im\{[\widehat{\widetilde{\mathbf{r}}}_{\mathbf{x}}]_q\}\right)}{2}} + e^{-\frac{j\pi\left(\Re\{[\widehat{\widetilde{\mathbf{r}}}_{\mathbf{x}}]_p\}+\Im\{[\widehat{\widetilde{\mathbf{r}}}_{\mathbf{x}}]_q\}\right)}{2}}\Bigg\}.
\end{aligned} \tag{108}$$

Considering that $\widehat{\widetilde{\mathbf{r}}}_{\mathbf{x}} \xrightarrow{D} \mathcal{CN}(\ddot{\mathbf{r}}_{\mathbf{x}}, \frac{4}{\pi^2 N}\boldsymbol{\Sigma})$, the expectations in (108) can be computed using the characteristic function of the Gaussian distribution as follows:

$$\begin{aligned}
\mathbb{E}\{[\widetilde{\widetilde{\mathbf{r}}}]_p[\widetilde{\widetilde{\mathbf{r}}}]_q^*\} &= \frac{e^{\frac{-[\boldsymbol{\Sigma}]_{p,p}-[\boldsymbol{\Sigma}]_{q,q}}{4N}}}{2} \\
&\times \Bigg[\cos\left(\frac{\pi}{2}[\Re\{[\ddot{\mathbf{r}}_{\mathbf{x}}]_p\} - \Re\{[\ddot{\mathbf{r}}_{\mathbf{x}}]_q\}]\right)e^{\frac{\Re\{[\boldsymbol{\Sigma}]_{p,q}\}}{2N}} \\
&- \cos\left(\frac{\pi}{2}[\Re\{[\ddot{\mathbf{r}}_{\mathbf{x}}]_p\} + \Re\{[\ddot{\mathbf{r}}_{\mathbf{x}}]_q\}]\right)e^{-\frac{\Re\{[\boldsymbol{\Sigma}]_{p,q}\}}{2N}} \\
&+ \cos\left(\frac{\pi}{2}[\Im\{[\ddot{\mathbf{r}}_{\mathbf{x}}]_p\} - \Im\{[\ddot{\mathbf{r}}_{\mathbf{x}}]_q\}]\right)e^{\frac{\Re\{[\boldsymbol{\Sigma}]_{p,q}\}}{2N}} \\
&- \cos\left(\frac{\pi}{2}[\Im\{[\ddot{\mathbf{r}}_{\mathbf{x}}]_p\} + \Im\{[\ddot{\mathbf{r}}_{\mathbf{x}}]_q\}]\right)e^{-\frac{\Re\{[\boldsymbol{\Sigma}]_{p,q}\}}{2N}}
\end{aligned}$$

$$+ j \cos\left(\frac{\pi}{2}\left[\Im\{[\ddot{\mathbf{r}}_{\mathbf{x}}]_p\} - \Re\{[\ddot{\mathbf{r}}_{\mathbf{x}}]_q\}\right]\right) e^{\frac{\Im\{[\mathbf{\Sigma}]_{p,q}\}}{2N}}$$
$$- j \cos\left(\frac{\pi}{2}\left[\Im\{[\ddot{\mathbf{r}}_{\mathbf{x}}]_p\} + \Re\{[\ddot{\mathbf{r}}_{\mathbf{x}}]_q\}\right]\right) e^{-\frac{\Im\{[\mathbf{\Sigma}]_{p,q}\}}{2N}}$$
$$- j \cos\left(\frac{\pi}{2}\left[\Re\{[\ddot{\mathbf{r}}_{\mathbf{x}}]_p\} - \Im\{[\ddot{\mathbf{r}}_{\mathbf{x}}]_q\}\right]\right) e^{-\frac{\Im\{[\mathbf{\Sigma}]_{p,q}\}}{2N}}$$
$$+ j \cos\left(\frac{\pi}{2}\left[\Re\{[\ddot{\mathbf{r}}_{\mathbf{x}}]_p\} + \Im\{[\ddot{\mathbf{r}}_{\mathbf{x}}]_q\}\right]\right) e^{\frac{\Im\{[\mathbf{\Sigma}]_{p,q}\}}{2N}}\Bigg]. \tag{109}$$

