A Method for Reducing the Performance Gap Between Non-Coherent and Coherent Sub-Arrays

Tom Tirer and Oded Bialer

***Abstract*—We consider the problem of estimating the angles of arrival (AOAs) of multiple sources from a single snapshot obtained by a set of non-coherent sub-arrays, i.e., while the antenna elements in each sub-array are coherent, each sub-array observes a different unknown phase offset. Previous relevant works are based on eigendecomposition of the sample covariance, which requires a large number of snapshots, or on combining the sub-arrays using non-coherent processing. In this paper, we propose a technique to estimate the sub-arrays phase offsets for a given AOAs hypothesis, which facilitates approximate maximum likelihood estimation (MLE) of the AOAs from a single snapshot. We derive the Cramér-Rao lower bound (CRLB) for the problem at hand, and analytically show that for a single source it may suffice to use a simple non-coherent AOA estimation. However, as we demonstrate by computer simulations, for multiple sources the proposed approach clearly outperforms the non-coherent estimator, and even attains the CRLB in various scenarios. Furthermore, the performance of the proposed method is often close to the performance of MLE in the coherent case, and the gap between the estimators is unavoidable, as implied by the gap between the CRLB for the coherent and non-coherent cases.**

***Index Terms*—Angle of arrival, array processing, Cramér-Rao lower bound, maximum likelihood estimation, single snapshot.**

I. Introduction

The problem of estimating the angles of arrival (AOAs) of radio frequency sources using an array of antenna elements attracts much interest in different disciplines, such as signal processing, communications, vehicular technology, and underwater acoustics [1]. In these fields, it is beneficial to use an array with large aperture, since the estimation accuracy is improved when the array's elements are coherent and spread over a large area [2], [3]. *Coherent elements* are elements that are connected to the same local oscillator. Therefore, ignoring the noise, the phase differences between the observations of these elements are only due to the difference in the propagation delays between the location of the sources and the location of the elements.

Maintaining the coherence of all the elements is very demanding, especially at high carrier frequency, as it requires expensive hardware and a complicated calibration process which is prone to errors. Therefore, a promising alternative is to split a large array into several smaller sub-arrays, such that the coherence of the elements is kept *only within each sub-array*, while the large area captured by the union of these sub-arrays is exploited for increasing the angular resolution using signal processing techniques.

The authors are with General Motors - Advanced Technical Center Israel. T. Tirer is also with the School of Electrical Engineering, Tel Aviv University, Tel Aviv 69978, Israel. (email: tirer.tom@gmail.com, oded.bialer8@gmail.com)

Throughout this paper, the term *non-coherent sub-arrays* will be used to describe the latter physical setting, where the elements of each sub-array are coherent, but there are unknown phase offsets between different sub-arrays due to unsynchronized local oscillators. On the other hand, the term *coherent sub-arrays* will be used for the physical setting where the elements of the entire array (including all sub-arrays) are coherent.

Several previous works have considered estimating the AOAs of a known number of sources from non-coherent sub-arrays [4]–[10], or from the somewhat related model of partially calibrated array [11], [12]. These methods can be divided into two approaches. One approach, used in [7], [9], [10], is based on *non-coherent processing*. It assumes that each sub-array observes a different signal of the sources, which is translated to treating the observations of different sub-arrays as if they are independent realizations. The sub-arrays realizations can therefore be used to estimate the AOAs by likelihood arguments—under the independence assumption [10], or, if the sub-arrays have similar sensor configuration, by subspace-based methods like MUSIC [13] and ESPRIT [14]. It is important to note that while non-coherent processing does not require estimation of the nuisance phase offsets parameters, it does not use the phase information between the sub-arrays, and thus does not exploit the potential aperture of the entire array.

The second approach, used in [4]–[6], [8], [11], [12], relies on having a large number of snapshots (temporal observations) with the *same* phase offsets between the sub-arrays. These methods estimate the AOAs using subspace methods (e.g. variants of MUSIC [13] and root-MUSIC [15]) applied on the sample covariance of the entire array, with modifications that allow to get rid of the nuisance phase offsets parameters. Note, though, that when the sub-arrays are not coherent due to unsynchronized local oscillators, they have different phase offsets in each snapshot. Therefore, methods that require multiple snapshots that share the same phase offsets cannot be used.

There are other practical applications where it is required to estimate the AOAs from a single realization of the measurements. For example, this is the case when the sources and/or receivers are moving. Previous works that dealt with estimating the AOAs of multiple sources from a single snapshot are based on minimization (often greedy) of the negative log-likelihood function [16]–[18], sparse signal reconstructions [19]–[21], atomic norm minimization [22]–[24], or, recently, on machine learning techniques [25]. However, these works focus on coherent arrays, and we have not found their extension to the discussed non-coherent problem.

Lastly, we note that there is another line of work that considers source localization from spatially spread (sub-)arrays that observe *different* AOAs, e.g. see [26]–[31]. In contrast, here we consider co-located sub-arrays, i.e. the sources are at far-field with respect to the array aperture that includes all the sub-arrays. Therefore, as assumed also in [4]–[12], the AOA of each source is the *same* at all the sub-arrays.

Contribution. In this paper, we propose a novel method to obtain closed-form estimators for the unknown phase offsets of the different sub-arrays, which are functions of the AOAs hypothesis. Substituting the phase estimators into the likelihood function makes it depend only on the AOAs, which facilitates (approximate) maximum likelihood estimation (MLE) of the AOAs from a single snapshot (our preliminary results appeared in [32]). The proposed method is the first to exploit the entire array aperture, composed of all the non-coherent sub-arrays, from a single snapshot. We also derive the Cramér-Rao lower bound (CRLB) on the AOAs estimation error for non-coherent sub-arrays. We use the CRLB for evaluating the performance of the proposed method and for gaining insights on the performance gap from the case that the sub-arrays are coherent. We highlight situations in which the performance gap is small and thus the effort of maintaining a fully coherent RF system may be spared, as well as situations in which the gap is large and the RF synchronization effort is worthy.

We present several interesting results. First, we show, both empirically and analytically, that for a single source the CRLB (of the discussed non-coherent scenario) can be attained by an AOA estimator that is based on non-coherent processing. Therefore, there is no need to try to estimate the phase offsets. Second, we empirically show that for multiple sources the proposed approach is better than AOAs estimation with non-coherent processing by a large margin, and can even attain the CRLB in various scenarios. Third, we show that (perhaps surprisingly) in various scenarios the gap between the CRLB for the cases of coherent and non-coherent sub-arrays is very small, and accordingly, the performance of the proposed approach is very close to the performance of MLE when the entire array (including all sub-arrays) is coherent.

The remainder of the paper is organized as follows. Section II defines the problem at hand. Section III reviews a common AOA estimator based on non-coherent processing. In Section IV we present the proposed AOA estimation method. In Section V we derive the CRLB for the non-coherent problem and discuss the single source case. Section VI provides simulation results of our method, compared to non-coherent estimation, the estimation in the coherent case, and the CRLB. Finally, Section VII concludes the paper.

II. Problem Formulation

Consider Q far-field[1] sources, located at unknown angles $\boldsymbol{\theta} \triangleq [\theta_1, \ldots, \theta_Q]^T$ and transmitting unknown narrowband signals that impinge on an M-element receiving array. Let the array be composed of L sub-arrays, where the ℓth sub-array consists of M_ℓ elements ($\sum_{\ell=1}^{L} M_\ell = M$). We assume that each sub-array is coherent, while different sub-arrays are non-coherent, i.e. there is an unknown phase shift between sub-arrays. Furthermore, we assume that only a single snapshot is available.

The waveform observed by each of the L sub-arrays can be formulated as

$$\boldsymbol{x}_\ell = \mathrm{e}^{-j\phi_\ell} \boldsymbol{A}_\ell(\boldsymbol{\theta})\boldsymbol{s} + \boldsymbol{n}_\ell, \quad \ell = 1, \ldots, L, \tag{1}$$

where ϕ_ℓ is the unknown phase shift in the ℓth sub-array, the $M_\ell \times Q$ matrix $\boldsymbol{A}_\ell(\boldsymbol{\theta}) \triangleq [\mathbf{a}_\ell(\theta_1), \ldots, \mathbf{a}_\ell(\theta_Q)]$ is the ℓth sub-array manifold, where each column $\mathbf{a}_\ell(\theta)$ is the ℓth sub-array response to a signal which arrives at angle θ. The $Q \times 1$ vector $\boldsymbol{s}$ represents the unknown signal of the sources, and the $M_\ell \times 1$ vector $\boldsymbol{n}_\ell$ represents white, zero-mean, circular complex Gaussian noise. To simplify the formulation, we assume that the noise variance σ^2 is known, and equal at all receivers (the extension to non equal variances is straightforward). Note that without loss of generality we can assume that $\phi_1 = 0$, since any other phase can be included in the unknown vector $\boldsymbol{s}$.

Note that (1) covers arbitrary sub-array manifolds in planar geometry (i.e. with no height axis). Formally, assuming that the axis origin (i.e. the reference point) is defined in the middle of the entire array, and that the source angle is measured with respect to the boresight, then the response of the ith element in the ℓth sub-array to a signal which arrives at angle θ is given by

$$[\mathbf{a}_\ell(\theta)]_i = g_{\ell,i}(\theta)\mathrm{e}^{j\frac{2\pi}{\lambda}(x_{\ell,i}\sin\theta + y_{\ell,i}\cos\theta)}, \tag{2}$$

where λ is the signal wavelength, $(x_{\ell,i}, y_{\ell,i})$ denotes the antenna location, and $g_{\ell,i}$ denotes the antenna element pattern, which is assumed to be known. For example, for linear arrays of omnidirectional sensors we have that $g_{\ell,i}(\theta) = 1$ and $y_{\ell,i} = 0$.

Let us also further explain the applicability of our observation model. We assume that the sources are at far-field with respect to the array aperture that includes *all* the sub-arrays (as assumed also in [4]–[12]). The far field condition is that the distance of each source from the antenna array is larger than $2D^2/\lambda$, where D is the full array aperture. Under these conditions, the AOA of each source is the same at all the antenna elements, and the difference in the distances between each source and the antennas is on the order of the wavelength. For (typical) moderate to high carrier frequencies this implies that the distance differences are small and hence the amplitude differences due to the delay differences are negligible. With these model assumptions, the amplitude of the source signal at all the sub-arrays is considered the same. Therefore, the only variations between the sub-arrays, which are not due to the setting's geometry and the noise, are the phase offsets that result from each sub-array having a different local oscillator.

To conclude this section, let us formulate the problem. Collecting the observations from all sub-arrays, we obtain

$$\boldsymbol{x} = \boldsymbol{\Lambda}(\{\phi_\ell\})\boldsymbol{A}(\boldsymbol{\theta})\boldsymbol{s} + \boldsymbol{n}, \tag{3}$$

where

$$\boldsymbol{x} \triangleq \begin{bmatrix} \boldsymbol{x}_1 \\ \vdots \\ \boldsymbol{x}_L \end{bmatrix}, \boldsymbol{A}(\boldsymbol{\theta}) \triangleq \begin{bmatrix} \boldsymbol{A}_1(\boldsymbol{\theta}) \\ \vdots \\ \boldsymbol{A}_L(\boldsymbol{\theta}) \end{bmatrix}, \boldsymbol{n} \triangleq \begin{bmatrix} \boldsymbol{n}_1 \\ \vdots \\ \boldsymbol{n}_L \end{bmatrix},$$

[1] Below, we discuss in detail the far-field assumption.

$$\boldsymbol{\Lambda}(\{\phi_\ell\}) \triangleq \begin{bmatrix} \mathrm{e}^{-j\phi_1}\boldsymbol{I}_{M_1} & & \boldsymbol{0} \\ & \ddots & \\ \boldsymbol{0} & & \mathrm{e}^{-j\phi_L}\boldsymbol{I}_{M_L} \end{bmatrix},$$

and $\boldsymbol{I}_{M_\ell}$ denotes the $M_\ell \times M_\ell$ identity matrix. The problem at hand is to estimate the AOAs $\boldsymbol{\theta}$ from the observations given in (3), where the nuisance parameters are $\{\phi_\ell\}$ and $\boldsymbol{s}$. Note that $\boldsymbol{s}$ represents unknown *deterministic* signal, i.e. we make no statistical assumptions on $\boldsymbol{s}$.

III. Non-Coherent Processing

Before we present our approach, let us briefly discuss a common method to estimate $\boldsymbol{\theta}$ using non-coherent processing. This method will be used for comparisons later in the paper.

