Direction of Arrival Estimation of Wideband Sources Using Sparse Linear Arrays

Wang, Feiyu ; Tian, Zhi; Leus, Geert; Fang, Jun

DOI
10.1109/TSP.2021.3094718

Publication date
2021

Document Version
Final published version

Published in
IEEE Transactions on Signal Processing

Direction of Arrival Estimation of Wideband Sources Using Sparse Linear Arrays

Feiyu Wang, Zhi Tian, *Fellow, IEEE*, Geert Leus, *Fellow, IEEE*, and Jun Fang, *Senior Member, IEEE*

***Abstract*—In this paper, we study the problem of wideband direction of arrival (DoA) estimation with sparse linear arrays (SLAs), where a number of uncorrelated wideband signals impinge on an SLA and the data is collected from multiple frequency bins. To boost the performance and perform underdetermined DoA estimation, the difference co-array response matrices for all frequency bins are constructed first. Then, to merge the data from different frequency bins, we resort to the Jacobi-Anger approximation to transform the co-array response matrices of all frequency bins into a single virtual uniform linear array (ULA) response matrix. The major advantage of this approach is that the transformation matrices are all signal independent. For the special case where all sources share an identical distribution of the power spectrum, we develop two super-resolution off-the-grid DoA estimation approaches based on atomic norm minimization (ANM), one with and one without prior knowledge of the power spectrum. Our solution is able to resolve more sources than the number of antennas but also more than the number of degrees of freedom (DoF) of the difference co-array of the SLA. For the general case where each source has an arbitrary power spectrum, we propose a multi-task ANM method to exploit the joint sparsity from all frequency bins. Simulation results show that our proposed methods present a clear performance advantage over existing methods, and achieve an estimation accuracy close to the associated Cramér-Rao bounds (CRBs).**

***Index Terms*—Wideband direction-of-arrival (DoA) estimation, sparse linear array (SLA), Jacobi-Anger approximation, atomic norm minimization (ANM).**

I. Introduction

DIRECTION of arrival (DoA) estimation using a sensor array has been an active research area for decades, with broad applications in a number of fields such as wireless communication, acoustics and passive sonar [1]. Most of the conventional DoA estimation techniques focus on the narrowband far-field scenario and are subspace-based methods, where multiple signal classification (MUSIC) [2] and estimation of signal parameters via rotational invariant techniques (ESPRIT) [3] are two of the most successful methods. Whereas MUSIC works for any array configuration, ESPRIT requires a specific shift invariant structure. Inspired by compressed sensing (CS) theory [4], over the past few years, sparsity-inducing DoA estimation techniques [5]–[7] have been introduced to exploit the inherent sparsity in the angular domain and they work for arbitrary array configurations. All these CS-based methods rely on a fixed angular sampling grid that serves as the set of all candidates of DoA values, and the DoAs of interest constitute the support of the sparse signal to be recovered. The angular resolution is limited by the grid spacing, which can be refined in subsequent rounds. Recently, two pioneering super-resolution (or off-the-grid CS) techniques have been proposed that recover frequencies in a continuous domain; one method is based on total variation minimization [8] and the other on atomic norm minimization [9]. Both approaches theoretically guarantee that the continuous-valued frequencies can be exactly recovered provided the frequencies are well separated and there is no noise. Note that although [8], [9] focus on frequency estimation, they can equivalently be used for DoA estimation of narrowband far-field sources as well. Other related super-resolution techniques include [10]–[16], which eliminate grid mismatch of grid-based CS methods, and some theoretical advances in this field are achieved [17]–[21]. Besides, super-resolution can also be extended to the 2-D line spectrum estimation cases [22]–[24].

Apart from narrowband DoA estimation, DoA estimation methods for wideband sources have also been studied for decades. Conventional methods rely on a bank of narrowband filters, which decompose the received wideband signal into several narrowband signals, and then employ some combination of narrowband signal processing algorithms to obtain the DoA estimates. A simple subspace-based wideband method is the incoherent signal subspace method (ISSM) [25], which applies narrowband techniques, such as MUSIC [2] and ESPRIT [3], independently to the outputs of the filter bank. Alternatively, in the coherent signal subspace method (CSSM) [26], focusing matrices are designed to combine the information from different frequency subbands, leading to improved performance compared to ISSM. The focusing schemes for the CSSM are further developed in [27]. Some variants of the CSSM have also been proposed, such as the weighted average of signal subspaces (WAVES) method [28], the test of orthogonality of projected subspaces (TOPS) method [29] and the subspace-based

Manuscript received June 26, 2020; revised February 1, 2021 and June 19, 2021; accepted June 22, 2021. Date of publication July 7, 2021; date of current version August 16, 2021. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Yuejie Chi. This work was supported in part by the ASPIRE project (project 14926 within the OTP program of NWO-TTW), in part by the National Science Foundation of China under Grants 61871091 and 61934008, and in part by the China Scholarship Council under Grant 201806070119. Part of this work was presented at the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP), 2020, with the title "Wideband Direction of Arrival Estimation with Sparse Linear Arrays". *(Corresponding author: Feiyu Wang.)*

Feiyu Wang and Geert Leus are with the Department of Microelectronics, Delft University of Technology, 2628 Delft CD, The Netherlands (e-mail: fywang.ee@outlook.com; g.j.t.leus@tudelft.nl).

Zhi Tian is with the Department of Electrical and Computer Engineering, George Mason University, Fairfax, VA 22030 USA (e-mail: ztian1@gmu.edu).

Jun Fang is with the National Key Laboratory on Communications, University of Electronic Science and Technology of China, Chengdu 611731, China (e-mail: JunFang@uestc.edu.cn).

Digital Object Identifier 10.1109/TSP.2021.3094718

autofocusing approach [30], [31]. Based on CS theory [4], in the past decade some wideband DoA estimation approaches based on sparse signal recovery techniques have also been developed [32], [33].

Most of the narrowband or wideband DoA estimation methods have been confined to the case of uniform linear arrays (ULAs) and/or resolve up to $N-1$ sources with an N-element array. However, the topic of DoA estimation with more sources than sensors has raised considerable attention [34]–[37]. An efficient way to achieve this goal is to use a sparse linear array (SLA) and to construct a new difference co-array with more degrees of freedom (DoF) than that directly obtained from the physical SLA. From the co-array perspective, the minimum redundancy array (MRA) [38] and sparse ruler array (SRA) [39] have been considered as optimal SLA designs, yet their antenna locations cannot be computed in closed form. Hence, several more tractable SLA configurations have been proposed in the past decade, such as the nested array [35] and the coprime array [36]. Based on these SLA configurations, most works focus on developing DoA estimation algorithms under the narrowband assumption [40]–[42]. For the wideband scenario, the DoA estimation problem for SLAs becomes more involved. In [43], a combined spatial smoothing MUSIC (SS-MUSIC) angular spectrum is constructed to combine the information from all frequency bins similar to ISSM. In [44], a focusing Khatri-Rao (FKR) subspace-based approach is proposed, where the way to calculate the focusing matrices is similar to that in CSSM, but now extended to the difference co-array. Some DoA estimation methods based on grid-based CS and sparse reconstruction are also proposed for wideband sources [45]–[47], yet they suffer from leakage effects when the sources are off the grid. To deal with the issue of grid mismatch, in [48], the true difference co-array response matrices from all frequency bands are approximated as a summation of a presumed dictionary and a structured parameterized matrix via a first-order Taylor expansion. Note that for all the above methods, the number of sources to be recovered should be less than the number of DoF of the difference co-array. Furthermore, coherent wideband processing is desired to boost performance, but existing focusing-based techniques often require accurate DoA knowledge that is not readily available in practice.

In this paper, we focus on wideband DoA estimation with SLAs. Similar to previous works [43]–[47], the difference co-array response matrices for all frequency bins are constructed to enhance the effective number of DoF of the array. But as opposed to focusing, which generally requires an initial estimate of the DoAs or the signal subspace, we resort to the Jacobi-Anger approximation from the manifold separation technique (MST) [49]–[52] in array processing, to transform the difference co-array response matrices from the different frequency bins into a single virtual ULA response matrix. This transformation allows us to combine the data from different frequencies easily, while the transformation matrices constructed from MST are signal independent and only depend on the corresponding frequency bin. Furthermore, in contrast to existing methods, we also consider the special case where all sources share the same normalized power spectrum, which is the case for many practical scenarios. This assumption allows us to resolve more sources than the number of DoF of the difference co-array. Accordingly, two super-resolution off-the-grid DoA estimation approaches are proposed based on atomic norm minimization (ANM), one with and one without prior knowledge of the power spectrum. For the general case where each source has an arbitrary power spectrum, we propose a multi-task ANM method to exploit the joint sparsity from all frequency bins. The corresponding Cramér-Rao bound (CRB) results for the three cases are also derived. Simulation results show that, through efficiently merging the information from different frequency subbands, our proposed methods outperform the state of the art and can achieve an estimation accuracy close to the associated CRBs. Furthermore, under the identical normalized power spectrum assumption, the proposed methods are capable of resolving more sources than the number of DoF of the difference co-array. Note that in the literature of classical wideband DoA estimation methods, although the data from multiple frequency bins are combined and exploited, the number of resolvable sources still does not exceed the number of antennas of the physical array (or co-array). Also, in the classical theory of distributed compressed sensing [53], even with multiple random dictionaries and a joint sparsity pattern, practical algorithms cannot resolve more sources than the number of measurements.

We noticed that in the recent paper [54], the authors also used the Jacobi-Anger expansion for wideband DoA estimation with an arbitrary array geometry. Different from our work, which is based on the difference co-array, their work only considered wideband DoA estimation based on the physical array directly. This implicates that the proposed method cannot resolve more sources than the number of physical antennas, while our method can achieve this, and can even resolve more sources than the number of DoF of the difference co-array with an additional assumption about the signal power spectrum. Furthermore, the method in [54] used a coherent annihilating filter to achieve wideband DoA estimation, where an initial estimate of the DoAs is required to compute the coefficients of the annihilating filter. This means that the method is still signal dependent, while our proposed method is totally signal independent.

The rest of the paper is organized as follows. In Section II, we provide the signal model and basic assumptions. The joint data model and additional assumptions are discussed in Section III. In Section IV, we propose two super-resolution off-the-grid DoA estimation approaches under the identical normalized power spectrum assumption. A multi-task ANM method for the general case is developed in Section V. A CRB analysis is conducted in Section VI. Simulation results are provided in Section VII, followed by concluding remarks in Section VIII.

II. Signal Model

Consider K far-field wideband sources impinging on an SLA with N antenna elements. For wideband processing, at each antenna, the received signal is first sampled at Nyquist rate and partitioned into length-M segments, and then an M-point discrete Fourier transform (DFT) is applied to each segment to compute M frequency subbands (equivalently a filter bank can

be adopted). For simplicity, we start from the noiseless model using an infinite number of data segments. The effect of the noise and a finite observation time will be included later in the algorithm development, CRB analysis and simulation results. The noiseless version of the received signal for the nth antenna and mth frequency can be written as

$$x_{n,m}[l] = \sum_{k=1}^{K} a_{n,m}(\theta_k) s_{k,m}[l],\ n \in \mathcal{N},\ m \in \mathcal{M}, \quad (1)$$

where $\mathcal{N} \triangleq \{1, \ldots, N\}$, $\mathcal{M} \triangleq \{1, \ldots, M\}$, $l \in \mathbb{N}^+$ denotes the index of the segment, $s_{k,m}[l]$ represents the source signal related to the kth source and mth frequency, $\theta_k \in [0, \pi)$ is the DoA of the kth source signal, and $a_{n,m}(\theta)$ represents the channel response at angle θ for the nth antenna and mth frequency. The channel response can generally be expressed as

$$a_{n,m}(\theta) \triangleq e^{-j2\pi d_n \cos(\theta)/\lambda_m},\ n \in \mathcal{N},\ m \in \mathcal{M}, \quad (2)$$

where d_n is the distance from the nth antenna to the first antenna and λ_m is the wavelength corresponding to the mth frequency f_m. Here we have $d_1 = 0$. Note that in the literature on DoA estimation based on either the actual physical array or the difference co-array, to obtain a well-featured Vandermonde structure from the array manifold, d_n should be constrained as an integer multiple of the basic element spacing. In our model, we assume arbitrary arrays and do not have such a basic element spacing assumption on the antenna locations $\{d_n\}_{n=1}^N$. For the nth antenna and the mth frequency, the K source signals $\{s_{k,m}[l]\}_{k=1}^K$ are assumed to be mutually uncorrelated.