Exploiting the Taylor expansion of the exponential function, (109) can be approximated for sufficiently large N as

$$\mathbb{E}\{[\widetilde{\ddot{\mathbf{r}}}]_p[\widetilde{\ddot{\mathbf{r}}}]_q^*\} \simeq \frac{1}{2} \tag{110}$$
$$\times\Bigg[\cos\left(\frac{\pi}{2}\left[\Re\{[\ddot{\mathbf{r}}_{\mathbf{x}}]_p\} - \Re\{[\ddot{\mathbf{r}}_{\mathbf{x}}]_q\}\right]\right)\left(1 + \frac{\Re\{[\mathbf{\Sigma}]_{p,q}\}}{2N}\right)$$
$$- \cos\left(\frac{\pi}{2}\left[\Re\{[\ddot{\mathbf{r}}_{\mathbf{x}}]_p\} + \Re\{[\ddot{\mathbf{r}}_{\mathbf{x}}]_q\}\right]\right)\left(1 - \frac{\Re\{[\mathbf{\Sigma}]_{p,q}\}}{2N}\right)$$
$$+ \cos\left(\frac{\pi}{2}\left[\Im\{[\ddot{\mathbf{r}}_{\mathbf{x}}]_p\} - \Im\{[\ddot{\mathbf{r}}_{\mathbf{x}}]_q\}\right]\right)\left(1 + \frac{\Re\{[\mathbf{\Sigma}]_{p,q}\}}{2N}\right)$$
$$- \cos\left(\frac{\pi}{2}\left[\Im\{[\ddot{\mathbf{r}}_{\mathbf{x}}]_p\} + \Im\{[\ddot{\mathbf{r}}_{\mathbf{x}}]_q\}\right]\right)\left(1 - \frac{\Re\{[\mathbf{\Sigma}]_{p,q}\}}{2N}\right)$$
$$+ j \cos\left(\frac{\pi}{2}\left[\Im\{[\ddot{\mathbf{r}}_{\mathbf{x}}]_p\} - \Re\{[\ddot{\mathbf{r}}_{\mathbf{x}}]_q\}\right]\right)\left(1 + \frac{\Im\{[\mathbf{\Sigma}]_{p,q}\}}{2N}\right)$$
$$- j \cos\left(\frac{\pi}{2}\left[\Im\{[\ddot{\mathbf{r}}_{\mathbf{x}}]_p\} + \Re\{[\ddot{\mathbf{r}}_{\mathbf{x}}]_q\}\right]\right)\left(1 - \frac{\Im\{[\mathbf{\Sigma}]_{p,q}\}}{2N}\right)$$
$$- j \cos\left(\frac{\pi}{2}\left[\Re\{[\ddot{\mathbf{r}}_{\mathbf{x}}]_p\} - \Im\{[\ddot{\mathbf{r}}_{\mathbf{x}}]_q\}\right]\right)\left(1 - \frac{\Im\{[\mathbf{\Sigma}]_{p,q}\}}{2N}\right)$$
$$+ j \cos\left(\frac{\pi}{2}\left[\Re\{[\ddot{\mathbf{r}}_{\mathbf{x}}]_p\} + \Im\{[\ddot{\mathbf{r}}_{\mathbf{x}}]_q\}\right]\right)\left(1 + \frac{\Im\{[\mathbf{\Sigma}]_{p,q}\}}{2N}\right)\Bigg]$$
$$= \sin(\frac{\pi}{2}\Re\{[\ddot{\mathbf{r}}_{\mathbf{x}}]_p\})\sin(\frac{\pi}{2}\Re\{[\ddot{\mathbf{r}}_{\mathbf{x}}]_p\})$$
$$+ \sin(\frac{\pi}{2}\Im\{[\ddot{\mathbf{r}}_{\mathbf{x}}]_p\})\sin(\frac{\pi}{2}\Im\{[\ddot{\mathbf{r}}_{\mathbf{x}}]_p\})$$
$$+ j \sin(\frac{\pi}{2}\Im\{[\ddot{\mathbf{r}}_{\mathbf{x}}]_p\})\sin(\frac{\pi}{2}\Re\{[\ddot{\mathbf{r}}_{\mathbf{x}}]_p\})$$
$$- j \sin(\frac{\pi}{2}\Re\{[\ddot{\mathbf{r}}_{\mathbf{x}}]_p\})\sin(\frac{\pi}{2}\Im\{[\ddot{\mathbf{r}}_{\mathbf{x}}]_p\})$$
$$+ \frac{\Re\{[\mathbf{\Sigma}]_{p,q}\}}{2N}\Bigg[\cos(\frac{\pi}{2}\Re\{[\ddot{\mathbf{r}}_{\mathbf{x}}]_p\})\cos(\frac{\pi}{2}\Re\{[\ddot{\mathbf{r}}_{\mathbf{x}}]_p\})$$
$$+ \cos(\frac{\pi}{2}\Im\{[\ddot{\mathbf{r}}_{\mathbf{x}}]_p\})\cos(\frac{\pi}{2}\Im\{[\ddot{\mathbf{r}}_{\mathbf{x}}]_p\})\Bigg]$$
$$+ j + \frac{\Im\{[\mathbf{\Sigma}]_{p,q}\}}{2N}\Bigg[\cos(\frac{\pi}{2}\Im\{[\ddot{\mathbf{r}}_{\mathbf{x}}]_p\})\cos(\frac{\pi}{2}\Re\{[\ddot{\mathbf{r}}_{\mathbf{x}}]_p\})$$
$$+ \cos(\frac{\pi}{2}\Re\{[\ddot{\mathbf{r}}_{\mathbf{x}}]_p\})\cos(\frac{\pi}{2}\Im\{[\ddot{\mathbf{r}}_{\mathbf{x}}]_p\})\Bigg]. \tag{111}$$