Non-coherent processing assumes, or simplifies the problem by assuming, that each sub-array observes a different ("independent") signal of the sources $\boldsymbol{s}_\ell$ instead of having the same signal $\boldsymbol{s}$ shared by all the sub-arrays (cf. (1)). Since in this approach a phase shift ϕ_ℓ cannot be identified from the multiplication $\mathrm{e}^{-j\phi_\ell}\boldsymbol{s}_\ell$ where both ϕ_ℓ and $\boldsymbol{s}_\ell$ are unknown, there is no necessity (or hope) for estimating the phase shifts $\{\phi_\ell\}$. Therefore, by absorbing the unknown phase shift ϕ_ℓ into the unknown signal $\boldsymbol{s}_\ell$ the following observation model is used instead of (1)

$$\boldsymbol{x}_\ell = \boldsymbol{A}_\ell(\boldsymbol{\theta})\boldsymbol{s}_\ell + \boldsymbol{n}_\ell, \quad \ell = 1, \ldots, L, \tag{4}$$

where $\{\boldsymbol{x}_\ell\}$ are statistically independent, and each one of them is a circular complex Gaussian with mean $\boldsymbol{A}_\ell(\boldsymbol{\theta})\boldsymbol{s}_\ell$ and variance $\sigma^2\boldsymbol{I}_{M_\ell}$. Therefore, a natural way to estimate $\boldsymbol{\theta}$ from $\{\boldsymbol{x}_\ell\}$ is based on minimizing the following cost function

$$f_{\text{non-coh}}(\boldsymbol{\theta}, \{\boldsymbol{s}_\ell\}) \triangleq \sum_{\ell=1}^{L} \left\|\boldsymbol{x}_\ell - \boldsymbol{A}_\ell(\boldsymbol{\theta})\boldsymbol{s}_\ell\right\|^2, \tag{5}$$

where $\|\cdot\|$ stands for the Euclidean norm. This function is related to the negative log-likelihood function *under the assumption that* $\{\boldsymbol{s}_\ell\}$ *are independent unknown variables*. A similar likelihood approach has been used in [10]. However, we emphasize that in the non-coherent setting that is considered in our paper (and stated in Section II) the source signals that are observed by different sub-array are the same. Therefore, $f_{\text{non-coh}}$ in (5) *does not* correspond to the negative log-likelihood function for the problem addressed in this paper.

Let us proceed with the minimization of (5) with respect to each $\boldsymbol{s}_\ell$. This yields

$$\hat{\boldsymbol{s}}_\ell = \left(\boldsymbol{A}_\ell^H(\boldsymbol{\theta})\boldsymbol{A}_\ell(\boldsymbol{\theta})\right)^{-1}\boldsymbol{A}_\ell^H(\boldsymbol{\theta})\boldsymbol{x}_\ell, \tag{6}$$

and by substituting this result into (5), we get

$$\begin{aligned} f_{\text{non-coh}}(\boldsymbol{\theta}) &= \sum_{\ell=1}^{L} \left\|(\boldsymbol{I}_{M_\ell} - \boldsymbol{P}_{\boldsymbol{A}_\ell(\boldsymbol{\theta})})\boldsymbol{x}_\ell\right\|^2 \\ &= \sum_{\ell=1}^{L} \left\|\boldsymbol{Q}_{\boldsymbol{A}_\ell(\boldsymbol{\theta})}\boldsymbol{x}_\ell\right\|^2, \end{aligned} \tag{7}$$

where

$$\boldsymbol{P}_{\boldsymbol{A}_\ell(\boldsymbol{\theta})} \triangleq \boldsymbol{A}_\ell(\boldsymbol{\theta})\left(\boldsymbol{A}_\ell^H(\boldsymbol{\theta})\boldsymbol{A}_\ell(\boldsymbol{\theta})\right)^{-1}\boldsymbol{A}_\ell^H(\boldsymbol{\theta}), \tag{8}$$

$$\boldsymbol{Q}_{\boldsymbol{A}_\ell(\boldsymbol{\theta})} \triangleq \boldsymbol{I}_{M_\ell} - \boldsymbol{P}_{\boldsymbol{A}_\ell(\boldsymbol{\theta})}. \tag{9}$$

Note that $\boldsymbol{P}_{\boldsymbol{A}_\ell(\boldsymbol{\theta})}$ is the orthogonal projection onto the column space of $\boldsymbol{A}_\ell(\boldsymbol{\theta})$, and $\boldsymbol{Q}_{\boldsymbol{A}_\ell(\boldsymbol{\theta})}$ is its orthogonal complement. Equivalently, $\boldsymbol{Q}_{\boldsymbol{A}_\ell(\boldsymbol{\theta})}$ is the orthogonal projection onto the null space of $\boldsymbol{A}_\ell^H(\boldsymbol{\theta})$. The AOA estimation is obtained by minimizing (7) with respect to $\boldsymbol{\theta}$. Due to the non-convexity of (7) finding its global minimum requires a Q-dimensional search over all the hypotheses of $\boldsymbol{\theta}$.

As shown in the experiments section, this non-coherent processing suffers from insufficient results due to the fact that it does not exploit the phase information in the M sensors jointly. In other words, it does not utilize the full array aperture for estimating the AOAs.

IV. The Proposed Method

The proposed method is based on (exact or approximate) maximum likelihood estimation (MLE) of $\boldsymbol{\theta}$. This "joint" MLE approach, in which the phase shifts have to be estimated as well, allows us to benefit from the entire aperture of the M-element array, contrary to non-coherent processing methods.

Following the assumptions in Section II, note that the observations vector $\boldsymbol{x}$, which is defined in (3), is a circular complex Gaussian with mean $\boldsymbol{\Lambda}(\{\phi_\ell\})\boldsymbol{A}(\boldsymbol{\theta})\boldsymbol{s}$ and variance $\sigma^2\boldsymbol{I}_M$ (recall that $\{\boldsymbol{\theta}, \{\phi_\ell\}, \boldsymbol{s}\}$ are deterministic unknowns). Therefore, the MLE of the AOAs can be obtained by minimizing the negative log-likelihood function, which under our assumptions coincides with the following least squares objective

$$f_{\text{ML}}(\boldsymbol{\theta}, \{\phi_\ell\}, \boldsymbol{s}) \triangleq \left\|\boldsymbol{x} - \boldsymbol{\Lambda}(\{\phi_\ell\})\boldsymbol{A}(\boldsymbol{\theta})\boldsymbol{s}\right\|^2, \tag{10}$$

To simplify the formulation, let us define the "phase-corrected" observation vector

$$\overline{\boldsymbol{x}}(\{\phi_\ell\}) \triangleq \begin{bmatrix} \mathrm{e}^{j\phi_1}\boldsymbol{x}_1 \\ \vdots \\ \mathrm{e}^{j\phi_L}\boldsymbol{x}_L \end{bmatrix}, \tag{11}$$

which allows us to write

$$f_{\text{ML}}(\boldsymbol{\theta}, \{\phi_\ell\}, \boldsymbol{s}) = \left\|\overline{\boldsymbol{x}}(\{\phi_\ell\}) - \boldsymbol{A}(\boldsymbol{\theta})\boldsymbol{s}\right\|^2. \tag{12}$$

The minimizer of (12) with respect to $\boldsymbol{s}$ is given by

$$\hat{\boldsymbol{s}} = \left(\boldsymbol{A}^H(\boldsymbol{\theta})\boldsymbol{A}(\boldsymbol{\theta})\right)^{-1}\boldsymbol{A}^H(\boldsymbol{\theta})\overline{\boldsymbol{x}}(\{\phi_\ell\}), \tag{13}$$

and plugging it into (12) yields

$$\begin{aligned} f_{\text{ML}}(\boldsymbol{\theta}, \{\phi_\ell\}) &= \left\|(\boldsymbol{I}_M - \boldsymbol{P}_{\boldsymbol{A}(\boldsymbol{\theta})})\overline{\boldsymbol{x}}(\{\phi_\ell\})\right\|^2 \\ &= \left\|\boldsymbol{Q}_{\boldsymbol{A}(\boldsymbol{\theta})}\begin{bmatrix} \mathrm{e}^{j\phi_1}\boldsymbol{x}_1 \\ \vdots \\ \mathrm{e}^{j\phi_L}\boldsymbol{x}_L \end{bmatrix}\right\|^2. \end{aligned} \tag{14}$$

where

$$\boldsymbol{P}_{\boldsymbol{A}(\boldsymbol{\theta})} \triangleq \boldsymbol{A}(\boldsymbol{\theta})\left(\boldsymbol{A}^H(\boldsymbol{\theta})\boldsymbol{A}(\boldsymbol{\theta})\right)^{-1}\boldsymbol{A}^H(\boldsymbol{\theta}), \tag{15}$$

$$\boldsymbol{Q}_{\boldsymbol{A}(\boldsymbol{\theta})} \triangleq \boldsymbol{I}_M - \boldsymbol{P}_{\boldsymbol{A}(\boldsymbol{\theta})}. \tag{16}$$

In general, minimizing (14) with respect to all the unknown parameters requires a $(Q+L)$-dimensional search over all the hypotheses of $\boldsymbol{\theta}, \{\phi_\ell\}$. Therefore, it may not be tractable.

In what follows, we present a novel method to obtain closed-form estimators for the unknown phase shifts $\{\phi_\ell\}$ as functions

of $\boldsymbol{\theta}$. Given these expressions, it remains to minimize the negative log-likelihood function in (14) only with respect to $\boldsymbol{\theta}$, i.e. we eliminate the need to search over the L unknown phase offsets. Since the proposed approach still requires a Q-dimensional search, we also show how to integrate the proposed phase estimation with an alternating minimization scheme, which requires only one-dimensional searches.

A. Sub-arrays phase offsets estimation

In the following, we derive closed-form estimators for the unknown sub-arrays' phase shifts, which are functions of $\boldsymbol{\theta}$. The main idea is that for $L=2$ there is an analytic solution (a function of $\boldsymbol{\theta}$) for the MLE of ϕ_2 (recall that without loss of generality $\phi_1=0$). And, we show that a similar estimation technique can be used even for $L>2$. Given the analytical expressions of the estimated phases, it remains to minimize the negative log-likelihood function in (14) only with respect to $\boldsymbol{\theta}$. The obtained minimizer is the exact MLE for the case of $L=2$ and approximate MLE for $L>2$.

A key ingredient of our strategy is partition of projection matrices into horizontal blocks. We start with partitioning $\boldsymbol{Q}_{\boldsymbol{A}(\boldsymbol{\theta})}$ into L blocks $\boldsymbol{Q}_{\boldsymbol{A}(\boldsymbol{\theta})}=\left[\boldsymbol{Q}^{(1)}_{\boldsymbol{A}(\boldsymbol{\theta})}\cdots\boldsymbol{Q}^{(L)}_{\boldsymbol{A}(\boldsymbol{\theta})}\right]$, such that the ℓth block is of size $M\times M_\ell$. Note that $\boldsymbol{Q}^{(\ell)}_{\boldsymbol{A}(\boldsymbol{\theta})}$ is very different than $\boldsymbol{Q}_{\boldsymbol{A}_\ell(\boldsymbol{\theta})}$ in (7) (e.g. even their size is different). Using this partition, (14) can be written as

$$\begin{aligned} f_{\text{ML}}(\boldsymbol{\theta},\{\phi_\ell\}) &= \left\|\sum_{\ell=1}^{L}\boldsymbol{Q}^{(\ell)}_{\boldsymbol{A}(\boldsymbol{\theta})}\boldsymbol{x}_\ell \mathrm{e}^{j\phi_\ell}\right\|^2 \\ &= \sum_{\ell=1}^{L}\left\|\boldsymbol{Q}^{(\ell)}_{\boldsymbol{A}(\boldsymbol{\theta})}\boldsymbol{x}_\ell\right\|^2 \\ &\quad + \sum_{(\ell_1,\ell_2)\in P_2^L} 2\text{Re}\left(\boldsymbol{x}_{\ell_1}^H\boldsymbol{Q}^{(\ell_1)H}_{\boldsymbol{A}(\boldsymbol{\theta})}\boldsymbol{Q}^{(\ell_2)}_{\boldsymbol{A}(\boldsymbol{\theta})}\boldsymbol{x}_{\ell_2}\mathrm{e}^{j(\phi_{\ell_2}-\phi_{\ell_1})}\right), \end{aligned} \tag{17}$$

where the sum in the second term goes over the $\binom{L}{2}$ pairs of sub-arrays (each pair appears only once). Formally, P_2^L is defined by $P_2^L \triangleq \{(\ell_1,\ell_2)\ :\ 1\le\ell_1<\ell_2\le L\}$.