Stacking the signals of all antennas for the mth frequency, i.e, introducing $\boldsymbol{x}_m[l] \triangleq [x_{1,m}[l]\ x_{2,m}[l]\ \ldots\ x_{N,m}[l]]^T$ and $\boldsymbol{a}_m(\theta) \triangleq [a_{1,m}(\theta)\ a_{2,m}(\theta)\ \ldots\ a_{N,m}(\theta)]^T$, we obtain

$$\boldsymbol{x}_m[l] = \sum_{k=1}^{K} \boldsymbol{a}_m(\theta_k) s_{k,m}[l] = \boldsymbol{A}_m(\boldsymbol{\theta}) \boldsymbol{s}_m[l], \quad (3)$$

where $\boldsymbol{s}_m[l] \triangleq [s_{1,m}[l]\ \ s_{2,m}[l]\ \ \ldots\ \ s_{K,m}[l]]^T$, $\boldsymbol{A}_m(\boldsymbol{\theta}) \triangleq [\boldsymbol{a}_m(\theta_1)\, \boldsymbol{a}_m(\theta_2)\ \ldots\ \boldsymbol{a}_m(\theta_K)]$ and $\boldsymbol{\theta} \triangleq [\theta_1\, \theta_2\ \ldots\ \theta_K]^T$. Computing the output covariance matrix $\boldsymbol{R}_m \triangleq E\{\boldsymbol{x}_m[l]\boldsymbol{x}_m^H[l]\}$, we obtain

$$\begin{aligned} \boldsymbol{R}_m &= \boldsymbol{A}_m(\boldsymbol{\theta}) E\{\boldsymbol{s}_m[l]\boldsymbol{s}_m^H[l]\}\boldsymbol{A}_m^H(\boldsymbol{\theta}) \\ &= \boldsymbol{A}_m(\boldsymbol{\theta}) \text{diag}\,(\boldsymbol{\gamma}_m)\, \boldsymbol{A}_m^H(\boldsymbol{\theta}), \end{aligned} \quad (4)$$

where $E\{\cdot\}$ represents the expectation operator, and $\boldsymbol{\gamma}_m = [\gamma_{1,m}\, \gamma_{2,m}\ \ldots\ \gamma_{K,m}]^T$ with $\gamma_{k,m}$ being the source power related to the kth source and mth frequency, i.e., $\gamma_{k,m} \triangleq E\{|s_{k,m}[l]|^2\}$. Vectorizing this expression, we obtain

$$\tilde{\boldsymbol{r}}_m \triangleq \text{vec}\{\boldsymbol{R}_m\} = \widetilde{\boldsymbol{B}}_m(\boldsymbol{\theta}) \boldsymbol{\gamma}_m, \quad (5)$$

where $\widetilde{\boldsymbol{B}}_m(\boldsymbol{\theta}) \triangleq (\boldsymbol{A}_m^*(\boldsymbol{\theta}) \circ \boldsymbol{A}_m(\boldsymbol{\theta}))$ denotes the new array response matrix from the co-array perspective, with $(\cdot)^*$ and $\circ$ standing for the complex conjugate and Khatri-Rao product, respectively. Define the antenna locations of the difference co-array as

$$\mathcal{D} \triangleq \{d_i - d_j :\ i, j \in \mathcal{N}\}, \quad (6)$$

and the cardinality of $\mathcal{D}$ as N_{co}, which indicates the DoF of the co-array. Denote $\{\xi_n\}_{n=1}^{N_{\text{co}}}$ as the elements of $\mathcal{D}$ and let $\mathcal{N}_{\text{co}} \triangleq \{1, \ldots, N_{\text{co}}\}$. After averaging the repeated rows in $\widetilde{\boldsymbol{B}}_m(\boldsymbol{\theta})$ and the related entries in $\tilde{\boldsymbol{r}}_m$, we have

$$\boldsymbol{r}_m \triangleq \boldsymbol{J}\tilde{\boldsymbol{r}}_m = \boldsymbol{J}\widetilde{\boldsymbol{B}}_m(\boldsymbol{\theta})\boldsymbol{\gamma}_m = \boldsymbol{B}_m(\boldsymbol{\theta})\boldsymbol{\gamma}_m, \quad (7)$$

where $\boldsymbol{J} \in \{0, 1/n :\ n \in \mathbb{N}^+\}^{N_{\text{co}} \times N^2}$ is the corresponding select-and-average matrix, and

$$\boldsymbol{B}_m(\boldsymbol{\theta}) \triangleq \boldsymbol{J}\widetilde{\boldsymbol{B}}_m(\boldsymbol{\theta}) \quad (8)$$

is the (non-redundant) co-array response matrix related to the mth frequency. The kth column of $\boldsymbol{B}_m(\boldsymbol{\theta})$ can be expressed as $\boldsymbol{b}_m(\theta_k) = [b_{1,m}(\theta_k)\, b_{2,m}(\theta_k)\ \ldots\ b_{N_{\text{co}},m}(\theta_k)]^T$, with

$$b_{n,m}(\theta) \triangleq e^{-j2\pi\xi_n \cos(\theta)/\lambda_m},\ n \in \mathcal{N}_{\text{co}},\ m \in \mathcal{M}. \quad (9)$$

The problem in this paper is to recover the continuous-valued DoAs $\{\theta_k\}_{k=1}^K$, given $\{\boldsymbol{r}_m\}_{m=1}^M$. In the next section, we will introduce the joint data model for coherent wideband processing, as well as additional assumptions.

III. Joint Data Model

A. Jacobi-Anger Expansion for Data Merging

In order to merge the data models in (7) for all M frequencies, we make the following assumption.

- A1: We assume that all co-array response matrices $\boldsymbol{B}_m(\boldsymbol{\theta})$ can be transformed into a single common virtual ULA response matrix $\boldsymbol{V}(\boldsymbol{\theta}) = [\boldsymbol{v}(\theta_1)\, \boldsymbol{v}(\theta_2)\ \ldots\ \boldsymbol{v}(\theta_K)]$, i.e., $\boldsymbol{B}_m(\boldsymbol{\theta}) = \boldsymbol{G}_m \boldsymbol{V}(\boldsymbol{\theta})$, where $\boldsymbol{v}(\theta) \triangleq [e^{-j\theta N_{\text{virt}}}\ \ldots\ e^{j\theta N_{\text{virt}}}]^T$, $\boldsymbol{G}_m \in \mathbb{C}^{N_{\text{co}} \times (2N_{\text{virt}}+1)}$ denotes the corresponding transformation matrix which depends on the mth frequency f_m only, and $(2N_{\text{virt}} + 1)$ is an odd number denoting the number of antennas in the virtual ULA.

One way to achieve this is by focusing, which unfortunately hinges on the unknown angles. To circumvent this obstacle, we resort to the more accurate Jacobi-Anger expansion, which provides a general infinite series expansion of exponentials of trigonometric functions in the basis of their harmonics [49], [50]. Specifically, from the definition in (9), the (n, k)th entry of the co-array response matrix $\boldsymbol{B}_m(\boldsymbol{\theta})$ can be written as

$$b_{n,m}(\theta_k) = \sum_{n_v=-\infty}^{\infty} j^{n_v} J_{n_v}\left(2\pi\frac{\xi_n}{\lambda_m}\right) e^{j\theta_k n_v} \quad (10)$$

where $J_{n_v}(\cdot)$ is the Bessel function of the first kind of order n_v. Note that although (10) indicates an infinite sum, the amplitude of $J_{n_v}(\cdot)$ decays rapidly as the value of n_v increases beyond the argument of the related Bessel function $J_{n_v}(\cdot)$ [50]. In practice, the infinite series can hence be truncated by considering only a limited number of modes as

$$\begin{aligned} b_{n,m}(\theta_k) &\approx \sum_{n_v=-N_{\text{virt}}}^{N_{\text{virt}}} j^{n_v} J_{n_v}\left(2\pi\frac{\xi_n}{\lambda_m}\right) e^{j\theta_k n_v} \\ &= \boldsymbol{g}_{n,m}^T \boldsymbol{v}(\theta_k), \end{aligned} \quad (11)$$

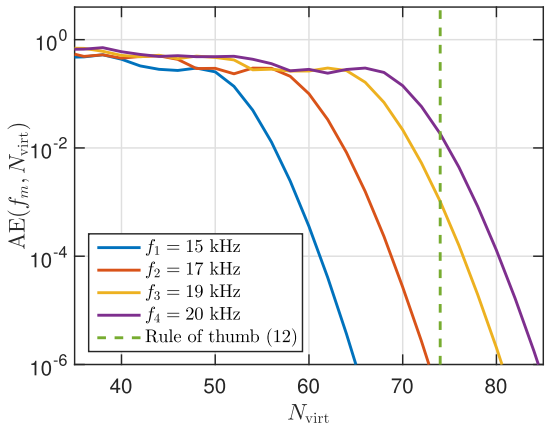


Fig. 1. The approximation error of the Jacobi-Anger expansion vs. the truncated order N_{virt}, where $\{\xi_1, \xi_2, \ldots, \xi_{N_{\mathrm{co}}}\} = \{-5\,d, -4\,d, \ldots, 5d\}$ with $N_{\mathrm{co}} = 11$ and $d = 0.04$ m.

where $\boldsymbol{g}_{n,m} \triangleq [g_{-N_{\mathrm{virt}}}^{(n,m)} \ \ldots \ g_{N_{\mathrm{virt}}}^{(n,m)}]^T$, and $g_{n_v}^{(n,m)} \triangleq j^{n_v} \cdot J_{n_v}(2\pi\xi_n/\lambda_m)$. We can make the resulting truncation error arbitrarily small by increasing the number of modes. Looking at (11), the entry of the virtual ULA steering vector $\boldsymbol{v}(\theta)$ for a large order n_v will have a marginal weight in the infinite series (10), and thus has little impact on the composition of $\boldsymbol{b}_m(\theta)$. Based on this fact, these terms with a large order n_v are negligible, contain little additional information and hence although they increase the number of virtual antennas, they do not increase the number of DoF since the related columns in $\boldsymbol{G}_m$ will be very small. Yet, the increased dimensions of the involved matrices will increase the complexity of any DoA estimation method. Hence, there is a trade-off between the accuracy of (11) and the computational complexity of the considered methods. This trade-off is usually determined by some rule of thumb. For example, in [50], the lower bound on the mode order N_{virt} is determined by the largest argument of the involved Bessel functions, i.e.,

$$N_{\mathrm{virt}} \geq \frac{2\pi}{\min_m \{\lambda_m\}} \cdot \max_n \{\xi_n\}. \tag{12}$$

To better illustrate this point, we define the approximation error of the Jacobi-Anger expansion as

$$\mathrm{AE}(f_m, N_{\mathrm{virt}}) \triangleq \max_{\theta\in[0,\pi)} \frac{1}{N_{\mathrm{co}}} ||\boldsymbol{b}_m(\theta) - \boldsymbol{G}_m \boldsymbol{v}(\theta)||_2^2$$

where $\boldsymbol{G}_m \triangleq [\boldsymbol{g}_{1,m} \ \boldsymbol{g}_{2,m} \ \ldots \ \boldsymbol{g}_{N_{\mathrm{co}},m}]^T$ is the corresponding transformation matrix. Fig. 1 depicts the approximation error of the Jacobi-Anger expansion vs. the truncated order N_{virt}, where we set $\{f_1, f_2, f_3, f_4\} = \{15, 17, 19, 20\}$ kHz and $\{\xi_1, \xi_2, \ldots, \xi_{N_{\mathrm{co}}}\} = \{-5\,d, -4\,d, \ldots, 5d\}$ with $N_{\mathrm{co}} = 11$ and $d = 0.04$ m. The choice of N_{virt} by the rule of thumb in (12) is also included. For a fixed truncated order N_{virt}, we observe that a higher frequency corresponds to a larger approximation error. This explains why in (12) the truncated order N_{virt} is only determined by the maximum frequency, i.e. $\max_m\{f_m\}$ (or the minimum wavelength). If we choose N_{virt} by the rule of thumb, the maximum approximation error is around 10^{-2} for all frequency bins. Further, when $N_{\mathrm{virt}} = 80$, which is a bit larger than the one from the rule of thumb, we can have a marginal approximation error as low as 10^{-4}. Now, according to (11) and (12), the kth column of the co-array response matrix $\boldsymbol{B}_m(\boldsymbol{\theta})$ can be expressed as $\boldsymbol{b}_m(\theta_k) = \boldsymbol{G}_m \boldsymbol{v}(\theta_k)$, which corroborates Assumption A1.

Under Assumption A1, we can rewrite (7) as

$$\boldsymbol{r}_m = \boldsymbol{G}_m \boldsymbol{V}(\boldsymbol{\theta})\boldsymbol{\gamma}_m = \boldsymbol{G}_m \boldsymbol{c}_m, \tag{13}$$

where $\boldsymbol{c}_m \triangleq \boldsymbol{V}(\boldsymbol{\theta})\boldsymbol{\gamma}_m$. Here we can see that the vector $\boldsymbol{c}_m$, which conveys the DoA information, is independent of the related frequency point f_m. Furthermore, through this transformation, all the $\boldsymbol{c}_m$, $m \in \mathcal{M}$, share a single virtual response matrix $\boldsymbol{V}(\boldsymbol{\theta})$ with Vandermonde structure, which generates a joint structure that can be used to merge the data from different frequencies.