Consequently, it follows from (28) that

$$[\mathbf{\Gamma}]_{p,q} = \mathbb{E}\{[\widetilde{\ddot{\mathbf{r}}}]_p[\widetilde{\ddot{\mathbf{r}}}]_q^*\} \simeq [\ddot{\mathbf{r}}]_p[\ddot{\mathbf{r}}]_q^* \tag{112}$$
$$+ \frac{1}{2N}\left(\sqrt{1 - [\Re\{[\ddot{\mathbf{r}}]_p\}]^2} \times \sqrt{1 - [\Re\{[\ddot{\mathbf{r}}]_q\}]^2}\right.$$
$$\left. + \sqrt{1 - [\Im\{[\ddot{\mathbf{r}}]_p\}]^2} \times \sqrt{1 - [\Im\{[\ddot{\mathbf{r}}]_q\}]^2}\right)\Re\{[\mathbf{\Sigma}]_{p,q}\}$$
$$+ \frac{j}{2N}\left(\sqrt{1 - [\Im\{[\ddot{\mathbf{r}}]_p\}]^2} \times \sqrt{1 - [\Re\{[\ddot{\mathbf{r}}]_q\}]^2}\right.$$
$$\left. + \sqrt{1 - [\Re\{[\ddot{\mathbf{r}}]_p\}]^2} \times \sqrt{1 - [\Im\{[\ddot{\mathbf{r}}]_q\}]^2}\right)\Im\{[\mathbf{\Sigma}]_{p,q}\}.$$

Inserting (112) into (107) and making use of (31) yields

$$\mathbb{E}\{\mathbf{\Psi}(\widehat{\phi}-\phi)(\widehat{\phi}-\phi)^H\mathbf{\Psi}^H\} \simeq \tag{113}$$
$$\left(\overline{\mathbf{J}}^H\mathbf{F}^H\text{diag}(\mathbf{b})\mathbf{F}^{-H}\mathbf{\Sigma}^{-1}\mathbf{F}^{-1}\text{diag}(\mathbf{b})\mathbf{F}\overline{\mathbf{J}}\right)^{-1} \times$$
$$\overline{\mathbf{J}}^H\mathbf{F}^H\text{diag}(\mathbf{b})\mathbf{F}^{-H}\mathbf{\Sigma}^{-1}\mathbf{F}^{-1}\text{diag}(\mathbf{b})\mathbf{F}\mathbf{\Gamma}\mathbf{F}^H\text{diag}(\mathbf{b})\mathbf{F}^{-H}\mathbf{\Sigma}^{-1}\mathbf{F}^{-1}$$
$$\times \text{diag}(\mathbf{b})\mathbf{F}\overline{\mathbf{J}}\left(\overline{\mathbf{J}}^H\mathbf{F}^H\text{diag}(\mathbf{b})\mathbf{F}^{-H}\mathbf{\Sigma}^{-1}\mathbf{F}^{-1}\text{diag}(\mathbf{b})\mathbf{F}\overline{\mathbf{J}}\right)^{-1},$$

Finally, substituting (113) into (106) and considering that $\overline{p}_k = \frac{p_k}{\sigma^2+\sum_{k=1}^K p_k}$ concludes the proof of Theorem 5.

APPENDIX I
PROOF OF COROLLARY 2

OCAB-MUSIC employs $\widetilde{\mathbf{r}} = \text{vec}(\widetilde{\overline{\mathbf{R}}})$ instead of $\widehat{\overline{\mathbf{r}}}$. Hence, its asymptotic estimation error is obtained by replacing $\Delta\overline{\mathbf{r}}$ with $\widetilde{\mathbf{r}}-\overline{\mathbf{r}}$ in (96). Following the same steps from (97) to (101), the covariance of the asymptotic distribution (as $N \to \infty$) of the DoA estimation errors for OCAB-MUSIC is obtained as

$$\mathcal{E}_{\theta_{k_1},\theta_{k_2}} = \Re\{\mathbf{z}_{k_1}^T\mathbf{T}\mathbf{J}^\dagger\mathbb{E}\{(\widetilde{\mathbf{r}}-\overline{\mathbf{r}})(\widetilde{\mathbf{r}}-\overline{\mathbf{r}})^H\}\mathbf{J}^{\dagger H}\mathbf{T}^H\mathbf{z}_{k_2}^*\}. \tag{114}$$