Let us consider the case of two sub-arrays ($L=2$) with phase offsets ϕ_1 and ϕ_2, In this case, we have $\boldsymbol{A}(\boldsymbol{\theta})=\begin{bmatrix}\boldsymbol{A}_1(\boldsymbol{\theta})\\ \boldsymbol{A}_2(\boldsymbol{\theta})\end{bmatrix}$ and

$$\begin{aligned} f_{\text{ML}}(\boldsymbol{\theta},\{\phi_\ell\}) &= \left\|\boldsymbol{Q}^{(1)}_{\boldsymbol{A}(\boldsymbol{\theta})}\boldsymbol{x}_1\right\|^2 + \left\|\boldsymbol{Q}^{(2)}_{\boldsymbol{A}(\boldsymbol{\theta})}\boldsymbol{x}_2\right\|^2 \\ &\quad + 2\text{Re}\left(\boldsymbol{x}_1^H\boldsymbol{Q}^{(1)H}_{\boldsymbol{A}(\boldsymbol{\theta})}\boldsymbol{Q}^{(2)}_{\boldsymbol{A}(\boldsymbol{\theta})}\boldsymbol{x}_2\mathrm{e}^{j(\phi_2-\phi_1)}\right). \end{aligned} \tag{18}$$

Note that the minimum of (18) with respect to $\phi_2-\phi_1$ is obtained for

$$\widehat{\phi_2-\phi_1} = \pi - \angle\left(\boldsymbol{x}_1^H\boldsymbol{Q}^{(1)H}_{\boldsymbol{A}(\boldsymbol{\theta})}\boldsymbol{Q}^{(2)}_{\boldsymbol{A}(\boldsymbol{\theta})}\boldsymbol{x}_2\right), \tag{19}$$

where $\angle x$ stands for the phase of the complex number x. This obviously implies that the closed-form expression (19) is the MLE of $\phi_2-\phi_1$, or equivalently, $\hat{\phi}_1=0$ and $\hat{\phi}_2=\pi-\angle\left(\boldsymbol{x}_1^H\boldsymbol{Q}^{(1)H}_{\boldsymbol{A}(\boldsymbol{\theta})}\boldsymbol{Q}^{(2)}_{\boldsymbol{A}(\boldsymbol{\theta})}\boldsymbol{x}_2\right)$ are MLEs of ϕ_1 and ϕ_2, respectively. Now, it is left to minimize $f_{\text{ML}}(\boldsymbol{\theta},\{\hat{\phi}_\ell\})$ with respect to $\boldsymbol{\theta}$ in order to obtain the MLE of the AOAs.

We turn to extend this approach to $L>2$ sub-arrays. Given a $\boldsymbol{\theta}$ hypothesis and a pair of sub-arrays (ℓ_1,ℓ_2), we propose to estimate the phase shift between the ℓ_2th sub-array and the ℓ_1th sub-array similarly to the way it is estimated for $\ell_1=1$ and $\ell_2=2$ in the case of $L=2$ sub-arrays.

Formally, let us define $\boldsymbol{A}_{\ell_1,\ell_2}(\boldsymbol{\theta}) \triangleq \begin{bmatrix}\boldsymbol{A}_{\ell_1}(\boldsymbol{\theta})\\ \boldsymbol{A}_{\ell_2}(\boldsymbol{\theta})\end{bmatrix}$, and let $\boldsymbol{Q}_{\boldsymbol{A}_{\ell_1,\ell_2}(\boldsymbol{\theta})}$ be the orthogonal projection onto the null space of $\boldsymbol{A}^H_{\ell_1,\ell_2}(\boldsymbol{\theta})$, i.e.

$$\begin{aligned} &\boldsymbol{Q}_{\boldsymbol{A}_{\ell_1,\ell_2}(\boldsymbol{\theta})} \\ &\triangleq \boldsymbol{I}_{M_{\ell_1}+M_{\ell_2}} - \boldsymbol{A}_{\ell_1,\ell_2}(\boldsymbol{\theta})\left(\boldsymbol{A}^H_{\ell_1,\ell_2}(\boldsymbol{\theta})\boldsymbol{A}_{\ell_1,\ell_2}(\boldsymbol{\theta})\right)^{-1}\boldsymbol{A}^H_{\ell_1,\ell_2}(\boldsymbol{\theta}). \end{aligned} \tag{20}$$

We partition $\boldsymbol{Q}_{\boldsymbol{A}_{\ell_1,\ell_2}(\boldsymbol{\theta})}$ into two horizontal blocks according to the size of the ℓ_1th and ℓ_2th sub-arrays $\boldsymbol{Q}_{\boldsymbol{A}_{\ell_1,\ell_2}(\boldsymbol{\theta})}=\left[\boldsymbol{Q}^{(1)}_{\boldsymbol{A}_{\ell_1,\ell_2}(\boldsymbol{\theta})},\boldsymbol{Q}^{(2)}_{\boldsymbol{A}_{\ell_1,\ell_2}(\boldsymbol{\theta})}\right]$. Thus, given a $\boldsymbol{\theta}$ hypothesis, the estimator of the phase offset for the pair of sub-arrays (ℓ_1,ℓ_2) is given by

$$\widehat{\phi_{\ell_2}-\phi_{\ell_1}} = \pi - \angle\left(\boldsymbol{x}_{\ell_1}^H\boldsymbol{Q}^{(1)H}_{\boldsymbol{A}_{\ell_1,\ell_2}(\boldsymbol{\theta})}\boldsymbol{Q}^{(2)}_{\boldsymbol{A}_{\ell_1,\ell_2}(\boldsymbol{\theta})}\boldsymbol{x}_{\ell_2}\right). \tag{21}$$

However, it is important to note that for $L>2$ sub-arrays, substituting (21) in all the $\binom{L}{2}=\frac{L(L-1)}{2}$ phase offsets $\{\phi_{\ell_2}-\phi_{\ell_1}\}$ in (17) would be wrong. It ignores the dependency between the $\binom{L}{2}$ pairs, which is due to the fact that there are only $L-1$ unknown phases $\{\phi_\ell\}_{\ell=2}^L$ that need to be estimated (recall that without loss of generality $\phi_1=0$). This means that for $L>2$ we should estimate $\{\phi_\ell\}_{\ell=2}^L$ directly.

As stated above, our goal is to reduce the complexity of minimizing f_{ML} from a $(Q+L)$-dimensional search to a Q-dimensional search (and beyond). Therefore, obtaining $L-1$ *closed-form estimators* for $\{\phi_\ell\}_{\ell=2}^L$ (as functions of $\boldsymbol{\theta}$) is desirable. Note that expressions for the estimators $\{\hat{\phi}_\ell\}_{\ell=2}^L$ can be straightforwardly formulated from a set of $L-1$ (out of the $\binom{L}{2}$) pairs of estimators $\{\widehat{\phi_{\ell_2}-\phi_{\ell_1}}\}$. On the other hand, obtaining $\{\hat{\phi}_\ell\}_{\ell=2}^L$ from more than $L-1$ (out of the $\binom{L}{2}$) pairs of $\{\widehat{\phi_{\ell_2}-\phi_{\ell_1}}\}$ is more complicated due to 2π ambiguity issues. Therefore, we propose to use (21) to estimate only $L-1$ out of the $\binom{L}{2}$ pairs of offsets $\{\phi_{\ell_2}-\phi_{\ell_1}\}$.

It remains to choose $L-1$ out of $\binom{L}{2}$ possible pairs of sub-arrays to which (21) is applied. We propose to use the $L-1$ pairs of *adjacent sub-arrays* because among all choices they are less prone to errors due to AOA ambiguities caused by widely spaced antenna elements. In non-adjacent sub-arrays there can be antenna elements with large gaps that lead to ambiguity in the AOA estimation [2], [3], which in turn harms the phase offset estimation. It is important to mention that we have verified our choice of $L-1$ adjacent sub-arrays by testing several other options of $L-1$ pairs, and observing that, indeed, choosing adjacent sub-arrays yields the best AOAs estimation performance.

To conclude this sub-section, let us assume that the sub-array indices are arranged such that two neighbor sub-arrays have subsequent numbers, i.e. the ℓth sub-array is close to the $(\ell-1)$th sub-array. Given a $\boldsymbol{\theta}$ hypothesis, we propose

Algorithm 1: Phase estimation for a given $\boldsymbol{\theta}$ hypothesis

Input: $\{\boldsymbol{x}_\ell\}_{\ell=1}^L$, $\{\boldsymbol{A}_\ell(\boldsymbol{\theta})\}_{\ell=1}^L$, where sub-array indices are arranged such that two neighbor sub-arrays have subsequent numbers.

$\hat{\phi}_1 = 0$;

for ℓ *from 2 to* L **do**

$\boldsymbol{A} = \begin{bmatrix} \boldsymbol{A}_{\ell-1}(\boldsymbol{\theta}) \\ \boldsymbol{A}_\ell(\boldsymbol{\theta}) \end{bmatrix}$;

$\boldsymbol{Q} = \boldsymbol{I} - \boldsymbol{A}(\boldsymbol{A}^H\boldsymbol{A})^{-1}\boldsymbol{A}^H$;

$\boldsymbol{Q}^{(1)} = \boldsymbol{Q}(:, 1 : M_{\ell-1})$;

$\boldsymbol{Q}^{(2)} = \boldsymbol{Q}(:, M_{\ell-1}+1 : end)$;

$\hat{\phi}_\ell = \hat{\phi}_{\ell-1} + \pi - \angle\left(\boldsymbol{x}_{\ell-1}^H \boldsymbol{Q}^{(1)H}\boldsymbol{Q}^{(2)}\boldsymbol{x}_\ell\right)$;

end

Output: $\{\hat{\phi}_\ell\}_{\ell=1}^L$ estimates of $\{\phi_\ell\}_{\ell=1}^L$.

Algorithm 2: Joint (approximate) maximum likelihood estimation (joint-MLE)

Input: $\{\boldsymbol{x}_\ell\}_{\ell=1}^L$, $\{\boldsymbol{A}_\ell(\boldsymbol{\theta})\}_{\ell=1}^L$.

$f_{min} \leftarrow +\infty$;

for $\tilde{\theta}_1$ *in* $[0°, 180°]$ **do**

$\vdots$

for $\tilde{\theta}_Q$ *in* $[0°, 180°]$ **do**

$\tilde{\boldsymbol{\theta}} \leftarrow [\tilde{\theta}_1, \ldots, \tilde{\theta}_Q]$;

$\{\hat{\phi}_\ell\} \leftarrow$ Algorithm 1 for $\tilde{\boldsymbol{\theta}}$;

$f \leftarrow f_{\text{ML}}(\tilde{\boldsymbol{\theta}}, \{\hat{\phi}_\ell\})$ (using (14));

if $f < f_{min}$ **then**

$f_{min} \leftarrow f$;

$\hat{\boldsymbol{\theta}} \leftarrow \tilde{\boldsymbol{\theta}}$;

end

end

$\vdots$

end

Output: $\hat{\boldsymbol{\theta}}$ estimate of $\boldsymbol{\theta}$.

to estimate $\{\widehat{\phi_\ell - \phi_{\ell-1}}\}_{\ell=2}^L$ by applying (21) with $(\ell_1, \ell_2) = (\ell-1, \ell)$, and then obtain estimators for $\{\phi_\ell\}$ by

$$\begin{aligned} &\hat{\phi}_1 = 0, \\ &\hat{\phi}_\ell = \hat{\phi}_{\ell-1} + \widehat{\phi_\ell - \phi_{\ell-1}}, \quad \ell = 2, \ldots, L. \end{aligned} \tag{22}$$

Note that (22) coincides with the exact MLE for $L = 2$, but is only an approximate MLE for $L > 2$.

The proposed phase-estimation procedure is summarized in Algorithm 1 with simplified notations. We emphasize that the $L-1$ matrix inversions are performed on $Q \times Q$ matrices, where Q is usually small. This procedure reduces the variable dimension of $f_{\text{ML}}(\boldsymbol{\theta}, \{\phi_\ell\})$ from $Q+L$ to Q, because it allows to go only over hypotheses of $\boldsymbol{\theta}$: For each given hypothesis of $\boldsymbol{\theta}$, the values of $\{\hat{\phi}_\ell\}_{\ell=1}^L$ can be obtained using Algorithm 1 and be substituted together with $\boldsymbol{\theta}$ in (14) or (17), to get a score for the hypothesis. The proposed method for approximate MLE of $\boldsymbol{\theta}$ is summarized in Algorithm 2.

B. Reduced complexity AOA estimation

For a large number of sources, e.g. $Q > 2$, the multidimensional search, which is required in order to obtain the global minimum of the non-convex function $f_{\text{ML}}(\boldsymbol{\theta}, \{\hat{\phi}_\ell\})$ in (14) or (17) (or even $f_{\text{non-coh}}(\boldsymbol{\theta})$ in (7)) is not tractable. Essentially, this is also the case in the coherent scenario, where all the sub-arrays are coherent, and can be considered as a single coherent array.

Over the years several computationally efficient techniques have been proposed for minimizing the negative log-likelihood function in the coherent case, e.g. [16]–[18]. These methods are iterative, and require only one-dimensional searches. However, they tend to converge to local minima (which is inevitable due to the non-convex optimization problem).

A very popular algorithm for AOA estimation in the coherent case, which has been proposed in [16], is based on alternating minimization (AM). Let us denote the negative log-likelihood function in the coherent case (where $\{\phi_\ell\}_{\ell=1}^L$ are zero or known) by $f_{\text{ML}}(\boldsymbol{\theta}|\{\phi_\ell\})$. In each AM iteration $f_{\text{ML}}(\boldsymbol{\theta}|\{\phi_\ell\})$ is minimized with respect to a single θ_q while the other AOAs remain fixed. More formally, the estimator of the qth AOA in the kth iteration, denoted by $\theta_q^{(k)}$, is the minimizer of

$$f_{\text{ML}}([\theta_1^{(k)}, \ldots, \theta_{q-1}^{(k)}, \theta_q, \theta_{q+1}^{(k-1)}, \ldots, \theta_Q^{(k-1)}]|\{\phi_\ell\}) \tag{23}$$

with respect to θ_q. Note that $\theta_q^{(k)}$ can be obtained using a *one-dimensional* search. The AM iterations stop when a stopping criterion is met, e.g. when $\|\boldsymbol{\theta}^{(k)} - \boldsymbol{\theta}^{(k-1)}\|$ is smaller than a predefined threshold. The initialization, which strongly affects the performance, is also done by estimating the AOAs one by one. In more details, $\theta_1^{(0)}$ is obtained by simple beamforming, and for $q > 1$, $\theta_q^{(0)}$ is a minimizer of $f_{\text{ML}}(\boldsymbol{\theta}|\{\phi_\ell\})$ under the assumption of q sources, and with the previous $\{\theta_{q'}\}_{q'=1}^{q-1}$ fixed to $\{\theta_{q'}^{(0)}\}_{q'=1}^{q-1}$ (see [16] for more details).