In the next subsection, we will introduce an additional assumption on the signal power spectrum, i.e., $\{\boldsymbol{\gamma}_m\}_{m=1}^M$, which occurs in some applications.

B. Identical Power Spectrum

In some scenarios, the following assumption can be adopted.

- A2: The normalized power spectrum profile of all the K sources is the same, indicated across frequencies by $\boldsymbol{p} = [p_1 \ p_2 \ \ldots \ p_M]^T$ with $\sum_{m=1}^M p_m = 1$. As a result, $\gamma_{k,m} = \alpha_k p_m$, where α_k is the total received power of the kth source.

This is for instance the case in a satellite communication system where all sources use the same modulation format and pulse shaping functions [55], which leads to an identical normalized power spectrum for every source. Furthermore, unlike terrestrial mobile communications, multipath has little effect on mobile-satellite links in most practical operating environments [56]. In particular, in open areas such as farm land or open fields, there are no obstacles on the line-of-sight (LoS) path between the receiver and the satellite. In the latter scenario, the channel is assumed to be flat block fading, i.e., the channel remains constant over a certain amount of coherence time, and hence the normalized power spectrum $\boldsymbol{p}$ is known *a priori* and identical for every source. Another case is the scenario where a single transmission is passing through a multipath channel, with large delays between the different paths. Then, from the eyes of the receiver, we have multiple uncorrelated sources impinging on the array with the same normalized power spectrum but different DoAs and attenuation factors, which leads to Assumption A2. Defining $\boldsymbol{\alpha} = [\alpha_1 \ \alpha_2 \ \ldots \ \alpha_K]^T$, we thus obtain $\boldsymbol{\gamma}_m = p_m \boldsymbol{\alpha}$.

Under Assumptions A1 and A2, we can rewrite (7) as

$$\boldsymbol{r}_m = p_m \boldsymbol{G}_m \boldsymbol{V}(\boldsymbol{\theta})\boldsymbol{\alpha} = \boldsymbol{H}_m \boldsymbol{V}(\boldsymbol{\theta})\boldsymbol{\alpha}, \tag{14}$$

where $\boldsymbol{H}_m \triangleq p_m \boldsymbol{G}_m$. Hence, aside from the frequency-dependent term $\boldsymbol{G}_m$ (or $\boldsymbol{H}_m$), the DoA-dependent components of the model are the same for all frequencies, which is useful for coherent wideband DoA estimation. On the other hand, it is still challenging to jointly utilize $\boldsymbol{r}_m$ from all frequencies in an efficient manner, because $\boldsymbol{G}_m$ (or $\boldsymbol{H}_m$) of size $N_{\mathrm{co}} \times (2N_{\mathrm{virt}} + 1)$ is not invertible for large N_{virt}.

C. Problem Setups

In the following two sections, effective wideband DoA estimation methods are developed for different cases depending on whether the power spectrum profile is identical to all sources or not, and whether it is known or not. The three cases (C1–C3) of interest, along with the corresponding parameters to be estimated, are summarized below:

- C1 Assumption A2 holds and the power spectrum $\boldsymbol{p}$ is known. The parameters to be estimated are $\boldsymbol{\theta}$ and $\boldsymbol{\alpha}$.
- C2 Assumption A2 holds and the power spectrum $\boldsymbol{p}$ is unknown. The parameters to be estimated are $\boldsymbol{\theta}$, $\boldsymbol{\alpha}$ and $\boldsymbol{p}$.
- C3 General case without Assumption A2. The parameters to be estimated are $\boldsymbol{\theta}$ and $\{\gamma_m\}_{m=1}^M$.

IV. Super-Resolution DoA Estimation: Identical Power Spectrum

In this section, we assume Assumption A2 holds, which means that all K sources share the same normalized power spectrum $\boldsymbol{p}$. Based on the joint data model derived in Section III, we investigate super-resolution techniques to achieve wideband DoA estimation. In the following, both the scenarios with and without prior knowledge of the power spectrum $\boldsymbol{p}$, are considered.

A. Known Power Spectrum (Case C1)

If the power spectrum $\boldsymbol{p}$ is known *a priori*, we can stack the different vectors $\boldsymbol{r}_m$ into $\boldsymbol{r} \triangleq [\boldsymbol{r}_1^T\ \boldsymbol{r}_2^T\ \dots\ \boldsymbol{r}_M^T]^T$ and merge the different equations (14) into $\boldsymbol{r} = \boldsymbol{H}\boldsymbol{V}(\boldsymbol{\theta})\boldsymbol{\alpha}$, where $\boldsymbol{H} \triangleq [\boldsymbol{H}_1^T\ \boldsymbol{H}_2^T\ \dots\ \boldsymbol{H}_M^T]^T$. We can then develop an algorithm to recover $\boldsymbol{\alpha}$ under the ANM framework. Note that the term $\boldsymbol{c} \triangleq \boldsymbol{V}(\boldsymbol{\theta})\boldsymbol{\alpha}$ is a linear combination of K complex sinusoids, and hence it has a sparse representation over the atom set

$$\mathcal{A} \triangleq \{\boldsymbol{v}(\theta):\ \theta \in [0, \pi)\}. \tag{15}$$

As a penalty function catered to the structure of the atom set $\mathcal{A}$, the atomic norm of $\boldsymbol{c}$ over $\mathcal{A}$ is defined as

$$\|\boldsymbol{c}\|_{\mathcal{A}} \triangleq \inf\{t > 0:\ \boldsymbol{c} \in t \cdot \mathrm{conv}(\mathcal{A})\}, \tag{16}$$

where $\mathrm{conv}(\mathcal{A})$ denotes the convex hull of $\mathcal{A}$. Hence, we can first recover $\boldsymbol{c}$ by solving the ANM problem as

$$\min_{\boldsymbol{c}} \quad \frac{\lambda_1}{2}\|\boldsymbol{r} - \boldsymbol{H}\boldsymbol{c}\|_2^2 + \|\boldsymbol{c}\|_{\mathcal{A}}, \tag{17}$$

where λ_1 denotes the regularization parameter to balance the tradeoff between the ANM and the data fitting error. Using an approach similar to [9], [10], the ANM problem (17) can be represented in an equivalent semi-definite programming (SDP) form as

$$\begin{aligned} \min_{t,\boldsymbol{u},\boldsymbol{c}} \quad & \frac{\lambda_1}{2}\|\boldsymbol{r} - \boldsymbol{H}\boldsymbol{c}\|_2^2 + \mathrm{trace}(\mathcal{T}(\boldsymbol{u})) + t \\ s.t. \quad & \begin{bmatrix} t & \boldsymbol{c}^H \\ \boldsymbol{c} & \mathcal{T}(\boldsymbol{u}) \end{bmatrix} \succeq 0, \end{aligned} \tag{18}$$

where $t \in \mathbb{R}^+$, and $\mathcal{T}(\boldsymbol{u})$ is a Hermitian Toeplitz matrix with the first column being $\boldsymbol{u}$. This SDP problem can be solved by some off-the-shelf solvers such as SeDuMi and SDPT3 [57], or some first-order fast algorithms such as the accelerated proximal gradient method or alternating direction method of multipliers (ADMM). Given $\boldsymbol{c}$ or $\mathcal{T}(\boldsymbol{u})$, DoA estimation can be performed using, for instance, any subspace-based method. In the simulations, root MUSIC [58] on $\mathcal{T}(\boldsymbol{u})$ is used to obtain the final DoA estimates $\boldsymbol{\theta}$. Note that the choice of $\boldsymbol{c}$ versus $\mathcal{T}(\boldsymbol{u})$ would not influence the DoA estimation performance. About the choice of the regularization parameter, similar to other regularization-based optimization problems, in our simulation, we pick the best one through a trial-and-error process. For a more practical implementation, cross validation [59] can be used to determine the regularization parameter.

B. Unknown Power Spectrum (Case C2)

If the power spectrum $\{p_m\}_{m=1}^M$ is unknown, we can first transform (13) into $\boldsymbol{r}_m q_m = \boldsymbol{G}_m \boldsymbol{V}(\boldsymbol{\theta})\boldsymbol{\alpha}$, where $q_m \triangleq 1/p_m$. Note that we consider only those frequencies for which p_m is not too small. Next, stacking the vectors $\boldsymbol{r}_m$ as $\boldsymbol{R} = \mathrm{blkdiag}(\boldsymbol{r}_1, \boldsymbol{r}_2, \dots, \boldsymbol{r}_M)$ and introducing $\boldsymbol{q} = [q_1\ q_2,\ \dots\ q_M]^T$, we can form $\boldsymbol{R}\boldsymbol{q} = \boldsymbol{G}\boldsymbol{V}(\boldsymbol{\theta})\boldsymbol{\alpha}$, where $\boldsymbol{G} \triangleq [\boldsymbol{G}_1^T\ \boldsymbol{G}_2^T\ \dots\ \boldsymbol{G}_M^T]^T$, and $\mathrm{blkdiag}(\cdot)$ denotes the block diagonal operator.

We now seek to jointly estimate the angles $\{\theta_k\}_{k=1}^K$ and the unknown power profile $\{p_m\}_{m=1}^M$. The unknown vectors $\boldsymbol{c} \triangleq \boldsymbol{V}(\boldsymbol{\theta})\boldsymbol{\alpha}$ and $\boldsymbol{q}$ can be solved from the following least squares problem with atomic-norm regularization:

$$\begin{aligned} \min_{\boldsymbol{c},\boldsymbol{q}} \quad & \frac{\lambda_2}{2}\|\boldsymbol{R}\boldsymbol{q} - \boldsymbol{G}\boldsymbol{c}\|_2^2 + \|\boldsymbol{c}\|_{\mathcal{A}} \\ s.t. \quad & \mathbf{1}^T\boldsymbol{q} = 1, \end{aligned} \tag{19}$$

where λ_2 denotes the regularization parameter for the data fitting error and the linear constraint $\mathbf{1}^T\boldsymbol{q} = 1$ is added to avoid the trivial solution $\boldsymbol{q} = \mathbf{0}$, with $\mathbf{1}$ and $\mathbf{0}$ denoting the vector/matrix of ones and zeros respectively. As before, the ANM problem (19) can be reformulated as an SDP problem as:

$$\begin{aligned} \min_{t,\boldsymbol{u},\boldsymbol{c},\boldsymbol{q}} \quad & \frac{\lambda_2}{2}\|\boldsymbol{R}\boldsymbol{q} - \boldsymbol{G}\boldsymbol{c}\|_2^2 + \mathrm{trace}(\mathcal{T}(\boldsymbol{u})) + t \\ s.t. \quad & \begin{bmatrix} t & \boldsymbol{c}^H \\ \boldsymbol{c} & \mathcal{T}(\boldsymbol{u}) \end{bmatrix} \succeq 0, \quad \mathbf{1}^T\boldsymbol{q} = 1. \end{aligned} \tag{20}$$

Based on $\boldsymbol{q}$ and $\boldsymbol{c}$ or $\mathcal{T}(\boldsymbol{u})$, the power spectrum $\{p_m\}_{m=1}^M$ and the angles $\{\theta_k\}_{k=1}^K$ can be readily obtained. For the simulations, we use element-wise inversion to obtain the power spectrum and root MUSIC [58] on $\mathcal{T}(\boldsymbol{u})$ to obtain the angles.