Considering the fact that the diagonal elements of $\widetilde{\overline{\mathbf{R}}}$ and $\overline{\mathbf{R}}$ are equal to one, (114) is simplified as

$$\mathcal{E}_{\theta_{k_1},\theta_{k_2}} = \Re\{\mathbf{z}_{k_1}^T\overline{\mathbf{T}}\mathbf{J}^\dagger\left[\mathbb{E}\{\widetilde{\ddot{\mathbf{r}}}\widetilde{\ddot{\mathbf{r}}}^H\} - \ddot{\mathbf{r}}\ddot{\mathbf{r}}^H\right]\overline{\mathbf{J}}^{\dagger H}\overline{\mathbf{T}}^H\mathbf{z}_{k_2}^*\}. \tag{115}$$

Substituting $\mathbb{E}\{\widetilde{\ddot{\mathbf{r}}}\widetilde{\ddot{\mathbf{r}}}^H\}$ from (112) completes the proof.

APPENDIX J
PROOF OF THEOREM 6

To derive $\lim_{SNR\to\infty}\mathcal{E}_{\theta_k}$, we need to calculate $\lim_{SNR\to\infty}(\sigma^2 + \sum_{k'=1}^K p_{k'})^2/p_k^2$, $\mathbf{W}_\infty = \lim_{SNR\to\infty}\mathbf{W}$ and $\mathbf{\Gamma}_\infty = \lim_{SNR\to\infty}\mathbf{\Gamma}$. It is obtained from (87) that

$$\lim_{SNR\to\infty}\frac{(\sigma^2 + \sum_{k'=1}^K p_{k'})^2}{p_k^2} = \lim_{SNR\to\infty}\frac{1}{\overline{p}_k^2} = K^2. \tag{116}$$

In addition, it follows from (53) and (54) that $\mathbf{W}$ and $\mathbf{\Gamma}$ depend on SNR through $\overline{\mathbf{R}}$, γ and $\ddot{\mathbf{r}}$. Hence, for calculating $\mathbf{W}_\infty$ and $\mathbf{\Gamma}_\infty$, it is sufficient to first compute $\overline{\mathbf{R}}_\infty = \lim_{SNR\to\infty}\overline{\mathbf{R}}$, $\gamma_\infty = \lim_{SNR\to\infty}\gamma$ and $\ddot{\mathbf{r}}_\infty = \lim_{SNR\to\infty}\ddot{\mathbf{r}}$, and then insert them back into the expressions of $\mathbf{W}$ and $\mathbf{\Gamma}$ given in (53) and (54). $\overline{\mathbf{R}}_\infty$ is obtained in (88). Given $\overline{\mathbf{R}}_\infty$, $\gamma_\infty = \lim_{SNR\to\infty}\gamma$ is equal to the $(M^2 - M) \times 1$ vector containing the real and imaginary parts of the elements of $\overline{\mathbf{R}}_\infty$ above its main diagonal elements. Then, exploiting (36), we have $\ddot{\mathbf{r}}_\infty = \lim_{SNR\to\infty}\mathbf{\Psi}^{-1}\overline{\mathbf{J}}^\dagger\mathbf{F}^{-1}\gamma = \mathbf{\Psi}^{-1}\overline{\mathbf{J}}^\dagger\mathbf{F}^{-1}\gamma_\infty$. This completes the proof.

REFERENCES

[1] H. Van Trees, *Optimum Array Processing (Detection, Estimation, and Modulation Theory, Part IV)*. New York: John Wiley and Sons Inc., 2002.
[2] S. S. Haykin, J. Litva, , and T. J. Shepherd, Eds., *Radar Array Processing*. Berlin, Germany: Springer-Verlag, 1993.
[3] B. Ottersten, "Array processing for wireless communications," in *Proceedings of 8th Workshop on Statistical Signal and Array Processing*, Jun 1996, pp. 466–473.
[4] A. Paulraj, B. Ottersten, R. Roy, A. Swindlehurst, G. Xu, and T. Kailath, "subspace methods for directions-of-arrival estimation," *Handbook of Statistics*, vol. 10, pp. 693–739, 1993.
[5] F. Li, H. Liu, and R. J. Vaccaro, "Performance analysis for DoA estimation algorithms: unification, simplification, and observations," *IEEE Trans. Aerosp. Electron. Syst*, vol. 29, no. 4, pp. 1170–1184, Oct 1993.
[6] P. Stoica and A. Nehorai, "Music, maximum likelihood, and cramer-rao bound," *IEEE Trans. Acoustics, Speech, and Signal Process.*, vol. 37, no. 5, pp. 720–741, May 1989.