Considering the case of non-coherent sub-arrays that we discuss in this paper, a similar AM scheme (including the initialization) can be applied on the non-coherent cost function $f_{\text{non-coh}}(\boldsymbol{\theta})$ given in (7). Regarding our joint-minimization of $f_{\text{ML}}(\boldsymbol{\theta}, \{\phi_\ell\})$, it is still possible to construct an AM optimization that requires only one-dimensional searches. Basically, the main difference from the other implementations is that for every hypothesis $[\theta_1^{(k)}, \ldots, \theta_{q-1}^{(k)}, \theta_q, \theta_{q+1}^{(k-1)}, \ldots, \theta_Q^{(k-1)}]$ in the one-dimensional search, it is required to obtain the estimated phases $\{\hat{\phi}_\ell\}$ via Algorithm 1. However, note that Algorithm 1 requires all the AOAs (to formulate $\{\boldsymbol{A}_\ell(\boldsymbol{\theta})\}_{\ell=1}^L$). Therefore, it cannot be used in the initialization strategy that is mentioned above. To resolve this issue, we suggest to initialize the AM of $f_{\text{ML}}(\boldsymbol{\theta}, \{\phi_\ell\})$ with the AOAs which are obtained by the AM of $f_{\text{non-coh}}(\boldsymbol{\theta})$. The proposed joint alternating-minimization scheme is summarized in Algorithm 3.

C. Complexity analysis

In the following, we present complexity analysis for Algorithm 2 and Algorithm 3, compare them to each other, and show that each of them has similar complexity as its counterpart in the coherent case, where the phase offsets of the

Algorithm 3: Joint alternating-minimization (joint-AM)

Input: $\{\boldsymbol{x}_\ell\}_{\ell=1}^L$, $\{\boldsymbol{A}_\ell(\boldsymbol{\theta})\}_{\ell=1}^L$, $\boldsymbol{\theta}^{(0)}$.
while *stopping criterion not met* **do**
  $k = k+1$;
  for q *from 1 to* Q **do**
    $f_{min} \leftarrow +\infty$;
    for $\tilde{\theta}_q$ *in* $[0^\circ, 180^\circ]$ **do**
      $\tilde{\boldsymbol{\theta}} \leftarrow [\theta_1^{(k)}, \ldots, \theta_{q-1}^{(k)}, \tilde{\theta}_q, \theta_{q+1}^{(k-1)}, \ldots, \theta_Q^{(k-1)}]$;
      $\{\hat{\phi}_\ell\} \leftarrow$ Algorithm 1 for $\tilde{\boldsymbol{\theta}}$;
      $f \leftarrow f_{\text{ML}}(\tilde{\boldsymbol{\theta}}, \{\hat{\phi}_\ell\})$ (using (14));
      if $f < f_{min}$ **then**
        $f_{min} \leftarrow f$;
        $\theta_q^{(k)} \leftarrow \tilde{\theta}_q$;
      end
    end
  end
end
Output: $\boldsymbol{\theta}^{(k)}$ estimate of $\boldsymbol{\theta}$.

sub-arrays are zero or known (and thus applying Algorithm 1 is spared). Before we continue, we remind the reader that M is the number of entire array elements, L is the number of sub-arrays, and Q is the number of sources. For simplification we assume that all the sub-arrays have the same number of elements, i.e. $\frac{M}{L}$ elements. We also denote by N_θ the number of θ hypotheses (the length of the one-dimensional AOA search grid), and by K the number of iterations used in Algorithm 3.

We start with calculating the complexity of (14), which is used by both Algorithm 2 and Algorithm 3. First, we compute the complexity of the $M \times M$ matrix $\boldsymbol{Q}_{\boldsymbol{A}(\boldsymbol{\theta})}$, which is defined by $\boldsymbol{Q}_{\boldsymbol{A}} \triangleq \boldsymbol{I}_M - \boldsymbol{A}(\boldsymbol{A}^H\boldsymbol{A})^{-1}\boldsymbol{A}^H$, where $\boldsymbol{A}$ is an $M \times Q$ matrix and for brevity we omitted the dependencies on $\boldsymbol{\theta}$. The complexity of $\boldsymbol{A}^H\boldsymbol{A}$ is $\mathcal{O}(MQ^2)$, and the complexity of $(\boldsymbol{A}^H\boldsymbol{A})^{-1}$ is $\mathcal{O}(MQ^2) + \mathcal{O}(Q^3)$. It follows that the complexity of $\boldsymbol{A}(\boldsymbol{A}^H\boldsymbol{A})^{-1}\boldsymbol{A}^H$ is $3\mathcal{O}(MQ^2) + \mathcal{O}(Q^3)$, and thus the complexity of $\boldsymbol{Q}_{\boldsymbol{A}}$ is $\mathcal{O}(M) + 3\mathcal{O}(MQ^2) + \mathcal{O}(Q^3) = \mathcal{O}(MQ^2)$, where we used $M > Q$. Computing (14) includes also matrix-vector multiplication, where the matrix is of size $M \times M$ and the vector is of size $M \times 1$. The complexity of this step is $\mathcal{O}(M^2)$. To conclude, the overall complexity of (14) is $\mathcal{O}(MQ^2) + \mathcal{O}(M^2)$.

Let us now calculate the complexity of Algorithm 1, which is also used by both Algorithm 2 and Algorithm 3. In this algorithm there are $L-1$ iterations. Each iteration includes computation of an $2\frac{M}{L} \times 2\frac{M}{L}$ matrix $\boldsymbol{Q}_{\boldsymbol{A}_{\ell-1,\ell}(\boldsymbol{\theta})}$ (cf. (20)). Similar to the complexity calculation of $\boldsymbol{Q}_{\boldsymbol{A}}$ above, it can be shown that the complexity of computing this matrix is $\mathcal{O}(2\frac{M}{L}Q^2)$. Each iteration of Algorithm 1 includes also computation of (21), which is dominated by two matrix-vector multiplications, where the matrix is of size $2\frac{M}{L} \times \frac{M}{L}$ and the vector is of size $\frac{M}{L} \times 1$. The complexity of this step is $2\mathcal{O}(2\frac{M^2}{L^2})$. To conclude, the overall complexity of Algorithm 1 is $\mathcal{O}((L-1)(\frac{M}{L}Q^2)) + \mathcal{O}((L-1)\frac{M^2}{L^2})) = \mathcal{O}(MQ^2) + \mathcal{O}(\frac{M^2}{L})$.

We turn to compute the complexity of Algorithm 2. This algorithm goes over N_θ^Q hypotheses of $\boldsymbol{\theta}$, and for each hypothesis it computes both Algorithm 1 and (14). Therefore, the total complexity of Algorithm 2 is $\mathcal{O}(N_\theta^Q(MQ^2 + \frac{M^2}{L} + MQ^2 + M^2)) = \mathcal{O}(N_\theta^Q(MQ^2 + M^2))$.

Similarly, we compute the complexity of Algorithm 3. In this algorithm there are K outer iterations and Q inner iterations. Each of them goes over N_θ hypotheses of θ (one-dimensional search) and for each hypothesis computes both Algorithm 1 and (14). Therefore, the total complexity of Algorithm 3 is $\mathcal{O}(KQN_\theta(MQ^2 + \frac{M^2}{L} + MQ^2 + M^2)) = \mathcal{O}(KQN_\theta(MQ^2 + M^2))$.

The results above show that Algorithm 3 has lower complexity than Algorithm 2 by a factor of $\frac{N_\theta^{Q-1}}{KQ}$. Since typically $N_\theta \gg KQ$, it follows that Algorithm 3 is more suitable to handle the cases with multiple sources ($Q > 1$). Furthermore, the results show that both Algorithm 2 and Algorithm 3 have similar complexity as their counterparts in the coherent case, when the entire array is coherent (and thus Algorithm 1 is not applied). This follows from the fact that the complexity of Algorithm 1 is not larger than the complexity of computing the cost function in (14).

D. Extension to multiple snapshots

In this sub-section we briefly discuss how to extend our approach to multiple snapshots (temporal observations). First, recall that the paper considers sub-arrays that are non-coherent due to using unsynchronized local oscillators (which reduces the system complexity). In this case, in each received snapshot there is a different phase offset between the sub-arrays, i.e. while the AOAs are the same over the different snapshots, the phase offsets are changing. Therefore, methods like [4]–[6], [8], [11], [12] that assume multiple snapshots that share the same nuisance sub-arrays phase offsets parameters cannot be used.

Under these conditions, extending the proposed approach to multiple snapshots is straightforward. In this case, for N snapshots there are N received signals per each sub-array, $\{\boldsymbol{x}_\ell(n)\}_{\ell=1,n=1}^{L,N}$, N phase offsets for each sub-array, $\{\phi_\ell(n)\}_{\ell=1,n=1}^{L,N}$, and N different source signals, $\{\boldsymbol{s}(n)\}_{n=1}^N$, where ℓ is the sub-array index and n is the snapshot index. It follows that the negative log-likelihood function is given by

$$f_{\text{ML}}(\boldsymbol{\theta}, \{\phi_\ell(n)\}, \{\boldsymbol{s}(n)\}) = \sum_{n=1}^{N} \left\| \begin{bmatrix} \mathrm{e}^{j\phi_1(n)}\boldsymbol{x}_1(n) \\ \vdots \\ \mathrm{e}^{j\phi_L(n)}\boldsymbol{x}_L(n) \end{bmatrix} - \boldsymbol{A}(\boldsymbol{\theta})\boldsymbol{s}(n) \right\|^2 . \quad (24)$$

Note that this formula subsumes the one in (12), which is given for a single snapshot $N = 1$. After substituting the estimations of the different signals, $\{\boldsymbol{s}(n)\}_{n=1}^N$, in it, we get

$$f_{\text{ML}}(\boldsymbol{\theta}, \{\phi_\ell(n)\}) = \sum_{n=1}^{N} \left\| \boldsymbol{Q}_{\boldsymbol{A}(\boldsymbol{\theta})} \begin{bmatrix} \mathrm{e}^{j\phi_1(n)}\boldsymbol{x}_1(n) \\ \vdots \\ \mathrm{e}^{j\phi_L(n)}\boldsymbol{x}_L(n) \end{bmatrix} \right\|^2 . \quad (25)$$

Note that this formula is the same as (14), except a sum over the snapshots and that the number of unknown phase offsets is increased from L to LN. For a given $\boldsymbol{\theta}$ hypothesis, Algorithm 1 can be used, without any change, to estimate the unknown phase offsets in each snapshot.

Extending our approach to another slightly different multiple snapshots problem, where the phase offsets are the same for different snapshots, is also straightforward. Note, though, that this problem is related to a partial calibration issue rather than to non-coherency of sub-arrays. In this case $\phi_\ell(n) = \phi_\ell$ for all n. Therefore, after substituting the estimations of $\{\boldsymbol{s}(n)\}_{n=1}^N$ in the negative log-likelihood function, we get

$$f_{\text{ML}}(\boldsymbol{\theta}, \{\phi_\ell\}) = \sum_{n=1}^{N} \left\| \boldsymbol{Q}_{\boldsymbol{A}(\boldsymbol{\theta})} \begin{bmatrix} \mathrm{e}^{j\phi_1}\boldsymbol{x}_1(n) \\ \vdots \\ \mathrm{e}^{j\phi_L}\boldsymbol{x}_L(n) \end{bmatrix} \right\|^2. \tag{26}$$

It then follows that, for a given $\boldsymbol{\theta}$ hypothesis, Algorithm 1 can still be used to estimate the unknown phases with merely changing $\angle\left(\boldsymbol{x}_{\ell-1}^H \boldsymbol{Q}^{(1)H}\boldsymbol{Q}^{(2)}\boldsymbol{x}_\ell\right)$ to $\angle\left(\frac{1}{N}\sum\limits_{n=1}^{N} \boldsymbol{x}_{\ell-1}^H(n)\boldsymbol{Q}^{(1)H}\boldsymbol{Q}^{(2)}\boldsymbol{x}_\ell(n)\right)$.

V. Cramér-Rao Lower Bound

In this section we derive the Cramér-Rao Lower Bound (CRLB) for AOA estimation with non-coherent sub-arrays. Note that CRLBs for systems of non-coherent sub-arrays have been derived in [9], [10]. But the CRLB that is derived in this section is different from these previous results. The main difference is due to the fact that in [9], [10] the CRLB is derived for a model in which each sub-array observes a different signal of the sources $\boldsymbol{s}_\ell$ (cf. (4)). In contrast, in our problem formulation all the sub-arrays receive the same $\boldsymbol{s}$ (cf. (1)), however, each sub-array has an unknown phase offset due to the non-coherency between the sub-arrays (as a result of having different local oscillators).