V. Super-Resolution DoA Estimation: General Case

In this section we consider super-resolution wideband DoA estimation for the general case, where the normalized power spectrum of the K sources is not necessarily the same. Resorting to the data model in (13), we can see that each $\boldsymbol{c}_m$, $m \in \mathcal{M}$, is a linear combination of K complex sinusoids and is parameterized by the same K DoAs. In addition, all $\boldsymbol{c}_m$, $m \in \mathcal{M}$ share a single virtual response matrix $\boldsymbol{V}(\boldsymbol{\theta})$ with Vandermonde structure, which provides us a joint sparse recovery framework [60] to merge the data collected from different frequency bins. As shown in [60], [61], for matrix $\boldsymbol{C} \triangleq [\boldsymbol{c}_1\ \boldsymbol{c}_2\ \dots\ \boldsymbol{c}_M]$, one can represent

it by defining an atom set $\bar{\mathcal{A}}$ as follows:

$$\bar{\mathcal{A}} \triangleq \{\boldsymbol{v}(\theta)\boldsymbol{\phi}^H : \theta \in [0,\pi), \|\boldsymbol{\phi}\|_2 = 1\}. \tag{21}$$

where $\boldsymbol{\phi} \in \mathbb{C}^M$. Then, as a penalty function catered to the structure of the atom set $\bar{\mathcal{A}}$, the corresponding atomic norm of matrix $\boldsymbol{C}$ over $\bar{\mathcal{A}}$ is defined as

$$\|\boldsymbol{C}\|_{\bar{\mathcal{A}}} \triangleq \inf\{t > 0 : \boldsymbol{C} \in t \cdot \mathrm{conv}(\bar{\mathcal{A}})\}. \tag{22}$$

Given $\{\boldsymbol{r}_m\}_{m=1}^M$ and introducing the transformation matrices $\{\boldsymbol{G}_m\}_{m=1}^M$, we can recover $\{\boldsymbol{c}_m\}_{m=1}^M$ under a multi-task ANM framework as

$$\begin{aligned} \min_{\boldsymbol{c}_1 \ldots \boldsymbol{c}_M} \quad & \frac{\lambda_3}{2}\sum_{m=1}^{M}\|\boldsymbol{r}_m - \boldsymbol{G}_m\boldsymbol{c}_m\|_2^2 + \|\boldsymbol{C}\|_{\bar{\mathcal{A}}} \\ s.t. \quad & \boldsymbol{C} = [\boldsymbol{c}_1\ \boldsymbol{c}_2\ \ldots\ \boldsymbol{c}_M] \end{aligned} \tag{23}$$

where λ_3 denotes the regularization parameter for the data fitting error. Note that this formulation can be seen as an extension of the multi-task compressed sensing framework [62] to the continuous domain. The multi-task ANM problem (23) can be represented in an equivalent SDP form as

$$\begin{aligned} \min_{\boldsymbol{T},\boldsymbol{u},\{\boldsymbol{c}_m\}_{m=1}^M} \quad & \frac{\lambda_3}{2}\sum_{m=1}^{M}\|\boldsymbol{r}_m - \boldsymbol{G}_m\boldsymbol{c}_m\|_2^2 \\ & + \mathrm{trace}(\mathcal{T}(\boldsymbol{u})) + \mathrm{trace}(\boldsymbol{T}) \\ s.t. \quad & \begin{bmatrix} \boldsymbol{T} & \boldsymbol{C}^H \\ \boldsymbol{C} & \mathcal{T}(\boldsymbol{u}) \end{bmatrix} \succeq 0, \\ & \boldsymbol{C} = [\boldsymbol{c}_1\ \boldsymbol{c}_2\ \ldots\ \boldsymbol{c}_M] \end{aligned} \tag{24}$$

where $\boldsymbol{T} \in \mathbb{C}^{(2N_{\mathrm{virt}}+1)\times(2N_{\mathrm{virt}}+1)}$. Given $\boldsymbol{C}$ or $\mathcal{T}(\boldsymbol{u})$, any subspace-based method could be used for DoA estimation. In the simulations, final DoA estimation is performed by applying root MUSIC [58] on $\mathcal{T}(\boldsymbol{u})$.

To avoid the high computational complexity of the SDP-based approaches to solve (24), a fast iterative algorithm could be developed based on ADMM [63]. This can be achieved in a similar way as in [64].

VI. CRB Analysis

In this section, we derive the Cramér-Rao bound (CRB) for the wideband DoA estimation problem considered in this paper. As is well known, the CRB is a lower bound on the variance of any unbiased estimator [65]. For classical subspace-based DoA estimation methods (both for narrowband and wideband scenarios), it is always hard to rigorously prove an estimator is unbiased or asymptotically unbiased, while the CRB can still be used to evaluate the performance of any method. Here we also use the CRB as a benchmark for evaluating the performance of our proposed methods. In addition, the CRB results illustrate the behavior of the resulting bounds, which helps us better understand the effect of different system parameters, including the noise power and the number of data segments.

We first specify the signal model considered in the CRB analysis. The received signals are a superposition of several harmonics corrupted by additive white Gaussian noise under a finite observation time. More specifically, we consider model (3) with the extra noise term as

$$\boldsymbol{x}_m[l] = \boldsymbol{A}_m(\boldsymbol{\theta})\boldsymbol{s}_m[l] + \boldsymbol{w}_m[l], \quad l \in \mathcal{L} \tag{25}$$

where $\boldsymbol{w}_m[l] \in \mathbb{C}^N$ is the additive noise vector, which is independent of the source signal, $\mathcal{L} \triangleq \{1,\ldots,L\}$ and L denotes the number of segments. We assume that the source and noise are distributed as $\boldsymbol{w}_m[l] \sim \mathcal{CN}(\boldsymbol{0}, \sigma_m^2\boldsymbol{I}_N)$ and $\boldsymbol{s}_m[l] \sim \mathcal{CN}(\boldsymbol{0}, \mathrm{diag}(\boldsymbol{\gamma}_m))$ respectively, where σ_m^2 denotes the noise power in the mth frequency bin, and $\mathcal{CN}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ denotes the circularly-symmetric complex Gaussian distribution with mean vector $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$. We assume the received signals from different segments and different frequency bins are mutually uncorrelated, i.e.,

$$E\{\boldsymbol{x}_m[l_1]\boldsymbol{x}_m^H[l_2]\} = \boldsymbol{0}, \quad l_1 \neq l_2,\ \forall m \in \mathcal{M}, \tag{26}$$

$$E\{\boldsymbol{x}_{m_1}[l_1]\boldsymbol{x}_{m_2}^H[l_2]\} = \boldsymbol{0}, \quad m_1 \neq m_2,\ \forall l_1, l_2 \in \mathcal{L}. \tag{27}$$

Note that for our proposed algorithms, we do not rely on assumptions (26) and (27), which should be verified depending on the application, and the additive noise $\boldsymbol{w}_m[l]$ and the source signal $\boldsymbol{s}_m[l]$ are not restricted to be circularly-symmetric complex Gaussian. Here we make these assumptions in order to facilitate the CRB analysis.

Notice that the array response matrix $\boldsymbol{A}_m(\boldsymbol{\theta})$ is parameterized by the DoA vector $\boldsymbol{\theta}$. For simplicity and notational convenience, we, instead, analyze the CRB for the following new vector $\boldsymbol{\psi} = [\psi_1\ \psi_2\ \ldots\ \psi_K]^T$ where ψ_k is defined as

$$\psi_k \triangleq \cos(\theta_k), \quad k = 1,\ldots,K. \tag{28}$$

Note that also in co-array based CRB studies for narrowband DoA estimation [66], [67], the CRB for $\boldsymbol{\psi}$ is investigated instead of $\boldsymbol{\theta}$ for the same reason. For notational convenience, in the following, we use $\boldsymbol{A}_m(\boldsymbol{\psi})$ to represent the array response matrix of the mth frequency bin. From (2) and the definition of $\{\psi_k\}_{k=1}^K$, the channel response can be expressed as

$$a_{n,m}(\psi_k) = e^{-j2\pi d_n\psi_k/\lambda_m},\ n \in \mathcal{N},\ m \in \mathcal{M}, \tag{29}$$

which is the (n,k)th element of $\boldsymbol{A}_m(\boldsymbol{\psi})$. Similarly, the kth column of $\boldsymbol{A}_m(\boldsymbol{\psi})$ can be written as $\boldsymbol{a}_m(\psi_k) \triangleq [a_{1,m}(\psi_k)\ a_{2,m}(\psi_k)\ \ldots\ a_{N,m}(\psi_k)]^T$.

Under the above definitions and assumptions, we can readily verify that $\boldsymbol{x}_m[l]$ also follows a circularly-symmetric complex Gaussian distribution, i.e., $\boldsymbol{x}_m[l] \sim \mathcal{CN}(\boldsymbol{0}, \boldsymbol{R}_m)$, where $\boldsymbol{R}_m$ can be easily derived as

$$\boldsymbol{R}_m = \boldsymbol{A}_m(\boldsymbol{\psi})\mathrm{diag}\,(\boldsymbol{\gamma}_m)\,\boldsymbol{A}_m^H(\boldsymbol{\psi}) + \sigma_m^2\boldsymbol{I}_N. \tag{30}$$

Stacking the received signals of all frequency bins into a vector $\boldsymbol{x}[l] \triangleq [\boldsymbol{x}_1^T[l]\ \boldsymbol{x}_2^T[l]\ \ldots\ \boldsymbol{x}_M^T[l]]^T$, and resorting to (27), we have

$$\boldsymbol{x}[l] \sim \mathcal{CN}(\boldsymbol{0}, \boldsymbol{R}_x), \quad l \in \mathcal{L} \tag{31}$$

where

$$\boldsymbol{R}_x \triangleq \mathrm{blkdiag}\,(\boldsymbol{R}_1, \boldsymbol{R}_2, \ldots, \boldsymbol{R}_M) \tag{32}$$

From (26) and (27), it can be easily verified that

$$E\{\boldsymbol{x}[l_1]\boldsymbol{x}^H[l_2]\} = \boldsymbol{0}, \quad l_1 \neq l_2. \tag{33}$$

Based on the distribution of the received signal shown in (31)–(33), the corresponding CRB analysis can be conducted for the three cases of wideband DoA estimation considered in this paper, i.e., Cases C1–C3, respectively. The details of the derivation are provided in the Appendix.

VII. Simulation Results

For the simulations, we generate the received signal according to the model presented in Section VI, which is a superposition of several harmonics corrupted by additive white Gaussian noise under a finite observation time. The signal-to-noise ratio (SNR) in the simulation is defined as

$$\text{SNR} \triangleq \frac{\sum_{m=1}^{M}\left(\sum_{k=1}^{K}\gamma_{k,m}\right)}{\sum_{m=1}^{M}\sigma_m^2} \tag{34}$$

An unbiased estimate for the covariance matrix $\boldsymbol{R}_m$ can be obtained as

$$\hat{\boldsymbol{R}}_m = \frac{1}{L}\left(\sum_{l=1}^{L}\boldsymbol{x}_m[l]\boldsymbol{x}_m^H[l]\right). \tag{35}$$

For a fair comparison, we assume that the knowledge of the number of source signals K is available to all algorithms in our experiments. For the regularization parameter in our proposed algorithms, we pick the best one through a trial-and-error process.

A. DoA Estimation for Identical Power Spectrum

In this part, for the first experiment we let Assumption A2 hold, i.e., we have $\boldsymbol{\gamma}_m = p_m\boldsymbol{\alpha},\ m \in \mathcal{M}$, while in the second simulation we consider the case where Assumption A2 is slightly violated. The speed of the signal wave is assumed to be $c = 340$ m/s and the wavelength corresponding to the mth frequency f_m is given by $\lambda_m = c/f_m$. Here we consider evenly-spaced frequency points $\{f_m\}_{m=1}^{M}$ over the frequency range of the signal. In addition, we consider a 2-level nested array of 4 antennas at locations $\{d_1, d_2, d_3, d_4\} = \{0, d, 2\,d, 5d\}$ where we set $d = 0.04$ m (corresponding to a frequency of $f = c/(2\,d) = 4.25$ kHz) as the basic element spacing. Note that under this setup, the co-array consists of $N_{\text{co}} = 11$ antennas. To evaluate the performance, we calculate the mean square error (MSE) for the DoAs as

$$\text{MSE}(\boldsymbol{\theta}) \triangleq \sum_{k=1}^{K}|\theta_k - \hat{\theta}_k|^2 \tag{36}$$

and the MSE for the parameter $\boldsymbol{\psi}$, which is defined in (28), as

$$\text{MSE}(\boldsymbol{\psi}) \triangleq \sum_{k=1}^{K}|\psi_k - \hat{\psi}_k|^2. \tag{37}$$

In the first example, we set the number of sources to $K = 5$ with DoAs given by $\{20, 53, 84, 127, 155\}\pi/180$. The frequency range of the sources is from 4 kHz to 20 kHz with $M = 21$, and the mode order of the Jacobi-Anger approximation is set as $N_{\text{virt}} = 85$. All sources share a common normalized power spectrum $\boldsymbol{p}$, as shown in Fig. 2. The received signal power