[7] P. Stoica and A. Nehorai, "Performance study of conditional and unconditional direction-of-arrival estimation," *IEEE Trans. Acoustics, Speech, and Signal Process.*, vol. 38, no. 10, pp. 1783–1795, Oct 1990.

[8] 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.

[9] M. Jansson, B. Goransson, and B. Ottersten, "A subspace method for direction of arrival estimation of uncorrelated emitter signals," *IEEE Trans. Signal Process.*, vol. 47, no. 4, pp. 945–956, April 1999.

[10] A. Moffet, "Minimum-redundancy linear arrays," *IEEE Trans. Antennas. Propag.*, vol. 16, no. 2, pp. 172–175, Mar 1968.

[11] P. P. Vaidyanathan and P. Pal, "Sparse sensing with co-prime samplers and arrays," *IEEE Trans. Signal Process.*, vol. 59, no. 2, pp. 573–586, Feb 2011.

[12] P. Pal and P. P. Vaidyanathan, "Nested arrays: A novel approach to array processing with enhanced degrees of freedom," *IEEE Trans. Signal Process.*, vol. 58, no. 8, pp. 4167–4181, Aug 2010.

[13] C. L. Liu and P. P. Vaidyanathan, "Cramér-Rao bounds for coprime and other sparse arrays, which find more sources than sensors," *Digit. Signal Process.*, vol. 61, pp. 43 – 61, 2017.

[14] Y. D. Zhang, M. G. Amin, and B. Himed, "Sparsity-based DoA estimation using co-prime arrays," in *2013 IEEE International Conference on Acoustics, Speech and Signal Processing*, May 2013, pp. 3967–3971.

[15] Q. Shen, W. Liu, W. Cui, and S. Wu, "Underdetermined DoA estimation under the compressive sensing framework: A review," *IEEE Access*, vol. 4, pp. 8865–8878, 2016.

[16] P. Pal and P. P. Vaidyanathan, "Pushing the limits of sparse support recovery using correlation information," *IEEE Trans. Signal Process.*, vol. 63, no. 3, pp. 711–726, Feb 2015.

[17] ——, "Correlation-aware techniques for sparse support recovery," in *IEEE Statistical Signal Processing Workshop (SSP)*, Aug 2012, pp. 53–56.

[18] Y. Chi, A. Pezeshki, L. Scharf, and R. Calderbank, "Sensitivity to basis mismatch in compressed sensing," in *IEEE International Conference on Acoustics, Speech and Signal Processing*, March 2010, pp. 3930–3933.

[19] Z. Tan and A. Nehorai, "Sparse direction of arrival estimation using co-prime arrays with off-grid targets," *IEEE Signal Process. Lett.*, vol. 21, no. 1, pp. 26–29, Jan 2014.

[20] Z. Yang, L. Xie, and C. Zhang, "A discretization-free sparse and parametric approach for linear array signal processing," *IEEE Trans. Signal Process.*, vol. 62, no. 19, pp. 4959–4973, Oct 2014.

[21] M. Wang and A. Nehorai, "Coarrays, MUSIC, and the Cramér-Rao bound," *IEEE Trans. Signal Process.*, vol. 65, no. 4, pp. 933–946, Feb 2017.

[22] S. Sedighi, R. B. S. Mysore, S. Maleki, and B. Ottersten, "Consistent least squares estimator for co-array-based DoA estimation," in *IEEE 10th Sensor Array and Multichannel Signal Processing Workshop (SAM)*, July 2018, pp. 524–528.

[23] S. Sedighi, B. S. M. R. Rao, and B. Ottersten, "An asymptotically efficient weighted least squares estimator for co-array-based DoA estimation," *IEEE Trans. Signal Process.*, vol. 68, pp. 589–604, 2020.

[24] S. Sedighi, M. R. Bhavani Shankar, and B. Ottersten, "A statistically efficient estimator for co-array based DoA estimation," in *52nd Asilomar Conference on Signals, Systems, and Computers*, 2018, pp. 880–883.

[25] R. H. Walden, "Analog-to-digital converter survey and analysis," *IEEE J. Sel. Areas Commun.*, vol. 17, no. 4, pp. 539–550, April 1999.

[26] H. Sun, A. Nallanathan, C. Wang, and Y. Chen, "Wideband spectrum sensing for cognitive radio networks: a survey," *IEEE Wirel. Commun.*, vol. 20, no. 2, pp. 74–81, 2013.

[27] J. Lunden, V. Koivunen, and H. V. Poor, "Spectrum exploration and exploitation for cognitive radio: Recent advances," *IEEE Signal Process. Mag.*, vol. 32, no. 3, pp. 123–140, 2015.