We start with establishing a general result for any number of sources. Then, we simplify the result for a single source ($Q = 1$), and compare it with the CRLB for the case where each sub-array observes a different signal of the sources, which essentially lower bounds the performance of methods that use non-coherent processing, such as minimizing (7).

A. Cramér-Rao Lower Bound for $Q \geq 1$

Let us define the unknown parameter vector

$$\boldsymbol{\psi} \triangleq \left[\bar{\boldsymbol{s}}^T, \boldsymbol{\phi}^T, \boldsymbol{\theta}^T\right]^T, \tag{27}$$

where

$$\begin{aligned} \boldsymbol{\phi} &\triangleq [\phi_1, \ldots, \phi_L]^T, \\ \bar{\boldsymbol{s}} &\triangleq \left[\mathrm{Re}\{\boldsymbol{s}\}^T, \mathrm{Im}\{\boldsymbol{s}\}^T\right]^T. \end{aligned} \tag{28}$$

The observations vector $\boldsymbol{x}_\ell$ in (1) is a proper complex Gaussian, random vector, whose mean and covariance are given by

$$\begin{aligned} \boldsymbol{m}_\ell &\triangleq \mathrm{e}^{-j\phi_\ell}\boldsymbol{A}_\ell(\boldsymbol{\theta})\boldsymbol{s}, \\ \boldsymbol{\Sigma}_\ell &\triangleq \sigma^2 \boldsymbol{I}_{M_\ell}, \end{aligned} \tag{29}$$

where σ^2 is the known noise variance. The $M \times 1$ vector that is obtained by concatenating $\{\boldsymbol{x}_\ell\}_{\ell=1}^L$ also has proper complex Gaussian distribution with mean and covariance

$$\begin{aligned} \boldsymbol{m}(\psi) &\triangleq \begin{bmatrix} \boldsymbol{m}_1 \\ \vdots \\ \boldsymbol{m}_L \end{bmatrix} = \check{\boldsymbol{A}}(\boldsymbol{\theta}, \boldsymbol{\phi})\boldsymbol{s}, \\ \boldsymbol{\Sigma} &\triangleq \sigma^2 \boldsymbol{I}_M, \end{aligned} \tag{30}$$

where $\check{\boldsymbol{A}}(\boldsymbol{\theta}, \boldsymbol{\phi}) \triangleq \begin{bmatrix} \mathrm{e}^{-j\phi_1}\boldsymbol{A}_1(\boldsymbol{\theta}) \\ \vdots \\ \mathrm{e}^{-j\phi_L}\boldsymbol{A}_L(\boldsymbol{\theta}) \end{bmatrix}$.

The Fisher information matrix (FIM) of the parameter vector $\boldsymbol{\psi}$ can be partitioned into blocks associated with the different combinations of the components in (27)

$$\boldsymbol{J}(\boldsymbol{\psi}) \triangleq \begin{bmatrix} \boldsymbol{J}_{\bar{\boldsymbol{s}}\bar{\boldsymbol{s}}} & \boldsymbol{J}_{\bar{\boldsymbol{s}}\boldsymbol{\phi}} & \boldsymbol{J}_{\bar{\boldsymbol{s}}\boldsymbol{\theta}} \\ \boldsymbol{J}_{\bar{\boldsymbol{s}}\boldsymbol{\phi}}^T & \boldsymbol{J}_{\boldsymbol{\phi}\boldsymbol{\phi}} & \boldsymbol{J}_{\boldsymbol{\phi}\boldsymbol{\theta}} \\ \boldsymbol{J}_{\bar{\boldsymbol{s}}\boldsymbol{\theta}}^T & \boldsymbol{J}_{\boldsymbol{\phi}\boldsymbol{\theta}}^T & \boldsymbol{J}_{\boldsymbol{\theta}\boldsymbol{\theta}} \end{bmatrix}. \tag{31}$$

The entries of the FIM for proper complex Gaussian observations, which depend on the parameters only through their expectation, are given by [2]

$$[\boldsymbol{J}(\boldsymbol{\psi})]_{n,m} = 2\,\mathrm{Re}\left\{\left(\frac{\partial \boldsymbol{m}(\boldsymbol{\psi})}{\partial \psi_n}\right)^H \boldsymbol{\Sigma}^{-1}\left(\frac{\partial \boldsymbol{m}(\boldsymbol{\psi})}{\partial \psi_m}\right)\right\}. \tag{32}$$

Let us calculate the necessary derivatives

$$\begin{aligned} \frac{\partial \boldsymbol{m}(\boldsymbol{\psi})}{\partial \bar{\boldsymbol{s}}} &= \left[\check{\boldsymbol{A}}(\boldsymbol{\theta}, \boldsymbol{\phi}) \quad j\check{\boldsymbol{A}}(\boldsymbol{\theta}, \boldsymbol{\phi})\right], \\ \frac{\partial \boldsymbol{m}(\boldsymbol{\psi})}{\partial \boldsymbol{\theta}} &= \check{\boldsymbol{A}}_{\boldsymbol{\theta}}(\boldsymbol{\theta}, \boldsymbol{\phi})\boldsymbol{S}, \\ \frac{\partial \boldsymbol{m}(\boldsymbol{\psi})}{\partial \boldsymbol{\phi}} &= \begin{bmatrix} -j\mathrm{e}^{-j\phi_1}\boldsymbol{A}_1(\boldsymbol{\theta})\boldsymbol{s} & & \boldsymbol{0} \\ & \ddots & \\ \boldsymbol{0} & & -j\mathrm{e}^{-j\phi_L}\boldsymbol{A}_L(\boldsymbol{\theta})\boldsymbol{s} \end{bmatrix}, \end{aligned} \tag{33}$$

where $\boldsymbol{S} \triangleq \mathrm{diag}\{s_1, \ldots, s_Q\}$ is a diagonal matrix with $\{s_1, \ldots, s_Q\}$ on the main diagonal, and $\left[\check{\boldsymbol{A}}_{\boldsymbol{\theta}}(\boldsymbol{\theta}, \boldsymbol{\phi})\right]_{m,q} \triangleq \frac{\partial\left[\check{\boldsymbol{A}}(\boldsymbol{\theta}, \boldsymbol{\phi})\right]_{m,q}}{\partial \theta_q}$. Note that $\check{\boldsymbol{A}}_{\boldsymbol{\theta}}(\boldsymbol{\theta}, \boldsymbol{\phi}) = \begin{bmatrix} \mathrm{e}^{-j\phi_1}\dot{\boldsymbol{A}}_1(\boldsymbol{\theta}) \\ \vdots \\ \mathrm{e}^{-j\phi_L}\dot{\boldsymbol{A}}_L(\boldsymbol{\theta}) \end{bmatrix}$, where

$\left[\dot{\boldsymbol{A}}_\ell(\boldsymbol{\theta})\right]_{m,q} = \frac{d[\boldsymbol{A}_\ell(\boldsymbol{\theta})]_{m,q}}{d\theta_q}$. Therefore, we have

$$
\begin{aligned}
\boldsymbol{J}_{\boldsymbol{\theta\theta}} &= \frac{2}{\sigma^2}\operatorname{Re}\left\{\boldsymbol{S}^H \check{\boldsymbol{A}}_{\boldsymbol{\theta}}^H \check{\boldsymbol{A}}_{\boldsymbol{\theta}} \boldsymbol{S}\right\} \\
&= \frac{2}{\sigma^2}\operatorname{Re}\left\{\boldsymbol{S}^H \left(\sum_{\ell=1}^{L} \dot{\boldsymbol{A}}_\ell^H \dot{\boldsymbol{A}}_\ell\right) \boldsymbol{S}\right\} \\
\boldsymbol{J}_{\boldsymbol{\phi\phi}} &= \frac{2}{\sigma^2}\operatorname{Re}\left\{\operatorname{diag}\{\boldsymbol{s}^H \boldsymbol{A}_1^H \boldsymbol{A}_1 \boldsymbol{s}, \ldots, \boldsymbol{s}^H \boldsymbol{A}_L^H \boldsymbol{A}_L \boldsymbol{s}\}\right\} \\
\boldsymbol{J}_{\bar{\boldsymbol{s}}\bar{\boldsymbol{s}}} &= \frac{2}{\sigma^2}\begin{bmatrix} \operatorname{Re}\left\{\check{\boldsymbol{A}}^H \check{\boldsymbol{A}}\right\} & -\operatorname{Im}\left\{\check{\boldsymbol{A}}^H \check{\boldsymbol{A}}\right\} \\ \operatorname{Im}\left\{\check{\boldsymbol{A}}^H \check{\boldsymbol{A}}\right\} & \operatorname{Re}\left\{\check{\boldsymbol{A}}^H \check{\boldsymbol{A}}\right\} \end{bmatrix}, \\
&= \frac{2}{\sigma^2}\begin{bmatrix} \operatorname{Re}\left\{\sum_{\ell=1}^{L} \boldsymbol{A}_\ell^H \boldsymbol{A}_\ell\right\} & -\operatorname{Im}\left\{\sum_{\ell=1}^{L} \boldsymbol{A}_\ell^H \boldsymbol{A}_\ell\right\} \\ \operatorname{Im}\left\{\sum_{\ell=1}^{L} \boldsymbol{A}_\ell^H \boldsymbol{A}_\ell\right\} & \operatorname{Re}\left\{\sum_{\ell=1}^{L} \boldsymbol{A}_\ell^H \boldsymbol{A}_\ell\right\} \end{bmatrix}, \\
\boldsymbol{J}_{\bar{\boldsymbol{s}}\boldsymbol{\phi}} &= \frac{2}{\sigma^2}\begin{bmatrix} \operatorname{Im}\left\{\begin{bmatrix} \boldsymbol{A}_1^H \boldsymbol{A}_1 \boldsymbol{s} & \ldots & \boldsymbol{A}_L^H \boldsymbol{A}_L \boldsymbol{s} \end{bmatrix}\right\} \\ -\operatorname{Re}\left\{\begin{bmatrix} \boldsymbol{A}_1^H \boldsymbol{A}_1 \boldsymbol{s} & \ldots & \boldsymbol{A}_L^H \boldsymbol{A}_L \boldsymbol{s} \end{bmatrix}\right\} \end{bmatrix}, \\
\boldsymbol{J}_{\bar{\boldsymbol{s}}\boldsymbol{\theta}} &= \frac{2}{\sigma^2}\begin{bmatrix} \operatorname{Re}\left\{\check{\boldsymbol{A}}^H \check{\boldsymbol{A}}_{\boldsymbol{\theta}} \boldsymbol{S}\right\} \\ \operatorname{Im}\left\{\check{\boldsymbol{A}}^H \check{\boldsymbol{A}}_{\boldsymbol{\theta}} \boldsymbol{S}\right\} \end{bmatrix}, \\
&= \frac{2}{\sigma^2}\begin{bmatrix} \operatorname{Re}\left\{\sum_{\ell=1}^{L} \boldsymbol{A}_\ell^H \dot{\boldsymbol{A}}_\ell \boldsymbol{S}\right\} \\ \operatorname{Im}\left\{\sum_{\ell=1}^{L} \boldsymbol{A}_\ell^H \dot{\boldsymbol{A}}_\ell \boldsymbol{S}\right\} \end{bmatrix}, \\
\boldsymbol{J}_{\boldsymbol{\phi\theta}} &= -\frac{2}{\sigma^2}\operatorname{Im}\left\{\begin{bmatrix} \boldsymbol{s}^H \boldsymbol{A}_1^H \dot{\boldsymbol{A}}_1 \boldsymbol{S} \\ \vdots \\ \boldsymbol{s}^H \boldsymbol{A}_L^H \dot{\boldsymbol{A}}_L \boldsymbol{S} \end{bmatrix}\right\}.
\end{aligned} \tag{34}
$$

Since any unknown phase can be included in the unknown vector $\boldsymbol{s}$, the row and column associated with one phase, e.g. ϕ_1, need to be removed from $\boldsymbol{J}(\boldsymbol{\psi})$ in order to make it invertible. The CRLB for the estimation of $\boldsymbol{\theta}$ is obtained by the $Q \times Q$ lower right block of $\boldsymbol{J}^{-1}(\boldsymbol{\psi})$. Formally, using block-wise inversion on (31), the lower right block of $\boldsymbol{J}^{-1}(\boldsymbol{\psi})$ is given by

$$
CRLB(\boldsymbol{\theta}) = \left(\boldsymbol{J}_{\boldsymbol{\theta\theta}} - \begin{bmatrix} \boldsymbol{J}_{\bar{\boldsymbol{s}}\boldsymbol{\theta}}^T \boldsymbol{J}_{\boldsymbol{\phi\theta}}^T \end{bmatrix} \begin{bmatrix} \boldsymbol{X}_{11} & \boldsymbol{X}_{12} \\ \boldsymbol{X}_{21} & \boldsymbol{X}_{22} \end{bmatrix} \begin{bmatrix} \boldsymbol{J}_{\bar{\boldsymbol{s}}\boldsymbol{\theta}} \\ \boldsymbol{J}_{\boldsymbol{\phi\theta}} \end{bmatrix}\right)^{-1}, \tag{35}
$$

where we use the notation

$$
\begin{bmatrix} \boldsymbol{J}_{\bar{\boldsymbol{s}}\bar{\boldsymbol{s}}} & \boldsymbol{J}_{\bar{\boldsymbol{s}}\boldsymbol{\phi}} \\ \boldsymbol{J}_{\bar{\boldsymbol{s}}\boldsymbol{\phi}}^T & \boldsymbol{J}_{\boldsymbol{\phi\phi}} \end{bmatrix}^{-1} \triangleq \begin{bmatrix} \boldsymbol{X}_{11} & \boldsymbol{X}_{12} \\ \boldsymbol{X}_{21} & \boldsymbol{X}_{22} \end{bmatrix}. \tag{36}
$$