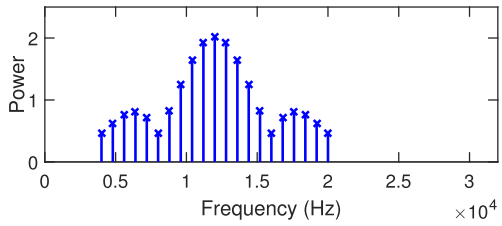
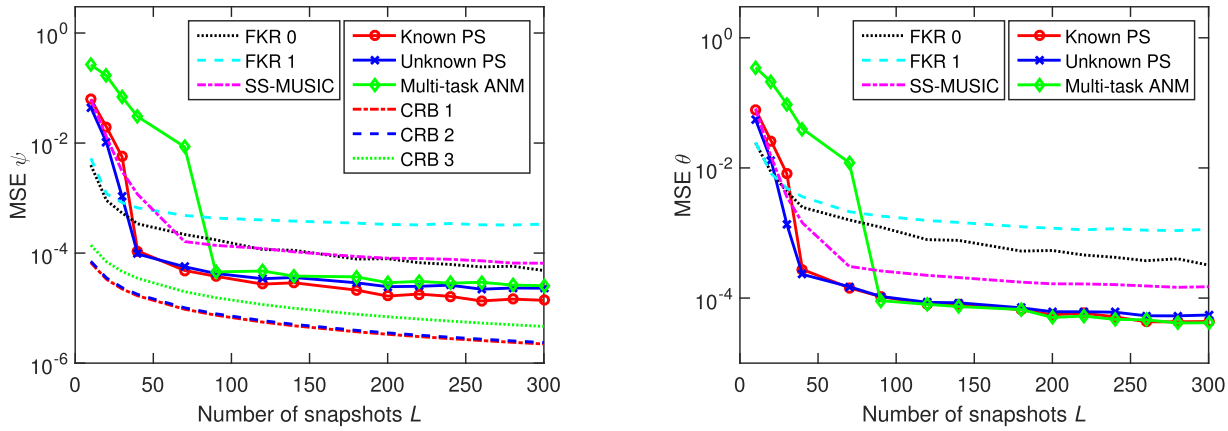


Fig. 2. Real power spectrum of sources, $M = 21$.

is $\boldsymbol{\alpha} = [1.3,\ 2.1,\ 0.5,\ 1.7,\ 0.9]^T$. Depending on whether the power spectrum is known or not, we test the proposed methods in (17) and (19), where the regularization parameters are set as $\lambda_1 = 0.4$ and $\lambda_2 = 1.3$, respectively. The performance of the multi-task ANM method, i.e. (23), is also evaluated with $\lambda_3 = 0.1$. Note that this method is for the general case, and thus the identical normalized power spectrum assumption is not exploited in the experiment. The SS-MUSIC [43] and the FKR subspace approach [44] are also evaluated as benchmarks. To implement the FKR subspace approach, we let f_{m_0}, $m_0 \in \mathcal{M}$, be the reference frequency and apply the rotational signal subspace (RSS) focusing algorithm [27]. The focusing matrix $\boldsymbol{F}_m$ of the mth frequency bin is DoA-dependent, which is obtained through solving the following constrained minimization problem

$$\begin{aligned}\min_{\boldsymbol{F}_m}\quad & \|\boldsymbol{B}_{m_0}(\hat{\boldsymbol{\theta}}) - \boldsymbol{F}_m\boldsymbol{B}_m(\hat{\boldsymbol{\theta}})\|_F^2 \\ s.t.\quad & \boldsymbol{F}_m^H\boldsymbol{F}_m = \boldsymbol{I}\end{aligned} \tag{38}$$

Here $\boldsymbol{B}_m(\boldsymbol{\theta})$ defined in (8) is the co-array response matrix related to the mth frequency, $\|\cdot\|_F$ denotes the Frobenius matrix norm, and $\hat{\boldsymbol{\theta}}$ is an initial estimate of $\boldsymbol{\theta}$. Its solution is given by $\boldsymbol{F}_m = \boldsymbol{U}_m^{(2)}(\boldsymbol{U}_m^{(1)})^H$, where the column vectors of $\boldsymbol{U}_m^{(1)}$ and $\boldsymbol{U}_m^{(2)}$ are the left and right singular vectors of $\boldsymbol{B}_m(\hat{\boldsymbol{\theta}})\boldsymbol{B}_{m_0}^H(\hat{\boldsymbol{\theta}})$, respectively. The initial angles $\hat{\boldsymbol{\theta}}$ have a great impact on the performance of focusing methods due to the focusing error. In this simulation, the initial DoA estimates are set in a similar way as in [29], [44], which consists of a perturbation of the true DoAs by adding Gaussian random noise as $\hat{\boldsymbol{\theta}} \sim \mathcal{N}(\boldsymbol{\theta}, \sigma_\theta^2\boldsymbol{I}_K)$, where σ_θ^2 denotes the estimation variance of the initial DoA estimates. Here we consider the two cases $\sigma_\theta^2 = 0$ and $\sigma_\theta^2 = 0.25$, and the reference frequency is chosen as $f_{m_0} = 4$ kHz. To improve the robustness, we also performed spatial smoothing for FKR before the focusing procedure. Fig. 3 depicts the MSE of the respective sets of parameters vs. the number of snapshots L, where we set SNR $= 10$ dB and the results are averaged over 200 independent runs. Our proposed methods (17), (19) and (23) are respectively referred to as the Known PS case (Case C1), the Unknown PS case (Case C2) and Multi-task ANM (Case C3). The CRB results for Case C1, Case C2 and Case C3, as well as the MSE results from SS-MUSIC and FKR, are also included (FKR 0 and FKR 1 represent the cases $\sigma_\theta^2 = 0$ and $\sigma_\theta^2 = 0.25$, respectively). From Fig. 3(a)-(b), we observe that our proposed methods achieve a higher estimation accuracy than the other methods when $L > 100$, both for parameter $\boldsymbol{\psi}$ and $\boldsymbol{\theta}$. The Known PS and Unknown PS methods even reach a better performance for a much smaller L. From Fig. 3(a),



(a) MSE for $\boldsymbol{\psi}$ vs. L. (b) MSE for $\boldsymbol{\theta}$ vs. L

Fig. 3. MSE vs. the number of snapshots L, $M = 21$ and SNR $= 10$ dB. (The case of identical power spectrum).

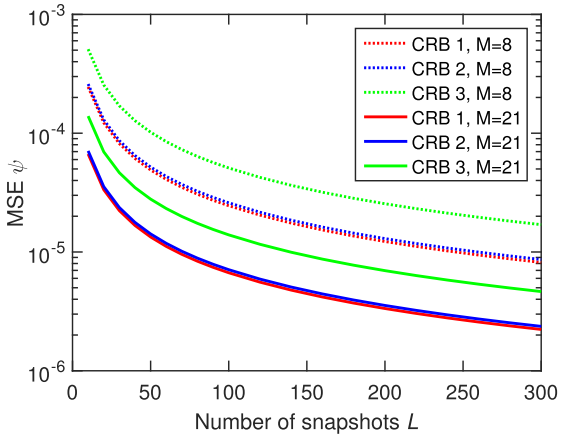


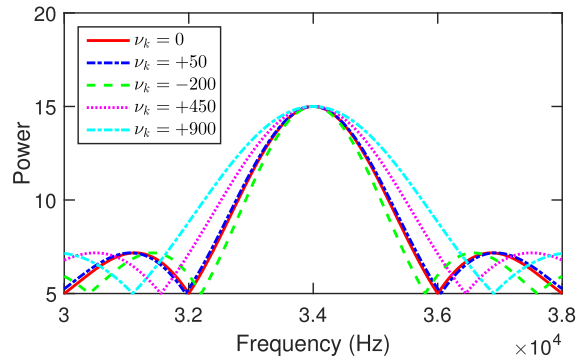
Fig. 4. CRB results vs. the number of snapshots L, SNR $= 10$ dB. (The case of identical power spectrum).

we also see that our proposed methods are not too far away from the CRB. Our techniques even achieve a much higher DoA estimation accuracy than FKR 0, where the true DoAs are used in the simulations to construct the focusing matrices, i.e., $\hat{\boldsymbol{\theta}} = \boldsymbol{\theta}$. This performance improvement is primarily due to the fact that, through the DoA-independent coherent wideband processing, the joint structural knowledge is well utilized in our proposed methods. Furthermore, when Assumption A2 is incorporated, Known PS and Unknown PS have a better performance compared with Multi-task ANM especially for a small value of L. Also, as expected, Known PS provides better DoA estimates than Unknown PS but the gap is very small. To better understand the benefits gained from the spectral diversity, we compute the CRB results for a different number of frequency bins M, as shown in Fig. 4, where the setup remains the same. These CRB results imply that a performance improvement could be achieved when the data from additional frequency bins are included.

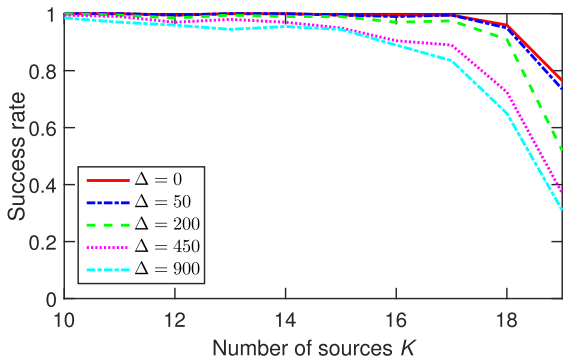
In the second simulation, we illustrate that our proposed method even works in the scenario where the number of DoF of the co-array is less than the number of sources, and the half wavelength spacing condition is violated. Also, in this simulation, we aim to evaluate the robustness of our proposed method in the case where Assumption A2 is slightly violated. In this extreme case, each signal has a common center frequency of $f_c = 34$ kHz and a common bandwidth of 8 kHz. The number of frequency bins is set to $M = 32$. The mode order of the Jacobi-Anger approximation is chosen as $N_{\text{virt}} = 150$, which satisfies (12). We consider the case where $K \in [10, 19]$ uncorrelated wideband sources impinge on the 2-level nested array mentioned above. The DoAs of these sources are uniformly generated over $[0, \pi)$ with an additional constraint on the minimal separation, i.e., $\min_{i,j} |\theta_i - \theta_j| \geq \pi/20$. The number of snapshots is set to $L = 5000$, with SNR $= 20$ dB. The source power related to the kth source and mth frequency, i.e., $\gamma_{k,m}$, is generated as

$$\gamma_{k,m} = \text{sinc}\left(\frac{f_m - f_c}{2000 + \nu_k}\right) \times 10 + 5 \tag{39}$$

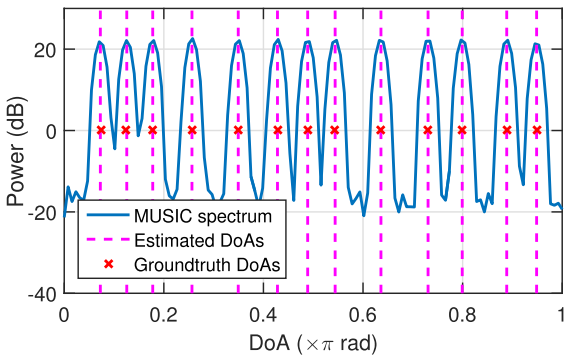
where $\text{sinc}(f) \triangleq \sin(\pi f)/(\pi f)$, and ν_k is a random variable which is uniformly distributed on $(-\Delta, +\Delta)$ with $\Delta \in \mathbb{R}^+$. Clearly, the value of Δ is to measure the "gap" between the identical power spectrum assumption and the real data model, and if $\Delta = 0$, then there is no model mismatch regarding to Assumption A2. The power spectrum of (39) with different values of ν_k is shown in Fig. 5(a). Note that although the power spectrum profile is not strictly identical to all sources here, we still use the formulation proposed in Section IV-B with $\lambda_2 = 0.5$ to recover the matrix $\mathcal{T}(\boldsymbol{u})$, which contains the information of the angles $\{\theta_k\}_{k=1}^K$. The success rate is introduced to evaluate the recovery performance, which is computed as the ratio of the number of successful trials to the total number of independent runs, where $\{\nu_k\}_{k=1}^K$ are randomly generated according to Δ for each run. A trial is considered successful if the maximal estimation error between the estimated DoAs $\{\hat{\theta}_k\}_{k=1}^K$ and the true DoAs $\{\theta_k\}_{k=1}^K$ is smaller than $\pi/150$, i.e., $\max_k |\theta_k - \hat{\theta}_k| < \pi/150$. In Fig. 5(b), we plot the success rates of our proposed method vs. the number of sources K, with different values of Δ. Fig. 5(c) depicts the MUSIC spectrum of $\mathcal{T}(\boldsymbol{u})$, the estimated DoAs and the groundtruth DoAs for a specific run of the experiments in Fig. 5(b), for $\Delta = 0$ and $K = 13$. From



(a)

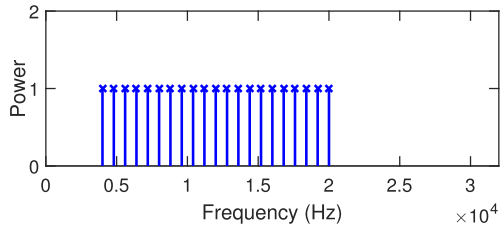


(b)

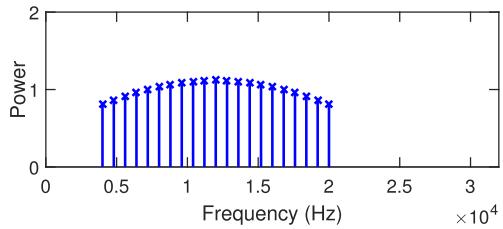


(c)

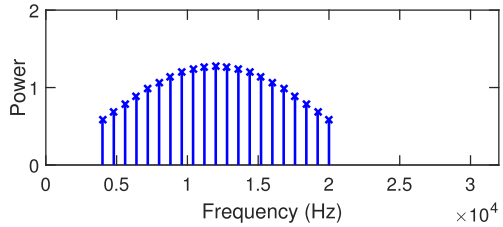
Fig. 5. (a) Power spectrum generated from (39) with different ν_k. (b) Success rates vs. the number of sources K, $M = 32$, $L = 5000$ and SNR $= 20$ dB. (c) MUSIC spectrum, estimated DoAs and groundtruth DoAs, $\Delta = 0$ and $K = 13$.