[28] J. Hasch, E. Topak, R. Schnabel, T. Zwick, R. Weigel, and C. Waldschmidt, "Millimeter-wave technology for automotive radar sensors in the 77 GHz frequency band," *IEEE Trans. Microw. Theory Techn.*, vol. 60, no. 3, pp. 845–860, 2012.

[29] B. F. Burke, F. Graham-Smith, and P. N. Wilkinson, *An introduction to radio astronomy*. Cambridge University Press, 2019.

[30] L. Lu, G. Y. Li, A. L. Swindlehurst, A. Ashikhmin, and R. Zhang, "An overview of massive MIMO: Benefits and challenges," *IEEE J. Sel. Topics Signal Process.*, vol. 8, no. 5, pp. 742–758, Oct 2014.

[31] M. J. Pelgrom, "Analog-to-digital conversion," in *Analog-to-Digital Conversion*. Springer, 2013, pp. 325–418.

[32] A. Gokceoglu, E. Björnson, E. G. Larsson, and M. Valkama, "Spatio-temporal waveform design for multiuser massive MIMO downlink with 1-bit receivers," *IEEE J. Sel. Topics Signal Process.*, vol. 11, no. 2, pp. 347–362, 2017.

[33] A. K. Saxena, I. Fijalkow, and A. L. Swindlehurst, "Analysis of one-bit quantized precoding for the multiuser massive MIMO downlink," *IEEE Trans. Signal Process.*, vol. 65, no. 17, pp. 4624–4634, 2017.

[34] S. Rao, A. Mezghani, and A. L. Swindlehurst, "Channel estimation in one-bit massive MIMO systems: Angular versus unstructured models," *IEEE J. Sel. Topics Signal Process.*, vol. 13, no. 5, pp. 1017–1031, Sep. 2019.

[35] H. Pirzadeh, G. Seco-Granados, S. Rao, and A. L. Swindlehurst, "Spectral efficiency of one-bit sigma-delta massive MIMO systems," *IEEE J. Sel. Areas Commun.*, pp. 1–1, 2020.

[36] Q. Wan, J. Fang, H. Duan, Z. Chen, and H. Li, "Generalized Bussgang LMMSE channel estimation for one-bit massive MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 19, no. 6, pp. 4234–4246, 2020.

[37] H. Zayyani, M. Korki, and F. Marvasti, "Dictionary learning for blind one-bit compressed sensing," *IEEE Signal Process. Lett.*, vol. 23, no. 2, pp. 187–191, 2015.

[38] B. Zhao, L. Huang, J. Li, M. Liu, and J. Wang, "Deceptive SAR jamming based on 1-bit sampling and time-varying thresholds," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 11, no. 3, pp. 939–950, 2018.

[39] A. Ameri, A. Bose, J. Li, and M. Soltanalian, "One-bit radar processing with time-varying sampling thresholds," *IEEE Trans. Signal Process.*, vol. 67, no. 20, pp. 5297–5308, Oct 2019.

[40] S. J. Zahabi, M. M. Naghsh, M. Modarres-Hashemi, and J. Li, "One-bit compressive radar sensing in the presence of clutter," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 56, no. 1, pp. 167–185, 2020.

[41] F. Xi, Y. Xiang, S. Chen, and A. Nehorai, "Gridless parameter estimation for one-bit MIMO radar with time-varying thresholds," *IEEE Trans. Signal Process.*, vol. 68, pp. 1048–1063, 2020.

[42] S. Sedighi, K. V. Mishra, M. R. B. Shankar, and B. Ottersten, "Localization with one-bit passive radars in narrowband internet-of-things using multivariate polynomial optimization," 2020. [Online]. Available: arXiv:2007.15108v1

[43] O. Bar-Shalom and A. J. Weiss, "DoA estimation using one-bit quantized measurements," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 38, no. 3, pp. 868–884, July 2002.

[44] M. Stein, K. Barbe, and J. A. Nossek, "DoA parameter estimation with 1-bit quantization bounds, methods and the exponential replacement," in *20th International ITG Workshop on Smart Antennas*, March 2016, pp. 1–6.

[45] X. Huang, S. Bi, and B. Liao, "Direction-of-arrival estimation based on quantized matrix recovery," *IEEE Commun. Lett.*, vol. 24, no. 1, pp. 349–353, 2020.

[46] I. Yoffe, N. Regev, and D. Wulich, "On direction of arrival estimation with 1-bit quantizer," in *IEEE Radar Conference (RadarConf)*, 2019, pp. 1–6.

[47] C. Stöckle, J. Munir, A. Mezghani, and J. A. Nossek, "1-bit direction of arrival estimation based on compressed sensing," in *IEEE 16th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, June 2015, pp. 246–250.

[48] X. Huang, P. Xiao, and B. Liao, "One-bit direction of arrival estimation with an improved fixed-point continuation algorithm," in *10th International Conference on Wireless Communications and Signal Processing (WCSP)*, 2018, pp. 1–4.