B. The special case of $Q = 1$ (single source)

Let us consider the case of a single source, i.e. $Q = 1$ and s is a scalar. In the following, we derive a simplified expression for the CRLB for non-coherent sub-arrays in this case. We start by simplifying the terms in (34), and removing the row and column associated with ϕ_1 (as explained in the end of Section V-A). We have

$$
\begin{aligned}
\boldsymbol{J}_{\theta\theta} &= \frac{2}{\sigma^2}|s|^2 \sum_{\ell=1}^{L} \|\dot{\mathbf{a}}_\ell\|^2 \\
\boldsymbol{J}_{\boldsymbol{\phi\phi}} &= \frac{2}{\sigma^2}|s|^2 \operatorname{diag}\{\|\mathbf{a}_2\|^2, \ldots, \|\mathbf{a}_L\|^2\} \\
\boldsymbol{J}_{\bar{\boldsymbol{s}}\bar{\boldsymbol{s}}} &= \frac{2}{\sigma^2}\sum_{\ell=1}^{L} \|\mathbf{a}_\ell\|^2 \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \\
\boldsymbol{J}_{\bar{\boldsymbol{s}}\boldsymbol{\phi}} &= \frac{2}{\sigma^2}\begin{bmatrix} \operatorname{Im}\{s\}\begin{bmatrix} \|\mathbf{a}_2\|^2 & \ldots & \|\mathbf{a}_L\|^2 \end{bmatrix} \\ -\operatorname{Re}\{s\}\begin{bmatrix} \|\mathbf{a}_2\|^2 & \ldots & \|\mathbf{a}_L\|^2 \end{bmatrix} \end{bmatrix}, \\
\boldsymbol{J}_{\bar{\boldsymbol{s}}\boldsymbol{\theta}} &= \frac{2}{\sigma^2}\begin{bmatrix} \operatorname{Re}\left\{s\sum_{\ell=1}^{L} \mathbf{a}_\ell^H \dot{\mathbf{a}}_\ell\right\} \\ \operatorname{Im}\left\{s\sum_{\ell=1}^{L} \mathbf{a}_\ell^H \dot{\mathbf{a}}_\ell\right\} \end{bmatrix}, \\
\boldsymbol{J}_{\boldsymbol{\phi\theta}} &= -\frac{2}{\sigma^2}|s|^2 \operatorname{Im}\left\{\begin{bmatrix} \mathbf{a}_2^H \dot{\mathbf{a}}_2 \\ \vdots \\ \mathbf{a}_L^H \dot{\mathbf{a}}_L \end{bmatrix}\right\}.
\end{aligned} \tag{37}
$$

A common mild assumption that simplifies the analysis of CRLB for a single source is that the array response $\mathbf{a}(\theta)$ is normalized by a complex scalar (which may be a function of θ), such that $\mathbf{a}^H(\theta)\dot{\mathbf{a}}(\theta) = 0$ [2]. For widely-used uniform linear arrays (ULAs) of omnidirectional sensors, this merely implies that the reference point (in which the phase is zero) should be set in the middle of the ULA.

The assumption $\mathbf{a}^H(\theta)\dot{\mathbf{a}}(\theta) = 0$ is translated in our case to $\sum_{\ell=1}^{L} \mathbf{a}_\ell^H(\theta)\dot{\mathbf{a}}_\ell(\theta) = 0$, which implies $\boldsymbol{J}_{\bar{\boldsymbol{s}}\theta} = \mathbf{0}$. Therefore, (35) turns into

$$
CRLB^{-1}(\theta) = \boldsymbol{J}_{\theta\theta} - \boldsymbol{J}_{\boldsymbol{\phi}\theta}^T \boldsymbol{X}_{22} \boldsymbol{J}_{\boldsymbol{\phi}\theta}. \tag{38}
$$

Using block-wise inversion on (36), we get

$$
\boldsymbol{X}_{22} = \left(\boldsymbol{J}_{\boldsymbol{\phi\phi}} - \boldsymbol{J}_{\bar{\boldsymbol{s}}\boldsymbol{\phi}}^T \boldsymbol{J}_{\bar{\boldsymbol{s}}\bar{\boldsymbol{s}}}^{-1} \boldsymbol{J}_{\bar{\boldsymbol{s}}\boldsymbol{\phi}}\right)^{-1}. \tag{39}
$$

Now, note that

$$
\begin{aligned}
&\boldsymbol{J}_{\boldsymbol{\phi\phi}} - \boldsymbol{J}_{\bar{\boldsymbol{s}}\boldsymbol{\phi}}^T \boldsymbol{J}_{\bar{\boldsymbol{s}}\bar{\boldsymbol{s}}}^{-1} \boldsymbol{J}_{\bar{\boldsymbol{s}}\boldsymbol{\phi}} \\
&= \frac{2|s|^2}{\sigma^2}\begin{bmatrix} \|\mathbf{a}_2\|^2 & & 0 \\ & \ddots & \\ 0 & & \|\mathbf{a}_L\|^2 \end{bmatrix} \\
&\quad - \frac{2|s|^2}{\sigma^2 \sum_{\ell=1}^{L} \|\mathbf{a}_\ell\|^2} \underbrace{\begin{bmatrix} \|\mathbf{a}_2\|^2\|\mathbf{a}_2\|^2 & \ldots & \|\mathbf{a}_2\|^2\|\mathbf{a}_L\|^2 \\ \vdots & \ddots & \vdots \\ \|\mathbf{a}_L\|^2\|\mathbf{a}_2\|^2 & \ldots & \|\mathbf{a}_L\|^2\|\mathbf{a}_L\|^2 \end{bmatrix}}_{=\boldsymbol{u}\boldsymbol{u}^T},
\end{aligned} \tag{40}
$$

where $\boldsymbol{u}^T \triangleq [\|\mathbf{a}_2\|^2, \ldots, \|\mathbf{a}_L\|^2]$. Using Sherman-Morrison

formula we have

$$\boldsymbol{X}_{22} = \frac{\sigma^2}{2|s|^2}\left[\begin{array}{cc} \begin{bmatrix} \|\mathbf{a}_2\|^{-2} & & 0 \\ & \ddots & \\ 0 & & \|\mathbf{a}_L\|^{-2} \end{bmatrix} + \frac{\frac{1}{\sum_{\ell=1}^{L}\|\mathbf{a}_\ell\|^2}\mathbf{1}_{L-1}\mathbf{1}_{L-1}^T}{1-\frac{\sum_{\ell=2}^{L}\|\mathbf{a}_\ell\|^2}{\sum_{\ell=1}^{L}\|\mathbf{a}_\ell\|^2}} \end{array}\right] = \frac{\sigma^2}{2|s|^2}\left(\mathrm{diag}\{\|\mathbf{a}_2\|^{-2},\ldots,\|\mathbf{a}_L\|^{-2}\} + \frac{1}{\|\mathbf{a}_1\|^2}\mathbf{1}_{L-1}\mathbf{1}_{L-1}^T\right). \tag{41}$$

Substituting $\boldsymbol{X}_{22}$ and $\boldsymbol{J}_{\phi\theta}$ in $\boldsymbol{J}_{\phi\theta}^T\boldsymbol{X}_{22}\boldsymbol{J}_{\phi\theta}$, we get

$$\begin{aligned} &\boldsymbol{J}_{\phi\theta}^T\boldsymbol{X}_{22}\boldsymbol{J}_{\phi\theta} \\ &= \frac{2}{\sigma^2}|s|^2\left(\sum_{\ell=2}^{L}\frac{\mathrm{Im}^2\left\{\mathbf{a}_\ell^H\dot{\mathbf{a}}_\ell\right\}}{\|\mathbf{a}_\ell\|^2} + \frac{\left(\mathrm{Im}\left\{\sum_{\ell=2}^{L}\mathbf{a}_\ell^H\dot{\mathbf{a}}_\ell\right\}\right)^2}{\|\mathbf{a}_1\|^2}\right) \\ &= \frac{2}{\sigma^2}|s|^2\sum_{\ell=1}^{L}\frac{\mathrm{Im}^2\left\{\mathbf{a}_\ell^H\dot{\mathbf{a}}_\ell\right\}}{\|\mathbf{a}_\ell\|^2}, \end{aligned} \tag{42}$$

where the last equality follows from $\sum_{\ell=2}^{L}\mathbf{a}_\ell^H\dot{\mathbf{a}}_\ell = \mathbf{a}^H\dot{\mathbf{a}} - \mathbf{a}_1^H\dot{\mathbf{a}}_1 = \mathbf{a}_1^H\dot{\mathbf{a}}_1$. Finally, substituting (42) and $\boldsymbol{J}_{\theta\theta}$ in (38) gives

$$CRLB^{-1}(\theta) = \frac{2}{\sigma^2}|s|^2\sum_{\ell=1}^{L}\left(\|\dot{\mathbf{a}}_\ell\|^2 - \frac{\mathrm{Im}^2\left\{\mathbf{a}_\ell^H\dot{\mathbf{a}}_\ell\right\}}{\|\mathbf{a}_\ell\|^2}\right). \tag{43}$$

Let us compare the CRLB in (43) to the CRLB for the case where each sub-array observes a different signal of the sources. Clearly, the latter CRLB bounds the performance of an estimator which is based on minimizing (7). This bound can be obtained by modifying the CRLB that has been derived in [3] for coherent arrays: The inverse CRLB (for a single snapshot) in [3] can be used for each sub-array independently, and the obtained L terms can be summed up, similar to L snapshots. Therefore, we obtain the following result

$$CRLB_{\text{unsuited}}^{-1}(\theta) = \frac{2}{\sigma^2}|s|^2\sum_{\ell=1}^{L}\left(\|\dot{\mathbf{a}}_\ell\|^2 - \frac{\left|\mathbf{a}_\ell^H\dot{\mathbf{a}}_\ell\right|^2}{\|\mathbf{a}_\ell\|^2}\right). \tag{44}$$

In the context of our paper, we consider (44) as the "unsuited CRLB", because it assumes each sub-array observes a different signal of the sources, which does not correspond with our model (1).

Let us compare (43) to (44). Note that in general $\mathrm{Im}^2\left\{\mathbf{a}_\ell^H\dot{\mathbf{a}}_\ell\right\} \le \left|\mathbf{a}_\ell^H\dot{\mathbf{a}}_\ell\right|^2$. Therefore, $CRLB(\theta) \le CRLB_{\text{unsuited}}(\theta)$. However, the two terms may be equal in certain cases. For example, if $\{\mathbf{a}_\ell(\theta)\}$ depend on θ only through their phase (e.g. in ULAs of omnidirectional sensors), we have $\dot{\mathbf{a}}_\ell = j\mathrm{diag}\{\frac{d\angle[\mathbf{a}_\ell]_1}{d\theta},\ldots,\frac{d\angle[\mathbf{a}_\ell]_{M_\ell}}{d\theta}\}\mathbf{a}_\ell$. Therefore, $\mathrm{Re}\left\{\mathbf{a}_\ell^H\dot{\mathbf{a}}_\ell\right\} = \mathrm{Re}\left\{j\sum_{m=1}^{M_\ell}\frac{d\angle[\mathbf{a}_\ell]_m}{d\theta}|[\mathbf{a}_\ell]_m|^2\right\} = 0$. This result implies that for ULAs and $Q=1$ using the joint MLE that minimizes (14) may not be better than estimation with the non-coherent cost function in (7).

VI. Numerical Results

In this section, we perform computer simulations in order to examine the performance of AOA estimation by Algorithm 2 and Algorithm 3, which will be referred to as joint-MLE and joint-AM, respectively. We compare our approach with two non-coherent processing methods. The first method is based on minimizing (7), similarly to the estimator used in [10] (as discussed in Section III). The second method is based on MUSIC [13] and treats the realizations of the L sub-arrays as L snapshots of a single sub-array. We use this method with spatial smoothing [33] and forward-backward smoothing [34] (within each sub-array) that empirically improved its performance. Applying this non-coherent MUSIC is possible because in our experiments the sub-arrays are ULAs with similar sensor configuration.

To evaluate the performance of the AOAs estimators we present the CRLB that is obtained in (35) for non-coherent sub-arrays, which will be referred to as the "non-coherent CRLB". Furthermore, to demonstrate the performance gap from the coherent case, i.e. from the case where all the $M = \sum_{\ell=1}^{L}M_\ell$ elements are coherent, we present the CRLB that is derived in [3], which will be referred to as the "coherent CRLB". Since CRLBs are often loose, we also present the results of the AOA MLE in the coherent case, to which $\{\phi_\ell\}$ are *known*. The performance of the latter can be regarded as a bound for the performance of any AOA estimation technique that includes (or avoids) estimation of the phase offsets in the non-coherent sub-arrays case.