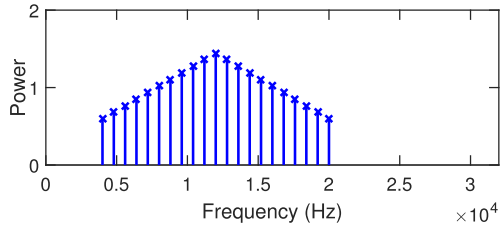
(a)



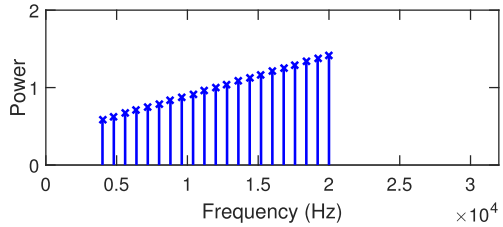
(b)



(c)



(d)

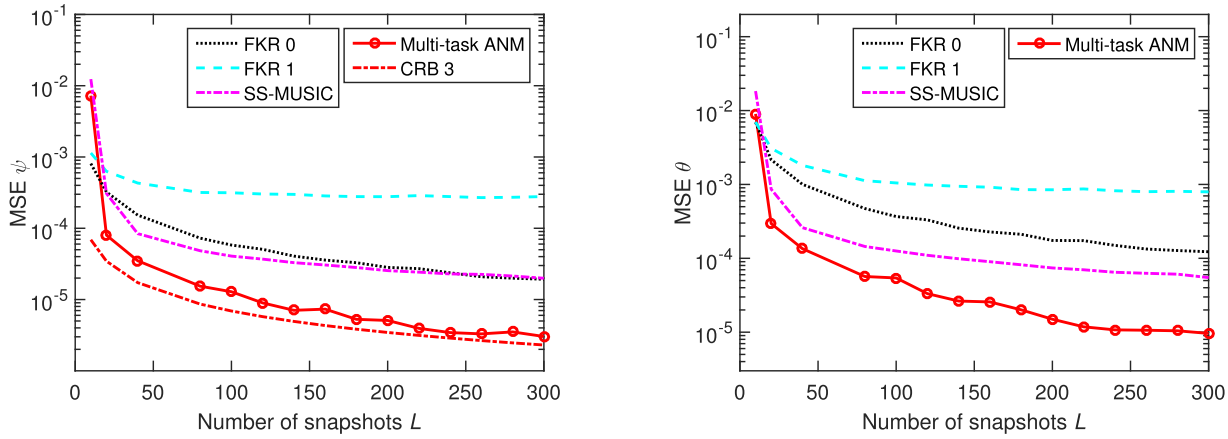


(e)

Fig. 6. (a)–(e) Real power spectrum of 5 sources, $M = 21$.

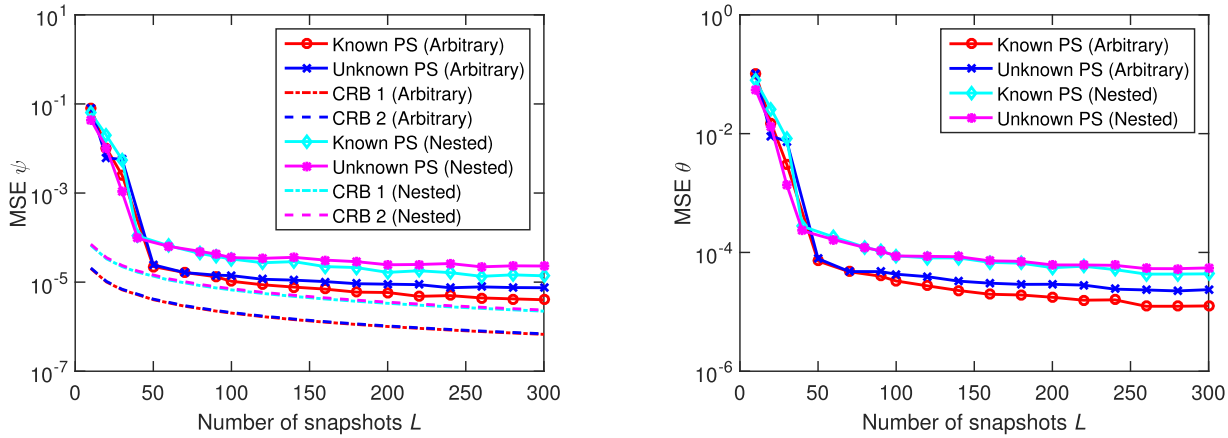
Fig. 5(b) and Fig. 5(c), we can see that, for $\Delta = 0$, although the basic element spacing is larger than half the largest wavelength, and the number of DoF of the difference co-array would be less than the number of sources, i.e., $d > \max_m\{\lambda_m/2\}$ and $N_{\rm co} < K$, our proposed method can still resolve up to $K = 17$ sources. Note that with Assumption A2, we can merge (14) from M frequency bins and construct a joint transformation matrix $\boldsymbol{G}$ (or $\boldsymbol{H}$), which is of size $(MN_{\rm co}) \times (2N_{\rm virt} + 1)$. When the rank of matrix $\boldsymbol{G}$ (or $\boldsymbol{H}$) is larger than $N_{\rm co}$, which always holds with a large M, this indicates that we have the capability of resolving more sources than the DoF of the difference co-array. Furthermore, ambiguities due to spatial aliasing can also be removed when the information from multiple frequency bins is involved, which allows our algorithms to work even if the half wavelength spacing condition is not satisfied. The issue of aliasing-free wideband DoA estimation was also discussed and investigated in [32].

Also, from Fig. 5(b), as expected, a large "gap" between Assumption A2 and the real data model (or a large Δ) would impair our capability of resolving more sources than the number



(a) MSE for $\boldsymbol{\psi}$ vs. L. (b) MSE for $\boldsymbol{\theta}$ vs. L

Fig. 7. MSE vs. the number of snapshots L, $M = 21$ and SNR $= 10$ dB. (General Case).



(a) MSE for $\boldsymbol{\psi}$ vs. L. (b) MSE for $\boldsymbol{\theta}$ vs. L

Fig. 8. MSE vs. the number of snapshots L, $M = 21$ and SNR $= 10$ dB. (Arbitrary array and nested array).

of DoF of the difference co-array. However, we can still resolve the sources when $K \leq 14$ with a large probability, even if the model mismatch has been included. To some extent, this verifies the robustness of our proposed method against model mismatch.

B. DoA Estimation for the General Case

In this simulation, we consider the general case without identical power spectra, obviating Assumption A2. We use the same array configuration as in the first simulation of the previous subsection, where $\{d_1, d_2, d_3, d_4\} = \{0, d, 2\,d, 5d\}$ and $d = 0.04$ m. The speed of the signal wave is assumed to be $c = 340$ m/s. Here we set $M = 21$ and consider evenly-spaced frequency points $\{f_m\}_{m=1}^{M}$ over the frequency range of the signal. The frequency range of the sources is from 4 kHz to 20 kHz. The mode order of the Jacobi-Anger approximation is chosen as $N_{\text{virt}} = 85$, which satisfies (12). We consider the case where $K = 5$ uncorrelated wideband sources impinge on this 2-level nested array. The DoAs of these 5 sources are given by $\{20, 53, 84, 127, 155\}\pi/180$. The power spectrum of each source, i.e., $\bar{\boldsymbol{\gamma}}_k \triangleq [\gamma_{k,1}\ \gamma_{k,2}\ \ldots\ \gamma_{k,M}]^T$, $k = 1, \ldots, K$, is shown in Fig. 6. Our proposed method for this general case is the multi-task ANM formulation (23), where the regularization parameter is set to $\lambda_3 = 1.9$. The CRB as well as the estimation results from the SS-MUSIC and FKR subspace approach are used as benchmarks. For the FKR subspace approach, the focusing matrix of the mth frequency bin is obtained through (38) and the reference frequency is chosen as $f_{m_0} = 4$ kHz.

Fig. 7 depicts the MSE vs. the number of snapshots L, where we set SNR $= 10$ dB and the results are averaged over 200 independent runs. From Fig. 7(a), it can be observed that the proposed method approaches the CRB closely. Also, for the estimation of both $\boldsymbol{\psi}$ and $\boldsymbol{\theta}$, from Fig. 7(a)-(b), we see that our proposed method has a better estimation performance than the other methods.

C. DoA Estimation With Arbitrary Arrays

In the array configuration of the previous two subsections, the antenna locations are constrained to be an integer multiple of the basic element spacing. While this constraint is essential for many traditional DoA estimators, our modeling

and algorithm design do not incur the assumption of basic element spacing. In this simulation, we assume arbitrary antenna locations and fix them to $\{0\,d, 2.1\,d, 5.9\,d, 8.7d\}$, with $d = 0.04$ m. Under this setup, the set of antenna locations of the corresponding difference co-array, which is defined in (6), is $\mathcal{D} = \{\pm 8.7\,d, \pm 6.6\,d, \pm 5.9\,d, \pm 3.8\,d, \pm 2.8\,d, \pm 2.1\,d, 0d\}$. Now the cardinality of $\mathcal{D}$ is 13, which indicates the number of DoF of the co-array. Furthermore, the number of sources is set to $K = 5$ with DoAs given by $\{20, 53, 84, 127, 155\}\pi/180$, the frequency range of the sources is from 4 kHz to 20 kHz with $M = 21$, and the mode order of the Jacobi-Anger approximation is set as $N_{\text{virt}} = 140$. We here adopt the setup in Section VII-A with identical normalized power spectra for all sources. Both the scenarios with and without prior knowledge of the power spectrum are considered (Case C1 and C2), with the regularization parameters $\lambda_1 = 0.4$ and $\lambda_2 = 1$, respectively. For comparison, we also consider a 2-level nested array with 4 antennas to collect the data and estimate the DoAs. In that case, the antenna locations are given by $\{d_1, d_2, d_3, d_4\} = \{0, d, 2\,d, 5d\}$. Fig. 8 plots the MSE as a function of the number of snapshots L, where we set SNR $= 10$ dB. In Fig. 8 (a), the corresponding CRB results for different scenarios are also included. We can see that the arbitrary array significantly outperforms the nested array, given the same number of antennas. This performance improvement is primarily due to the fact that the difference co-array constructed from the array with arbitrary physical locations has more DoF than the regular nested array. A similar concept was considered in [68], where temporal off-the-Nyquist-grid sampling is introduced, which yields more lags or differences for covariance estimation. And as before, Known PS provides better DoA estimates than Unknown PS yet with a small gap.

VIII. Conclusion

In this paper we have studied wideband DoA estimation with SLAs. We resorted to the Jacobi-Anger approximation to transform the difference co-array response matrices of all frequency bins into a single virtual ULA response matrix, which allows us to combine the data from different frequencies easily. In contrast to existing methods, we introduced an optional assumption that all sources share the same normalized power spectrum. Under this assumption, two super-resolution off-the-grid DoA estimation approaches are developed based on ANM, one with and one without prior knowledge of the power spectrum. Surprisingly, these methods can even resolve more sources than the number of DoF of the difference co-array. For the general case where each source has an arbitrary power spectrum, a multi-task ANM method is developed to exploit the joint sparsity from all frequency bins. Simulation results show that, through efficiently merging the information from different frequency subbands, our proposed methods outperform the state of the art and can achieve an estimation accuracy close to the associated CRBs.

Appendix
Derivation of Cramér-Rao Lower Bounds

In the following three subsections, we will provide the corresponding Cramér-Rao bound (CRB) analysis for the three cases of wideband direction of arrival (DoA) estimation considered in this paper, i.e., Cases C1-C3.