[49] X. Meng and J. Zhu, "A generalized sparse bayesian learning algorithm for 1-bit DoA estimation," *IEEE Commun. Lett.*, vol. 22, no. 7, pp. 1414–1417, 2018.

[50] T. Chen, M. Guo, and X. Huang, "Direction finding using compressive one-bit measurements," *IEEE Access*, vol. 6, pp. 41 201–41 211, 2018.

[51] X. Huang and B. Liao, "One-bit MUSIC," *IEEE Signal Process. Lett.*, vol. 26, no. 7, pp. 961–965, July 2019.

[52] J. H. Van Vleck and D. Middleton, "The spectrum of clipped noise," *Proceedings of the IEEE*, vol. 54, no. 1, pp. 2–19, 1966.

[53] C. Liu and P. P. Vaidyanathan, "One-bit sparse array DoA estimation," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, March 2017, pp. 3126–3130.

[54] K. N. Ramamohan, S. Prabhakar Chepuri, D. F. Comesaña, and G. Leus, "Blind calibration of sparse arrays for DoA estimation with analog and one-bit measurements," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2019, pp. 4185–4189.

[55] Z. Cheng, S. Chen, Q. Shen, J. He, and Z. Liu, "Direction finding of electromagnetic sources on a sparse cross-dipole array using one-bit measurements," *IEEE Access*, vol. 8, pp. 83 131–83 143, 2020.

[56] C. Zhou, Y. Gu, Z. Shi, and M. Haardt, "Direction-of-arrival estimation for coprime arrays via coarray correlation reconstruction: A one-bit perspective," in *IEEE 11th Sensor Array and Multichannel Signal Processing Workshop (SAM)*, 2020, pp. 1–4.

[57] E. L. Lehmann and G. Casella, *Theory of point estimation*. Springer Science & Business Media, 2006.

[58] A. W. Van der Vaart, *Asymptotic statistics*. Cambridge university press, 2000, vol. 3.

[59] S. M. Kay, *Fundamentals of statistical signal processing, Volume I: Estimation theory*. Englewood Cliffs, NJ: Prentice Hall, 1993.

[60] P. Stoica and P. Babu, "The Gaussian data assumption leads to the largest Cramér-Rao bound [lecture notes]," *IEEE Signal Process. Mag.*, vol. 28, no. 3, pp. 132–133, 2011.

[61] I. Abrahamson *et al.*, "Orthant probabilities for the quadrivariate normal distribution," *The Annals of Mathematical Statistics*, vol. 35, no. 4, pp. 1685–1703, 1964.

[62] C. L. Liu and P. P. Vaidyanathan, "Remarks on the spatial smoothing step in coarray MUSIC," *IEEE Signal Process. Lett.*, vol. 22, no. 9, pp. 1438–1442, Sept 2015.

[63] A. Papoulis, *Probability, random variables and stochastic processes*. New York: McGraw-hill, 2002.

[64] M. Kaveh and A. Barabell, "The statistical performance of the MUSIC and the minimum-norm algorithms in resolving plane waves in noise," *IEEE Trans. Acoustics, Speech, and Signal Process.*, vol. 34, no. 2, pp. 331–341, April 1986.

[65] D. N. Lal, "A note on a form of Tchebycheff's inequality for two or more variables," *Sankhyā: The Indian Journal of Statistics (1933-1960)*, vol. 15, no. 3, pp. 317–320, 1955.

[66] C. Carathéodory and L. Fejér, "Über den zusammenhang der extremen von harmonischen funktionen mit ihren koeffizienten und über den picard-landau'schen satz," *Rendiconti del Circolo Matematico di Palermo (1884-1940)*, vol. 32, no. 1, pp. 218–239, 1911.

[67] T. J. Rothenberg, "Identification in parametric models," *Econometrica: Journal of the Econometric Society*, pp. 577–591, 1971.

[68] G. P. Styan, "Hadamard products and multivariate statistical analysis," *Linear algebra and its applications*, vol. 6, pp. 217–240, 1973.

[69] J. R. Magnus and H. Neudecker, *Matrix differential calculus with applications in statistics and econometrics*. USA: New York: John Wiley and Sons, Inc., 1988.

Saeid Sedighi (S'17–M'2021) was born in Tehran, Iran, in 1987. He received the M.Sc. degree in Electrical Engineering from Imam Khomeini International University (IKIU), Qazvin, Iran, in 2012, and the Ph.D. degree in Electrical Engineering from University of Luxembourg, Luxembourg, in 2021. He held a research position with the Department of Electrical Engineering, Imam Khomeini International University (IKIU) from October 2012 to December 2013 as well as from January 2015 to August 2016. He was also a researcher at Iran's National Elites Foundation from January 2014 to January 2015. He is currently a postdoctoral researcher at the Interdisciplinary Centre for Security, Reliability and Trust (SnT), University of Luxembourg, Luxembourg. His research interests are statistical signal processing, array processing, radar, sparse sensing, low-resolution sampling, detection and estimation theory.