The first examined scenario demonstrates the performance for a single source ($Q=1$). We consider a ULA of $M=12$ elements, with half wavelength spacing, partitioned into $L=2$ sub-arrays of size $M_\ell = 6$. The single source is located at angle 8° (the boresight angle of the entire array is 0°), and its transmitted signal is a proper complex Gaussian random variable. The unknown phase of each sub-array is selected at random from a uniform distribution over $[0, 2\pi]$. The signal to noise ratio (SNR) value is changed from 0 dB to 30 dB. To obtain statistical results, at each SNR we perform $N_{\text{exp}} = 500$ Monte Carlo experiments, where the Gaussian noise $\{\boldsymbol{n}_\ell\}$, the signal s, and the unknown phases $\{\phi_\ell\}$ are drawn in each experiment.

We compare our joint-MLE approach with the coherent MLE and the non-coherent estimator that minimizes (7) (for a single source there is no point to apply MUSIC or AM). In Fig. 1 we present the AOA root mean square error (RMSE) for the different methods, which is defined by

$$\text{RMSE} = \sqrt{\frac{1}{N_{\text{exp}}\cdot Q}\sum_{i=1}^{N_{\text{exp}}}\|\hat{\boldsymbol{\theta}}(i) - \boldsymbol{\theta}\|^2}, \tag{45}$$

where $\hat{\boldsymbol{\theta}}(i)$ is the estimated AOA at the i-th experiment and $\boldsymbol{\theta}$ is the true AOA. The results agree with the analysis in Section V-B: For a single source, estimating the phases of non-coherent ULA's sub-arrays is not effective, and using non-coherent processing is suffice (as it attains the "non-coherent CRLB" even at quite low SNR).

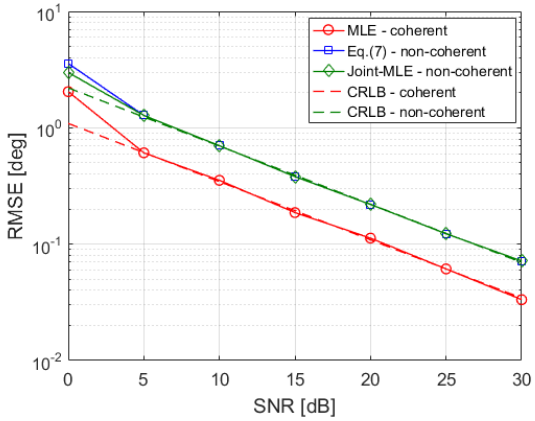


Fig. 1: RMSE vs. SNR for a single source at $8°$, using 2 uniform linear sub-arrays, each of 6 elements.

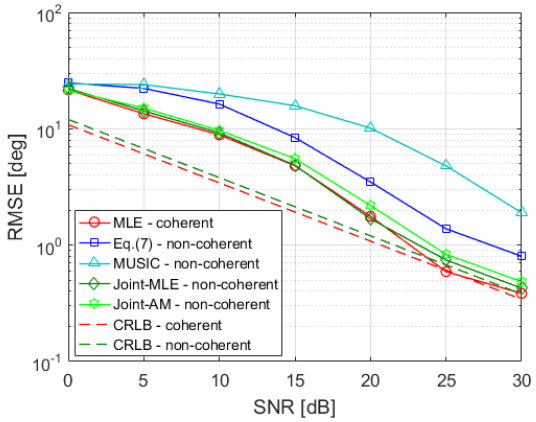


Fig. 2: RMSE vs. SNR for sources at $5°$ and $8°$, using 2 uniform linear sub-arrays, each of 6 elements.

Therefore, to demonstrate the usefulness of our method, we turn to examine a scenario with multiple sources: we add an additional source located at angle $5°$ (thus $Q = 2$). The two sources transmit Gaussian signals with equal power. The performance of the estimation techniques is measured using (45), where $\hat{\boldsymbol{\theta}}(i)$ is permuted such that $\|\hat{\boldsymbol{\theta}}(i) - \boldsymbol{\theta}\|^2$ is minimal. Since it is still tractable to perform a 2-dimensional search, estimations that are based on the global minimizers of the cost functions can be obtained. Yet, we also present the performance of our joint-AM method, which require only one-dimensional searches and is thus faster. The results are presented in Fig. 2. The joint-MLE is better than the non-coherent processing methods, and has very similar performance to the MLE of the *coherent* case. A small performance gap due to the non-coherent sub-arrays is observed only at high SNR, and apparently it is unavoidable, as demonstrated by the gap between the CRLBs for the coherent and non-coherent cases. Remarkably, the performance of joint-AM is very close to the performance of joint-MLE.

Note that so far we considered $L = 2$ sub-arrays. Therefore, the joint-MLE result given by Algorithm 2 is equivalent to the one obtained by minimizing (14) by a $Q + L$ dimensional search over $\boldsymbol{\theta}$ and $\{\phi_\ell\}$. Now, we turn to examine a scenario with $L = 4$ sub-arrays of size $M_\ell = 6$, i.e. in this scenario the entire ULA has $M = 24$ elements with half wavelength spacing. The other configuration settings are kept as before (e.g. the sources are at $5°$ and $8°$). Note that in this scenario Algorithm 2 yields an *approximate* MLE. The results are presented in Fig. 3. Our joint-MLE clearly outperforms the non-coherent estimators, and has only a small performance gap from the MLE of the coherent case. Again, this gap is unavoidable, as demonstrated by the CRLB for the coherent and non-coherent cases. The fact that the proposed joint-MLE estimator attains the "non-coherent CRLB", even though it is not an exact MLE for $L > 2$, shows the effectiveness of our phase estimation algorithm. Once again, the performance of joint-AM is very close to the performance of joint-MLE.

We turn to examine scenarios with $Q = 4$ sources, located at angles $-5°$, $0°$, $5°$ and $8°$, and transmitting Gaussian signals

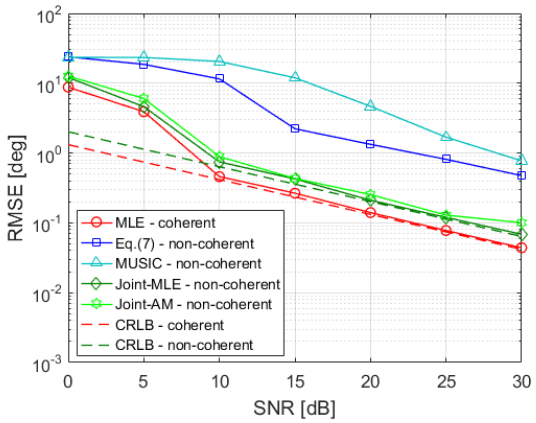


Fig. 3: RMSE vs. SNR for sources at $5°$ and $8°$, using 4 uniform linear sub-arrays, each of 6 elements.

with equal power. In this case, multidimensional search is not feasible. Therefore, we obtain the estimators via alternating minimization (AM) schemes applied on the cost functions, as explained in Section IV-B. The results for $L = 2$ uniform linear sub-arrays (each of 6 elements) and $L = 4$ uniform linear sub-arrays (each of 6 elements) are presented in Figs. 4 and 5, respectively. In both scenarios the proposed joint-AM is clearly better than the non-coherent processing reference methods. For $L = 2$ the gap from the "coherent AM" [16] (to which $\{\phi_\ell\}$ are known) is almost negligible. All the AM-based estimators (including the one for the coherent case) suffer from being trapped in local minima, which is implied by the fact that they are way above the CRLB and exhibit only moderate improvement as the SNR grows. This behavior is expected, because the number of sources is high and they are quite close. To mitigate this limitation, future work may consider incorporating the proposed phase estimation strategy with computationally tractable methods that may estimate the AOAs of dense sources better than AM. One such example is AOA estimation based on deep neural networks [25].

In order to compare the methods for $Q = 4$ sources, without the disastrous effect of being trapped in bad local minima, let

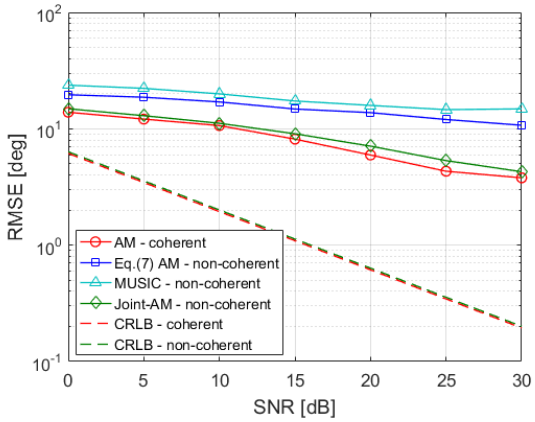


Fig. 4: RMSE vs. SNR for sources at $-5°$, $0°$, $5°$ and $8°$, using 2 uniform linear sub-arrays, each of 6 elements.

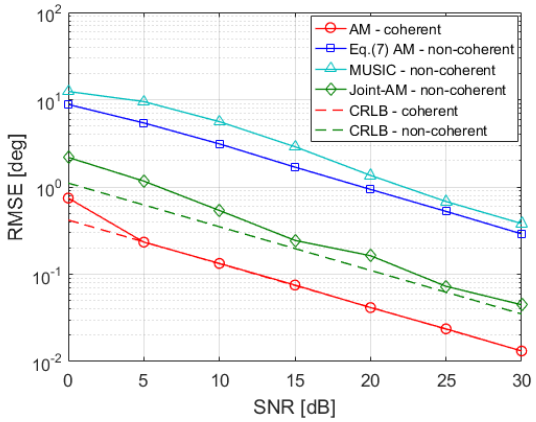


Fig. 6: RMSE vs. SNR for sources at $-15°$, $0°$, $15°$ and $30°$, using 4 uniform linear sub-arrays, each of 6 elements.

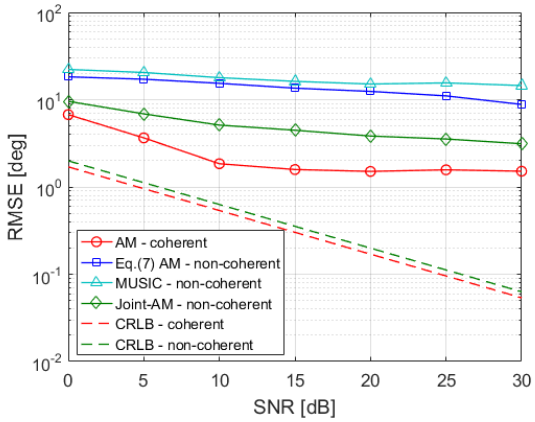


Fig. 5: RMSE vs. SNR for sources at $-5°$, $0°$, $5°$ and $8°$, using 4 uniform linear sub-arrays, each of 6 elements.

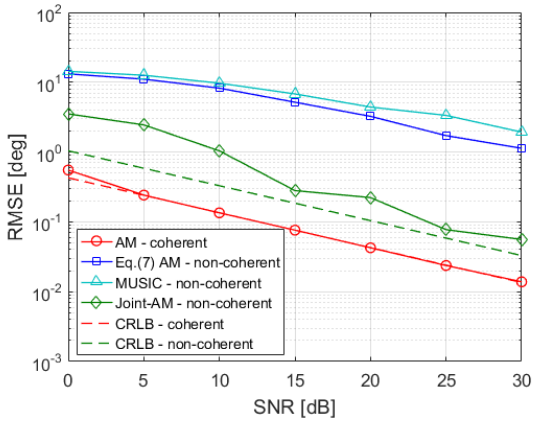
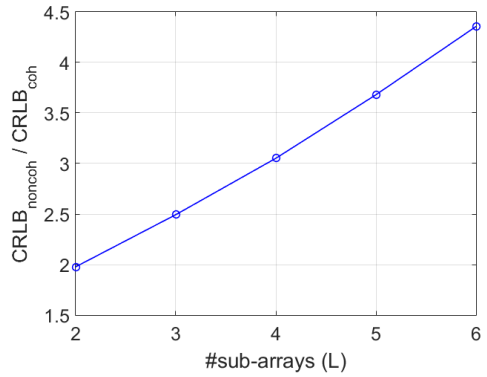


Fig. 7: RMSE vs. SNR for sources at $-30°$, $-15°$, $0°$, $15°$ and $30°$, using 4 uniform linear sub-arrays, each of 6 elements.