A. CRB for Case C1

In Case C1, we assume Assumption A2 holds, which means that all K sources share the same normalized power spectrum $\{p_m\}_{m=1}^M$ and we can write $\gamma_{k,m} = \alpha_k p_m$. We further assume that prior knowledge of $\{p_m\}_{m=1}^M$ is available. Hence, the parameters describing the distribution of the received signal $\boldsymbol{x}_m[l]$ are $\boldsymbol{\psi}$, $\boldsymbol{\alpha}$ and $\{\sigma_m^2\}_{m=1}^M$. We define the complete set of parameters to be estimated in Case C1 as

$$\boldsymbol{\xi}^{(1)} \triangleq \left[\boldsymbol{\psi}^T\ \boldsymbol{\alpha}^T\ \sigma_1^2\ \ldots\ \sigma_M^2\right]^T. \tag{40}$$

Since the random vector $\boldsymbol{x}[l]$ follows a circularly-symmetric complex Gaussian distribution, we can resort to the Slepian-Bangs formula [69], [70] to compute the Fisher information matrix (FIM). According to the Slepian-Bangs formula, and the fact that the $\boldsymbol{x}[l]$, $l \in \mathcal{L}$, are mutually uncorrelated for different segments l, i.e., (33), the (i,j)th element of the FIM $\boldsymbol{\Omega}^{(1)}$ is given by

$$\Omega_{ij}^{(1)} = L \cdot \text{trace}\left(\boldsymbol{R}_x^{-1}\frac{\partial \boldsymbol{R}_x}{\partial \xi_i^{(1)}}\boldsymbol{R}_x^{-1}\frac{\partial \boldsymbol{R}_x}{\partial \xi_j^{(1)}}\right) \tag{41}$$

where $\xi_i^{(1)}$ denotes the ith entry of $\boldsymbol{\xi}^{(1)}$.

We first compute the partial derivative of $\boldsymbol{R}_x$ with respect to ψ_k. From the definition of $\boldsymbol{a}_m(\psi_k)$, we have

$$\frac{\partial \boldsymbol{a}_m(\psi_k)}{\partial \psi_k} = -j \cdot \frac{2\pi}{\lambda_m} \cdot \boldsymbol{D}\boldsymbol{a}_m(\psi_k) \tag{42}$$

where $\boldsymbol{D} \triangleq \text{diag}(d_1, \ldots, d_N)$. Combining the above equation with (30) and (32), we have

$$\frac{\partial \boldsymbol{R}_x}{\partial \psi_k} = \text{blkdiag}\left(\frac{\partial \boldsymbol{R}_1}{\partial \psi_k}, \frac{\partial \boldsymbol{R}_2}{\partial \psi_k}, \ldots, \frac{\partial \boldsymbol{R}_M}{\partial \psi_k}\right) \tag{43}$$

where

$$\begin{aligned}\frac{\partial \boldsymbol{R}_m}{\partial \psi_k} &= \gamma_{k,m} \cdot \left(\frac{\partial \boldsymbol{a}_m(\psi_k)}{\partial \psi_k}\boldsymbol{a}_m^H(\psi_k) + \boldsymbol{a}_m(\psi_k)\frac{\partial \boldsymbol{a}_m^H(\psi_k)}{\partial \psi_k}\right) \\ &= \frac{j2\pi\gamma_{k,m}}{\lambda_m}\left(\boldsymbol{a}_m(\psi_k)\boldsymbol{a}_m^H(\psi_k)\boldsymbol{D} - \boldsymbol{D}\boldsymbol{a}_m(\psi_k)\boldsymbol{a}_m^H(\psi_k)\right)\end{aligned}$$

Similarly, we can obtain the partial derivatives with respect to the other parameters as follows

$$\frac{\partial \boldsymbol{R}_x}{\partial \alpha_k} = \text{blkdiag}\left(\frac{\partial \boldsymbol{R}_1}{\partial \alpha_k}, \frac{\partial \boldsymbol{R}_2}{\partial \alpha_k}, \ldots, \frac{\partial \boldsymbol{R}_M}{\partial \alpha_k}\right) \tag{44}$$

$$\frac{\partial \boldsymbol{R}_x}{\partial \sigma_m^2} = \text{diag}(\boldsymbol{i}_m) \otimes \boldsymbol{I}_N \tag{45}$$

where

$$\frac{\partial \boldsymbol{R}_m}{\partial \alpha_k} = \boldsymbol{a}_m(\psi_k)\boldsymbol{a}_m^H(\psi_k) \cdot p_m \tag{46}$$

and $\boldsymbol{i}_m$ denotes the mth column of $\boldsymbol{I}_M$. After obtaining the FIM $\boldsymbol{\Omega}^{(1)}$, the CRB matrix can be calculated as [65]

$$\text{CRB}^{(1)}(\boldsymbol{\xi}^{(1)}) = \left(\boldsymbol{\Omega}^{(1)}\right)^{-1}. \tag{47}$$

B. CRB for Case C2

As Case C2 relies on Assumption A2, but now $\{p_m\}_{m=1}^M$ is unknown and to be estimated. The distribution of $\boldsymbol{x}_m[l]$ is now characterized by the unknown parameters

$$\boldsymbol{\xi}^{(2)} \triangleq \left[\boldsymbol{\psi}^T\ \boldsymbol{\alpha}^T\ \boldsymbol{p}^T\ \sigma_1^2\ \ldots\ \sigma_M^2\right]^T. \tag{48}$$

Similar to the discussion in the previous subsection, we resort to the Slepian-Bangs formula to obtain the FIM $\boldsymbol{\Omega}^{(2)}$. The partial derivative of $\boldsymbol{R}_x$ with respect to $\boldsymbol{\psi}$, $\boldsymbol{\alpha}$ and σ_m^2, $m \in \mathcal{M}$, has already been computed in (43), (44) and (45), respectively. Here we only need to consider the partial derivative of $\boldsymbol{R}_x$ with respect to $\boldsymbol{p}$. From the definition of $\boldsymbol{R}_x$ in (32), we have

$$\frac{\partial \boldsymbol{R}_x}{\partial p_m} = \mathrm{diag}(\boldsymbol{i}_m) \otimes \left(\boldsymbol{A}_m(\boldsymbol{\psi})\mathrm{diag}(\boldsymbol{\alpha})\boldsymbol{A}_m^H(\boldsymbol{\psi})\right). \tag{49}$$

Similarly as before, after obtaining the FIM $\boldsymbol{\Omega}^{(2)}$, the CRB matrix can be obtained as

$$\mathrm{CRB}^{(2)}(\boldsymbol{\xi}^{(2)}) = \left(\boldsymbol{\Omega}^{(2)}\right)^{-1}. \tag{50}$$

C. CRB for Case C3

For the general case without Assumption A2, we have an arbitrary power spectrum for each source signal. Thus, the complete set of parameters corresponding to the distribution (31) is

$$\boldsymbol{\xi}^{(3)} \triangleq \left[\boldsymbol{\psi}^T\ \boldsymbol{\gamma}_1^T\ \ldots\ \boldsymbol{\gamma}_M^T\ \sigma_1^2\ \ldots\ \sigma_M^2\right]^T. \tag{51}$$

Apparently, if only one frequency bin is available, the problem reduces to the narrowband case. For such a scenario, several closed-form CRB expressions for potentially undetermined DoA estimation have been derived [66], [67], [71]. For the wideband case, we define the narrowband CRB for the mth frequency bin with respect to $\boldsymbol{\psi}$ as $\mathrm{CRB}_m(\boldsymbol{\psi}) \in \mathbb{C}^{K\times K}$, $m \in \mathcal{M}$. Resorting to the block diagonal property of $\boldsymbol{R}_x$ and the inversion of a partitioned matrix, the closed-form CRB expression for the wideband case can be expressed as

$$\mathrm{CRB}^{(3)}(\boldsymbol{\psi}) = \left[\sum_{m=1}^{M} \left(\mathrm{CRB}_m(\boldsymbol{\psi})\right)^{-1}\right]^{-1} \tag{52}$$

REFERENCES

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

[2] R. Schmidt, "Multiple emitter location and signal parameter estimation," *IEEE Trans. Antennas Propag.*, vol. AP-34, no. 3, pp. 276–280, Mar. 1986.

[3] R. Roy and T. Kailath, "ESPRIT-estimation of signal Parameters via rotational invariance techniques," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 37, no. 7, pp. 984–995, Jul. 1989.

[4] D. L. Donoho, "Compressive sensing," *IEEE Trans. Inf. Theory*, vol. 52, no. 4, pp. 1289–1306, Apr. 2006.

[5] D. Malioutov, M. Cetin, and A. S. Willsky, "A sparse signal reconstruction perspective for source localization with sensor arrays," *IEEE Trans. Signal Process.*, vol. 53, no. 8, pp. 3010–3022, Aug. 2005.

[6] P. Stoica, P. Babu, and J. Li, "Spice: A sparse covariance-based estimation method for array processing," *IEEE Trans. Signal Process.*, vol. 59, no. 2, pp. 629–638, Feb. 2011.

[7] M. M. Hyder and K. Mahata, "Direction-of-arrival estimation using a mixed $\ell_{2,0}$ norm approximation," *IEEE Trans. Signal Process.*, vol. 58, no. 9, pp. 4646–4655, Sep. 2010.

[8] E. J. Candès and C. Fernandez-Granda, "Towards a mathematical theory of super-resolution," *Commun. Pure Appl. Math.*, vol. 67, no. 6, pp. 906–956, 2014.

[9] G. Tang, B. N. Bhaskar, P. Shah, and B. Recht, "Compressed sensing off the grid," *IEEE Trans. Inf. Theory*, vol. 59, no. 11, pp. 7465–7490, Nov. 2013.

[10] B. N. Bhaskar, G. Tang, and B. Recht, "Atomic norm denoising with applications to line spectral estimation," *IEEE Trans. Signal Process.*, vol. 61, no. 23, pp. 5987–5999, Dec. 2013.

[11] G. Tang, B. N. Bhaskar, and B. Recht, "Near minimax line spectral estimation," *IEEE Trans. Inf. Theory*, vol. 61, no. 1, pp. 499–512, Jan. 2015.

[12] 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.

[13] Z. Yang and L. Xie, "On gridless sparse methods for line spectral estimation from complete and incomplete data," *IEEE Trans. Signal Process.*, vol. 63, no. 12, pp. 3139–3153, Jun. 2015.

[14] Z. Yang and L. Xie, "Enhancing sparsity and resolution via reweighted atomic norm minimization," *IEEE Trans. Signal Process.*, vol. 64, no. 4, pp. 995–1006, Feb. 2016.

[15] Y. Chen and Y. Chi, "Robust spectral compressed sensing via structured matrix completion," *IEEE Trans. Inf. Theory*, vol. 60, no. 10, pp. 6576–6601, Oct. 2014.

[16] J. Fang, F. Wang, Y. Shen, H. Li, and R. S. Blum, "Super-resolution compressed sensing for line spectral estimation: An iterative reweighted approach," *IEEE Trans. Signal Process.*, vol. 64, no. 18, pp. 4649–4662, Sep. 2016.

[17] Q. Li and G. Tang, "Approximate support recovery of atomic line spectral estimation: A tale of resolution and precision," in *Proc. IEEE Global Conf. Signal Inf. Process.*, Washington, DC, USA, 2016, pp. 153–156.

[18] Y. Chi and M. F. D. Costa, "Harnessing sparsity over the continuum: Atomic norm minimization for superresolution," *IEEE Signal Process. Mag.*, vol. 37, no. 2, pp. 39–57, Mar. 2020.

[19] M. F. Da Costa and Y. Chi, "On the stable resolution limit of total variation regularization for spike deconvolution," *IEEE Trans. Inf. Theory*, vol. 66, no. 11, pp. 7237–7252, Nov. 2020.

[20] S. Li, H. Mansour, and M. B. Wakin, "Recovery analysis of damped spectrally sparse signals and its relation to MUSIC," *Inf. Inference: A J. IMA,* 2020.

[21] S. Li, D. Yang, G. Tang, and M. B. Wakin, "Atomic norm minimization for modal analysis from random and compressed samples," *IEEE Trans. Signal Process.*, vol. 66, no. 7, pp. 1817–1831, Apr. 2018.

[22] Y. Chi and Y. Chen, "Compressive two-dimensional harmonic retrieval via atomic norm minimization," *IEEE Trans. Signal Process.*, vol. 63, no. 4, pp. 1030–1042, Feb. 2015.

[23] Z. Yang, L. Xie, and P. Stoica, "Vandermonde decomposition of multilevel toeplitz matrices with application to multidimensional super-resolution," *IEEE Trans. Inf. Theory*, vol. 62, no. 6, pp. 3685–3701, Jun. 2016.

[24] Z. Zhang, Y. Wang, and Z. Tian, "Efficient two-dimensional line spectrum estimation based on decoupled atomic norm minimization," *Signal Process.*, vol. 163, pp. 95–106, 2019.

[25] M. Wax, T.-J. Shan, and T. Kailath, "Spatio-temporal spectral analysis by eigenstructure methods," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. ASSP-32, no. 4, pp. 817–827, Aug. 1984.