M. R. Bhavani Shankar received Masters and Ph. D in Electrical Communication Engineering from Indian Institute of Science, Bangalore in 2000 and 2007 respectively. He was a Post Doc at the ACCESS Linnaeus Centre, Signal Processing Lab, Royal Institute of Technology (KTH), Sweden from 2007 to September 2009. He joined SnT in October 2009 as a Research Associate and is currently a Senior Research Scientist/ Assistant Professor at SnT and is leading the Radar Signal Processing activities. He was with Beceem Communications, Bangalore from 2006 to 2007 as a Staff Design Engineer working on Physical Layer algorithms for WiMAX compliant chipsets. He was a visiting student at the Communication Theory Group, ETH Zurich, headed by Prof. Helmut Bölcskei during 2004. Prior to joining Ph. D, he worked on Audio Coding algorithms in Sasken Communications, Bangalore as a Design Engineer from 2000 to 2001. His research interests include Design and Optimization of MIMO Communication Systems, Automotive Radar and Array Processing, polynomial signal processing, Satellite communication systems, Resource Allocation and Fast Algorithms for Structured Matrices.

He is currently on the Executive Committee of the IEEE Benelux joint chapter on communications and vehicular technology and serves as handling editor for Elsevier Signal Processing and member of the EURASIP Technical Area Committee on Theoretical and Methodological Trends in Signal Processing.. He was a co-recipient of the 2014 Distinguished Contributions to Satellite Communications Award, from the Satellite and Space Communications Technical Committee of the IEEE Communications Society.

Mojtaba Soltanalian (Senior Member, IEEE) received the Ph.D. degree in electrical engineering (with specialization in signal processing) from the Department of Information Technology, Uppsala University, Sweden, in 2014. He is currently with the faculty of the Electrical and Computer Engineering Department, University of Illinois at Chicago (UIC), Chicago, IL, USA. Before joining UIC, he held research positions with the Interdisciplinary Centre for Security, Reliability and Trust (SnT, University of Luxembourg), and California Institute of Technology, Pasadena, CA, USA. His research interests include interplay of signal processing, learning and optimization theory, and specifically different ways the optimization theory can facilitate a better processing and design of signals for collecting information, communication, and also to form a more profound understanding of data, whether it is in everyday applications or in large-scale, complex scenarios. Dr. Soltanalian serves as an Associate Editor for IEEE Transactions on Signal Processing and as the Chair of the IEEE Signal Processing Society Chapter in Chicago. He was the recipient of the 2017 IEEE Signal Processing Society Young Author Best Paper Award, an also the 2018 European Signal Processing Association Best PhD Award.

Björn Ottersten (S'87–M'89–SM'99–F'04) received the M.S. degree in electrical engineering and applied physics from Linköping University, Linköping, Sweden, in 1986, and the Ph.D. degree in electrical engineering from Stanford University, Stanford, CA, USA, in 1990. He has held research positions with the Department of Electrical Engineering, Linköping University, the Information Systems Laboratory, Stanford University, the Katholieke Universiteit Leuven, Leuven, Belgium, and the University of Luxembourg, Luxembourg. From 1996 to 1997, he was the Director of Research with ArrayComm, Inc., a start-up in San Jose, CA, USA, based on his patented technology. In 1991, he was appointed Professor of signal processing with the Royal Institute of Technology (KTH), Stockholm, Sweden. Dr. Ottersten has been Head of the Department for Signals, Sensors, and Systems, KTH, and Dean of the School of Electrical Engineering, KTH. He is currently the Director for the Interdisciplinary Centre for Security, Reliability and Trust, University of Luxembourg. He is a recipient of the IEEE Signal Processing Society Technical Achievement Award, the EURASIP Group Technical Achievement Award, and the European Research Council advanced research grant twice. He has co-authored journal papers that received the IEEE Signal Processing Society Best Paper Award in 1993, 2001, 2006, 2013, and 2019, and 8 IEEE conference papers best paper awards. He has been a board member of IEEE Signal Processing Society, the Swedish Research Council and currently serves of the boards of EURASIP and the Swedish Foundation for Strategic Research. Dr. Ottersten has served as Editor in Chief of EURASIP Signal Processing, and acted on the editorial boards of IEEE Transactions on Signal Processing, IEEE Signal Processing Magazine, IEEE Open Journal for Signal Processing, EURASIP Journal of Advances in Signal Processing and Foundations and Trends in Signal Processing. He is a fellow of EURASIP.