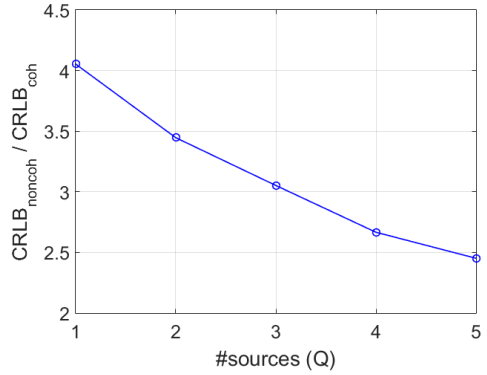
us consider a simpler scenario, for which the initialization strategies used for AM are good enough. Therefore, we consider the configuration of the $L = 4$ uniform linear 6-element sub-arrays, but this time with larger AOAs differences, i.e. the sources are located at angles $-15°$, $0°$, $15°$ and $30°$. Furthermore, we apply joint-AM (Algorithm 3) with nine different AOAs initializations and choose the AOAs hypothesis with the lowest negative log-likelihood value as the estimator. The initializations are made as follows. The first initialization is the AOAs estimated by AM applied on (7) (as done before), for which Algorithm 3 yields a new estimator, denoted by $\hat{\boldsymbol{\theta}}_{joint}$. Then, we apply Algorithm 3 eight more times, where at each time it is initialized with $\hat{\boldsymbol{\theta}}_{joint}$ modified by $\pm 1.5°$ in one of the AOAs. Even though we overall run Algorithm 3 nine times, each run requires only a few iterations (typically about five) consisted of one-dimensional searches. Therefore, our method is still much faster than any alternative that requires multidimensional search. The results are presented in Fig. 6. This time the "coherent AM" attains the "coherent CRLB", and the proposed joint-AM is very close to the "non-coherent CRLB". Most of the gap between joint-AM and "coherent AM" is unavoidable, as demonstrated by the CRLBs for the coherent and non-coherent cases. Once again, our method is much better than the non-coherent processing reference methods.

As a final demonstration of the effectiveness of the proposed joint-AM, we repeat the previous experiment with an additional source located at angle $-30°$ (i.e. now there are $Q = 5$ sources with AOAs of $\{0°, \pm 15°, \pm 30°\}$). For joint-AM we use the strategy of applying Algorithm 3 with different initializations, as described above. The results are presented in Fig. 7, and show again that the joint-AM significantly outperforms the non-coherent processing reference methods and is quite close to the "non-coherent CRLB".

Let us summarize the major observations from the experiments. We see that when there are multiple sources the proposed joint-MLE and joint-AM clearly outperform the non-coherent processing methods and attain a smaller performance gap from the MLE of the coherent case (to which the sub-arrays phase offsets are known). Examining the performance gap between the proposed joint-MLE (or joint-AM) and the MLE for coherent sub-arrays shows two interesting trends: (a) when the number of sub-arrays L grows the gap increases, and



(a)



(b)

Fig. 8: Ratio between the CRLBs for the non-coherent and coherent cases, for: (a) $Q = 3$ sources and the number of sub-arrays (L) is varied; (b) $L = 4$ sub-arrays and the number of sources (Q) is varied. See the text for full details.

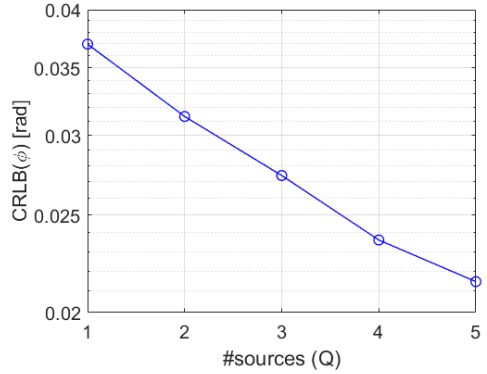


Fig. 9: The CRLB for estimating $\{\phi_\ell\}$ (in the non-coherent case) for $L = 4$ sub-arrays and varying number of sources Q, at SNR of 30 dB. See the text for full details.

(b) when the number of sources Q grows the gap decreases.

The two trends can be also observed by examining the ratio between the non-coherent and coherent CRLBs. Since the Fisher information matrix in (31) (whose terms appear in (34)) does not seem to have a tractable structure that allows to analytically compare the resulting "non-coherent CRLB" to the "coherent CRLB" in [3], we compare them numerically. In Fig. 8a we present the CRLB ratio (non-coherent/coherent) for $Q = 3$ sources, located at angles $\{0°, \pm 15°\}$, versus the number of sub-arrays L (each of 6 elements, as before), which varies from 2 to 6. In Fig. 8b we present the CRLB ratio for $L = 4$ sub-arrays (each of 6 elements) versus the number of sources Q, which is varied from 1 to 5 (with associated AOAs of $\{0°\}, \{0°, 15°\}, \{0°, \pm 15°\}, \{0°, \pm 15°, 30°\}$ and $\{0°, \pm 15°, \pm 30°\}$).

In Fig. 8a we see that increasing the number of sub-arrays increases the gap (ratio) between the CRLBs. This behavior results from the fact that while each additional sub-array does not add any unknown parameter in the coherent case, it does add a nuisance parameter (the unknown phase offset) in the non-coherent case. In Fig. 8b we see that increasing the number of sources reduces the gap between the CRLBs, even though in both coherent and non-coherent settings each additional source adds the same unknown parameters (another entry in $\boldsymbol{\theta}$ and in $\boldsymbol{s}$). To better understand this phenomenon, we show in Fig. 9 the CRLB for estimating the sub-arrays phase offsets $\{\phi_\ell\}$ versus the number of sources Q (for SNR of 30 dB and the same settings that are used in Fig. 8b). It is realized that the CRLB for estimating $\{\phi_\ell\}$ reduces with the increase in Q. This explains why the CRLB ratio for the AOAs estimation in Fig. 8b reduces with the increase in Q. The intuition for the behavior in Fig. 9 is that having more sources provides more constraints for matching the phase offsets to the observations, and thus improves the estimation of the phase offsets.

VII. Conclusions

We developed estimators for the AOAs of multiple sources from a single snapshot of an array which is consisted of non-coherent sub-arrays. To this end, we derived closed-form estimators for the sub-arrays phase offsets, which depend on the AOAs hypothesis, and then applied (approximate) MLE of the AOAs. While for a single source we showed (both empirically and by analyzing the CRLB) that estimating the sub-arrays phase offsets may not be useful, for multiple sources the proposed approach clearly outperforms non-coherent processing methods, and even attains the CRLB in various scenarios. Furthermore, for multiple sources the proposed approach has also shown that in some settings it attains close performance to MLE of the AOAs in the coherent case. From the simulation results and the CRLB we conclude that when the number of sources increases — the performance gap between non-coherent and coherent sub-arrays reduces. On the other hand, when the number of sub-arrays increases, more nuisance parameters need to be estimated, and the gap increases.

References

[1] H. Krim and M. Viberg, "Two decades of array signal processing research," *IEEE Signal Process. Mag.*, 1996.

[2] H. L. Van Trees, *Optimum array processing: Part IV of detection, estimation, and modulation theory*. John Wiley & Sons, 2004.

[3] P. Stoica and A. Nehorai, "MUSIC, maximum likelihood, and Cramér-Rao bound," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 37, no. 5, pp. 720–741, 1989.

[4] M. D. Zoltowski and K. T. Wong, "Closed-form eigenstructure-based direction finding using arbitrary but identical subarrays on a sparse uniform cartesian array grid," *IEEE Trans. Signal Process.*, vol. 48, no. 8, pp. 2205–2210, 2000.

[5] M. Pesavento, A. B. Gershman, and K. M. Wong, "Direction of arrival estimation in partly calibrated time-varying sensor arrays," in *IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP)*, vol. 5, pp. 3005–3008, IEEE, 2001.

[6] C. M. S. See and A. B. Gershman, "Direction-of-arrival estimation in partly calibrated subarray-based sensor arrays," *IEEE Trans. Signal Process.*, vol. 52, no. 2, pp. 329–338, 2004.

[7] D. W. Rieken and D. R. Fuhrmann, "Generalizing MUSIC and MVDR for multiple noncoherent arrays," *IEEE Trans. Signal Process.*, vol. 52, no. 9, pp. 2396–2406, 2004.

[8] P. Parvazi, M. Pesavento, and A. B. Gershman, "Direction-of-arrival estimation and array calibration for partly-calibrated arrays," in *IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP)*, pp. 2552–2555, IEEE, 2011.

[9] F. Wen, Q. Wan, R. Fan, and H. Wei, "Improved MUSIC algorithm for multiple noncoherent subarrays," *IEEE Signal Process. Lett.*, vol. 21, no. 5, pp. 527–530, 2014.

[10] W. Suleiman, P. Parvazi, M. Pesavento, and A. M. Zoubir, "Non-coherent direction-of-arrival estimation using partly calibrated arrays," *IEEE Trans. Signal Process.*, vol. 66, no. 21, pp. 5776–5788, 2018.

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

[12] A. L. Swindlehurst, P. Stoica, and M. Jansson, "Exploiting arrays with multiple invariances using MUSIC and MODE," *IEEE Trans. Signal Process.*, vol. 49, no. 11, pp. 2511–2521, 2001.

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

[14] 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, 1989.

[15] A. Barabell, "Improving the resolution performance of eigenstructure-based direction-finding algorithms," in *IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP)*, vol. 8, pp. 336–339, IEEE, 1983.

[16] I. Ziskind and M. Wax, "Maximum likelihood localization of multiple sources by alternating projection," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 36, no. 10, pp. 1553–1560, 1988.

[17] Y. Bresler and A. Macovski, "Exact maximum likelihood parameter estimation of superimposed exponential signals in noise," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 34, no. 5, pp. 1081–1089, 1986.

[18] M. Feder and E. Weinstein, "Parameter estimation of superimposed signals using the EM algorithm," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 36, no. 4, pp. 477–489, 1988.

[19] B. D. Jeffs, "Sparse inverse solution methods for signal and image processing applications," in *IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP)*, vol. 3, pp. 1885–1888, IEEE, 1998.

[20] 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, 2005.

[21] Z. Yang, J. Li, P. Stoica, and L. Xie, "Sparse methods for direction-of-arrival estimation," in *Academic Press Library in Signal Processing, Volume 7*, pp. 509–581, Elsevier, 2018.

[22] Z. Tian, Z. Zhang, and Y. Wang, "Low-complexity optimization for two-dimensional direction-of-arrival estimation via decoupled atomic norm minimization," in *IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP)*, pp. 3071–3075, IEEE, 2017.

[23] Z. Zhang, Y. Wang, and Z. Tian, "Efficient two-dimensional line spectrum estimation based on decoupled atomic norm minimization," *arXiv preprint arXiv:1808.01019*, 2018.

[24] Y. Wang and Z. Tian, "IVDST: A fast algorithm for atomic norm minimization in line spectral estimation," *IEEE Signal Process. Lett.*, vol. 25, no. 11, pp. 1715–1719, 2018.

[25] O. Bialer, N. Garnett, and T. Tirer, "Performance advantages of deep neural networks for angle of arrival estimation," in *IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP)*, pp. 3907–3911, IEEE, 2019.

[26] A. J. Weiss, "Direct position determination of narrowband radio frequency transmitters," *IEEE Signal Process. Lett.*, vol. 11, no. 5, pp. 513–516, 2004.

[27] O. Bialer, D. Raphaeli, and A. J. Weiss, "Maximum-likelihood direct position estimation in dense multipath," *IEEE Trans. Veh. Technol.*, vol. 62, no. 5, pp. 2069–2079, 2012.

[28] L. Tzafri and A. J. Weiss, "High-resolution direct position determination using MVDR," *IEEE Trans. Wireless Commun.*, vol. 15, no. 9, pp. 6449–6461, 2016.

[29] T. Tirer and A. J. Weiss, "Performance analysis of a high-resolution direct position determination method," *IEEE Trans. Signal Process.*, vol. 65, no. 3, pp. 544–554, 2016.

[30] D. Tollefsen, P. Gerstoft, and W. S. Hodgkiss, "Multiple-array passive acoustic source localization in shallow water," *The Journal of the Acoustical Society of America*, vol. 141, no. 3, pp. 1501–1513, 2017.

[31] X. Zhang, K. Huo, Y. Liu, and X. Li, "Direction of arrival estimation via joint sparse bayesian learning for bi-static passive radar," *IEEE Access*, vol. 7, pp. 72979–72993, 2019.

[32] T. Tirer and O. Bialer, "Effective approximate maximum likelihood estimation of angles of arrival for non-coherent sub-arrays," in *IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP)*, pp. 4557–4561, IEEE, 2020.

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

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

Tom Tirer received the B.Sc. degree (summa cum laude) in electrical and computer engineering from Ben-Gurion University of the Negev, Beersheba, Israel, in 2010, and the M.Sc. degree (summa cum laude) in electrical engineering from Tel Aviv University, Tel Aviv, Israel, in 2016. He is currently working toward the Ph.D. degree at Tel Aviv University. From 2010 to 2015, he was with the Israeli Semiconductor Industry, where he was involved in the design and development of 3GPP LTE modems. Since 2017, he has been developing signal processing and machine learning algorithms for radars mounted on autonomous cars. His research interests include signal and image processing, optimization, compressed sensing, and machine learning. He received the Weinstein Prize for research in signal processing (2016, 2017, and 2019) and the KLA excellence award (2020).

Oded Bialer received his B.Sc., M.Sc. and Ph.D. degrees in Electrical Engineering from Tel Aviv University, Tel Aviv, Israel, in 2000, 2008, and 2016, respectively. His area of expertise is signal processing and deep learning. Since 1999 he has been developing cutting-edge technologies for several high-tech companies. He is currently working at General Motors R&D, where he is developing smart sensing systems and perception algorithms for autonomous driving. His research interests include statistical signal processing, detection and estimation, sensor array processing, source localization, machine learning, and deep learning.