[26] H. Wang and M. Kaveh, "Coherent signal-subspace processing for the detection and estimation of angles of arrival of multiple wide-band sources," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. ASSP-33, no. 4, pp. 823–831, Aug. 1985.

[27] H. Hung and M. Kaveh, "Focussing matrices for coherent signal-subspace processing," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 36, no. 8, pp. 1272–1281, Aug. 1988.

[28] E. D. D. Claudio and R. Parisi, "WAVES: Weighted average of signal subspaces for robust wideband direction finding," *IEEE Trans. Signal Process.*, vol. 49, no. 10, pp. 2179–2191, Oct. 2001.

[29] Y.-S. Yoon, L. M. Kaplan, and J. H. McClellan, "TOPS: New DOA estimator for wideband signals," *IEEE Trans. Signal Process.*, vol. 54, no. 6, pp. 1977–1989, Jun. 2006.

[30] P. Pal and P. Vaidyanathan, "A novel autofocusing approach for estimating directions-of-arrival of wideband signals," in *Proc. Conf. Rec. 43rd Asilomar Conf. Signals Syst. Comput.*, 2009, pp. 1663–1667.

[31] Y. Jiang, D. Li, X. Wu, and W.-P. Zhu, "A gridless wideband DOA estimation based on atomic norm minimization," in *Proc. IEEE Sensor Array Multichannel Signal Process. Workshop*, Hangzhou, China, 2020, pp. 1–5.

[32] Z. Tang, G. Blacquière, and G. Leus, "Aliasing-free wideband beamforming using sparse signal representation," *IEEE Trans. Signal Process.*, vol. 59, no. 7, pp. 3464–3469, Jul. 2011.
[33] L. Wang, L. Zhao, G. Bi, C. Wan, L. Zhang, and H. Zhang, "Novel wideband DOA estimation based on sparse Bayesian learning with Dirichlet process priors," *IEEE Trans. Signal Process.*, vol. 64, no. 2, pp. 275–289, Jan. 2016.
[34] W.-K. Ma, T.-H. Hsieh, and C.-Y. Chi, "DOA estimation of quasi-stationary signals with less sensors than sources and unknown spatial noise covariance: A Khatri-Rao subspace approach," *IEEE Trans. Signal Process.*, vol. 58, no. 4, pp. 2168–2180, Apr. 2010.
[35] 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.
[36] 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.
[37] D. Romero, D. D. Arianananda, Z. Tian, and G. Leus, "Compressive covariance sensing: Structure-based compressive sensing beyond sparsity," *IEEE Signal Process. Mag.*, vol. 33, no. 1, pp. 78–93, Jan. 2016.
[38] A. T. Moffet, "Minimum-redundancy linear arrays," *IEEE Trans. Antennas Propag.*, vol. AP-16, no. 2, pp. 172–175, Mar. 1968.
[39] S. Shakeri, D. D. Arianananda, and G. Leus, "Direction of arrival estimation using sparse ruler array design," in *Proc. IEEE Int. Workshop Signal Process. Adv. Wireless Commun.*, Cesme, Turkey, 2012, pp. 525–529.
[40] P. Pal and P. P. Vaidyanathan, "A grid-less approach to underdetermined direction of arrival estimation via low rank matrix denoising," *IEEE Signal Process. Lett.*, vol. 21, no. 6, pp. 737–741, Jun. 2014.
[41] C. Zhou, Y. Gu, X. Fan, Z. Shi, G. Mao, and Y. D. Zhang, "Direction-of-arrival estimation for coprime array via virtual array interpolation," *IEEE Trans. Signal Process.*, vol. 66, no. 22, pp. 5956–5971, Nov. 2018.
[42] M. Wang, Z. Zhang, and A. Nehorai, "Grid-less DOA estimation using sparse linear arrays based on Wasserstein distance," *IEEE Signal Process. Lett.*, vol. 26, no. 6, pp. 838–842, Jun. 2019.
[43] K. Han and A. Nehorai, "Wideband Gaussian source processing using a linear nested array," *IEEE Signal Process. Lett.*, vol. 20, no. 11, pp. 1110–1113, Nov. 2013.
[44] D. Feng, M. Bao, Z. Ye, L. Guan, and X. Li, "A novel wideband DOA estimator based on Khatri-Rao subspace approach," *Signal Process.*, vol. 91, no. 10, pp. 2415–2419, Oct. 2011.
[45] Z.-Q. He, Z.-P. Shi, L. Huang, and H. C. So, "Underdetermined DOA estimation for wideband signals using robust sparse covariance fitting," *IEEE Signal Process. Lett.*, vol. 22, no. 4, pp. 435–439, Apr. 2015.
[46] Q. Shen, W. Liu, W. Cui, S. Wu, Y. D. Zhang, and M. G. Amin, "Focused compressive sensing for underdetermined wideband DOA estimation exploiting high-order difference coarrays," *IEEE Signal Process. Lett.*, vol. 24, no. 1, pp. 86–90, Jan. 2017.
[47] Y. Shi, X.-P. Mao, C. Zhao, and Y.-T. Liu, "Underdetermined DOA estimation for wideband signals via joint sparse signal reconstruction," *IEEE Signal Process. Lett.*, vol. 26, no. 10, pp. 1541–1545, Oct. 2019.
[48] Q. Shen, W. Cui, W. Liu, S. Wu, and M. G. Amin, "Underdetermined wideband DOA estimation of off-grid sources employing the difference co-array concept," *Signal Process.*, vol. 130, pp. 299–304, 2017.
[49] M. A. Doron, E. Doron, and A. J. Weiss, "Coherent wide-band processing for arbitrary array geometry," *IEEE Trans. Signal Process.*, vol. 41, no. 1, pp. 414–417, Jan. 1993.
[50] F. Belloni, A. Richter, and V. Koivunen, "DOA estimation via manifold separation for arbitrary array structures," *IEEE Trans. Signal Process.*, vol. 55, no. 10, pp. 4800–4810, Oct. 2007.
[51] J. Zhuang, C. Duan, W. Wang, and Z. Chen, "Joint estimation of azimuth and elevation via manifold separation for arbitrary array structures," *IEEE Trans. Veh. Technol.*, vol. 67, no. 7, pp. 5585–5596, Jul. 2018.
[52] J. Zhuang, H. Xiong, W. Wang, and Z. Chen, "Application of manifold separation to parametric localization for incoherently distributed sources," *IEEE Trans. Signal Process.*, vol. 66, no. 11, pp. 2849–2860, Jun. 2018.
[53] D. Baron, M. F. Duarte, M. B. Wakin, S. Sarvotham, and R. G. Baraniuk, "Distributed compressive sensing," 2009, *arXiv:0901.3403*.
[54] Y. Pan, G. Q. Luo, Z. Liao, B. Cai, and M. Yao, "Wideband direction-of-arrival estimation with arbitrary array via coherent annihilating," *IEEE Access*, vol. 7, pp. 51058–51068, 2019.
[55] R. E. Sheriff and Y. F. Hu, *Mobile Satellite Communication Networks*. Hoboken, NJ, USA: John Wiley & Sons, 2003.
[56] B. Vucetic and J. Du, "Channel modeling and simulation in satellite mobile communication systems," *IEEE J. Sel. Areas Commun.*, vol. 10, no. 8, pp. 1209–1218, Oct. 1992.
[57] R. H. Tüütüncü, K. C. Toh, and M. J. Todd, "Solving Semidefinite-quadratic-linear programs using SDPT3," *Math. Program.*, vol. 95, no. 2, pp. 189–217, 2003.
[58] B. D. Rao and K. V. S. Hari, "Performance analysis of root-MUSIC," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 37, no. 12, pp. 1939–1949, Dec. 1989.
[59] C. M. Bishop, *Pattern Recognition and Machine Learning*. Berlin, Germany: Springer, 2006.
[60] Z. Yang and L. Xie, "Exact joint sparse frequency recovery via optimization methods," *IEEE Trans. Signal Process.*, vol. 64, no. 19, pp. 5145–5157, Oct. 2016.
[61] Y. Li and Y. Chi, "Off-the-grid line spectrum denoising and estimation with multiple measurement vectors," *IEEE Trans. Signal Process.*, vol. 64, no. 5, pp. 1257–1269, Mar. 2016.
[62] S. Ji, D. Dunson, and L. Carin, "Multitask compressive sensing," *IEEE Trans. Signal Process.*, vol. 57, no. 1, pp. 92–106, Jan. 2009.
[63] S. Boyd *et al.*, "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Foundations Trends Mach. Learn.*, vol. 3, no. 1, pp. 1–122, 2011.
[64] Y. Wang, Y. Zhang, Z. Tian, G. Leus, and G. Zhang, "Super-resolution channel estimation for arbitrary arrays in hybrid millimeter-wave massive mimo systems," *IEEE J. Sel. Topics Signal Process.*, vol. 13, no. 5, pp. 947–960, Sep. 2019.
[65] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*. Upper Saddle River, NJ: Prentice Hall, 1993.
[66] C.-L. Liu and 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.
[67] A. Koochakzadeh and P. Pal, "Cramér-Rao bounds for underdetermined source localization," *IEEE Signal Process. Lett.*, vol. 23, no. 7, pp. 919–923, Jul. 2016.
[68] D. Cohen, Y. C. Eldar, and G. Leus, "Universal lower bounds on sampling rates for covariance estimation," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process.*, 2015, pp. 3272–3276.
[69] P. Stoica and A. Nehorai, "Performance study of conditional and unconditional direction-of-arrival estimation," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 38, no. 10, pp. 1783–1795, Oct. 1990.
[70] P. Stoica, E. G. Larsson, and A. B. Gershman, "The stochastic CRB for array processing: A textbook derivation," *IEEE Signal Process. Lett.*, vol. 8, no. 5, pp. 148–150, May 2001.
[71] 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.

Feiyu Wang received the B.Sc. degree from Zhejiang Ocean University, China, the M.Sc. degree from Ningbo University, China, and the Ph.D. degree from the University of Electronic Science and Technology of China (UESTC), China, all in electrical engineering. He was a Postdoctoral Researcher with the Department of Microelectronics, Delft University of Technology, the Netherlands, from January 2019 to October 2020. His current research interests include compressed sensing, tensor analysis, and statistical signal processing.

Zhi Tian (Fellow, IEEE) has been a professor with the Electrical and Computer Engineering Department of George Mason University since 2015. Prior to that, she was on the faculty of Michigan Tech University from 2000 to 2014. She served as a program director with the U.S. National Science Foundation from 2012 to 2014. Her current research focuses on massive MIMO, distributed network optimization and machine learning. She was an IEEE Distinguished Lecturer for both the IEEE Communications Society and the IEEE Vehicular Technology Society. She served as Associate Editor for IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS and IEEE TRANSACTIONS ON SIGNAL PROCESSING. She is a Member-at-Large of the IEEE Signal Processing Society (2019-2021). She received the IEEE Communications Society TCCN Publication Award in 2018.

Geert Leus (Fellow, IEEE) received the M.Sc. and Ph.D. degree in electrical engineering from the KU Leuven, Belgium, in June 1996 and May 2000, respectively. Currently, Geert Leus is a Full Professor at the Faculty of Electrical Engineering, Mathematics and Computer Science of the Delft University of Technology, The Netherlands. Geert Leus received the 2021 EURASIP Individual Technical Achievement Award, a 2005 IEEE Signal Processing Society Best Paper Award, and a 2002 IEEE Signal Processing Society Young Author Best Paper Award. He is a Fellow of the IEEE and a Fellow of EURASIP. Geert Leus was a Member-at-Large of the Board of Governors of the IEEE Signal Processing Society, the Chair of the IEEE Signal Processing for Communications and Networking Technical Committee, and the Editor in Chief of the EURASIP *Journal on Advances in Signal Processing*. Currently, he is the Chair of the EURASIP Technical Area Committee on Signal Processing for Multisensor Systems and the Editor in Chief of EURASIP *Signal Processing*.

Jun Fang (Senior Member, IEEE) received the B.S. and M.S. degrees from the Xidian University, Xi'an, China in 1998 and 2001, respectively, and the Ph.D. degree from the National University of Singapore, Singapore, in 2006, all in electrical engineering. During 2006, he was a postdoctoral research associate in the Department of Electrical and Computer Engineering, Duke University. From January 2007 to December 2010, he was a research associate with the Department of Electrical and Computer Engineering, Stevens Institute of Technology. Since 2011, he has been with the University of Electronic Science and Technology of China (UESTC). His research interests include compressed sensing and sparse theory, massive MIMO/mmWave communications, and statistical inference.

Dr. Fang received the IEEE Jack Neubauer Memorial Award in 2013 for the best systems paper published in the IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY. He serves as a Senior Associate Editor for IEEE SIGNAL PROCESSING LETTERS.