Delft University of Technology

Self-Calibration of Acoustic Scalar and Vector Sensor Arrays

Ramamohan, Krishnaprasad Nambur; Chepuri, Sundeep Prabhakar; Comesana, Daniel Fernandez; Leus, Geert

Important note
To cite this publication, please use the final published version (if applicable).
Please check the document version above.

Self-Calibration of Acoustic Scalar and Vector Sensor Arrays

Krishnaprasad Nambur Ramamohan, *Student Member,* Sundeep Prabhakar Chepuri, *Member, IEEE,* Daniel Fernandez Comesana, and Geert Leus, *Fellow, IEEE*

Abstract—In this work, we consider the self-calibration problem of joint calibration and direction-of-arrival (DOA) estimation using acoustic sensor arrays. Unlike many previous iterative approaches, we propose solvers that can be readily used for both linear and non-linear arrays for jointly estimating the sensor gain, phase errors, and the source DOAs. We derive these algorithms for both the conventional element-space and covariance data models. We focus on sparse and regular arrays formed using scalar sensors as well as vector sensors. The developed algorithms are obtained by transforming the underlying non-linear calibration model into a linear model, and subsequently by using convex relaxation techniques to estimate the unknown parameters. We also derive identifiability conditions for the existence of a unique solution to the self-calibration problem. To demonstrate the effectiveness of the developed techniques, numerical experiments, and comparisons to the state-of-the-art methods are provided. Finally, the results from an experiment that was performed in an anechoic chamber using an acoustic vector sensor array are presented to demonstrate the usefulness of the proposed self-calibration techniques.

I. Introduction

The problem of estimating the direction-of-arrival (DOA) of multiple far-field events impinging on an array of spatially distributed sensors has received considerable interest in various fields including communications, radio astronomy, acoustics, and seismology. Usually scalar sensor arrays, such as acoustic pressure sensor (APS) arrays, are used for DOA estimation. In recent times, transducers that measure vector quantities are becoming practically feasible [1], enabling new processing capabilities. An acoustic vector sensor (AVS) is such a device that is capable of measuring both the acoustic pressure and particle velocity. Unlike an APS, a single AVS can measure the DOA of a far-field event [2] and arrays of such AVSs have proven to have distinct advantages compared to conventional microphone arrays [3]. In practice, all sensors and their arrays are highly sensitive to model errors [4]. Among those model errors, the gain and phase mismatches between sensors, known as calibration errors are the dominant ones that degrade the DOA estimation results. Those are the focus of this work.

Many advanced subspace based algorithms, e.g., multiple signal classification (MUSIC) [5], minimum variance distortionless response (MVDR) [6], and estimation of signal parameters via rotational invariance technique (ESPRIT) [7], have

K. Nambur Ramamohan is with Microflown AVISA, NL. S. P. Chepuri is with the Department of Electrical Communications Engineering, Indian Institute of Science, India. D. F. Comesana is with Microflown Technologies, Arnhem, NL. G. Leus is with the Faculty of EEMCS, Delft University of Technology. This work was supported in part by the Pratiksha Trust Fellowship, IISc and is part of the ASPIRE project (project 14926 within the STW OTP).

been developed for DOA estimation. Further sparse recovery techniques have also been widely used, whenever only a few sources are present [8]. These traditional algorithms require more physical sensors than the number of sources and use the data acquired in the *element-space* domain (i.e., at the output of the sensor elements) or in the covariance (or *co-array*) domain. Also nowadays, to reduce sensing and data processing costs, sparse sensing methods are gaining attention [9]. One can resolve and estimate DOAs of as many as $\mathcal{O}(M^2)$ sources using only M physical elements by smartly and irregularly placing the sensor elements. Such sensor placements are generally referred to as sparse arrays [10]–[12]. Most of the discussed algorithms developed for APS arrays can be used directly or adapted for equivalent AVS arrays as well [2], [3].

DOA estimates obtained from these aforementioned standard algorithms in the presence of calibration errors are severely degraded [4]. These errors originate from the variability in the analog electronics and the manufacturing technology across sensors in the array. They affect both the signal-of-interest and the noise part of the measurement data [4], [13]. On the other hand, it is also possible to have calibration errors that only affect the signal part of the measurements [4], whenever there are position or orientation errors of the sensors (channels) in the array for instance or perturbations in the sensors' gain and phase patterns. Usually labor intensive and expensive calibration procedures are applied to correct for these mismatches [14]–[16], which are impractical for large number of sensors. Furthermore, such calibration errors vary with time and changes in the environment, and as a result, the deployed sensors require periodic re-calibration. In such scenarios, self-calibration methods are inevitable. The term *self-calibration* refers to using the information collected by the array to simultaneously estimate the calibration errors and source DOAs without any reference sources with known direction and/or pre-defined waveform.

A. Self-calibration methods

Self-calibration techniques for scalar sensor arrays in the presence of gain and phase uncertainties between sensors have been widely studied [13], [17]–[21]. It is a non-linear estimation problem with unknown calibration and array manifold matrix. Specific conditions should be satisfied such that they are independently identifiable [17], [19]. In some cases it is impossible to independently resolve both of them [17], [22].

Maximum likelihood (ML) and maximum a posteriori (MAP) based gain and phase estimation algorithms have

This article has been accepted for publication in IEEE Transactions on Signal Processing. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TSP.2022.3214383

2

been proposed in [23], [24] to solve the self-calibration problem. Although both ML and MAP based estimators are asymptotically efficient, they are computationally expensive and not suitable for practical applications. In contrast, there are also computationally friendly self-calibration techniques which can be broadly classified into two categories: geometry-dependent and geometry-independent approaches.

The first kind of self-calibration techniques are developed for specific/regular array geometries, where spatial redundancies are used to eliminate the array manifold information in order to estimate the calibration errors. Specifically, in [13], a self-calibration procedure for scalar sensor arrays arranged in a uniform linear array (ULA) configuration was presented, where the Toeplitz structure of the data covariance matrix was utilized. Extensions and adaptations of this self-calibration approach are presented in [18], [25]–[27]. A self-calibration technique based on ESPRIT for an APS ULA is derived in [17].

The second kind of self-calibration approaches are applicable to arbitrary array geometries where the array manifold and calibration matrices are estimated mostly with iterative techniques [19], [20], [28]. However, these approaches suffer from the choice of the initial estimate of the calibration errors and the algorithm might only converge to a local minimum leading to a sub-optimal solution.

Aforementioned two categories of self-calibration approaches are developed specifically for scenarios with more sensors than sources. Nonetheless, with the increased attention on sparse sensing, a self-calibration algorithm for sparse arrays was proposed in [29], where a sub-optimal method was used to estimate the phase errors.

On the other hand, apart from the DOA estimation techniques proposed in [30]–[34], not much attention is given to the self-calibration problem for AVS arrays. It may seem possible to adapt the second category of aforementioned self-calibration techniques for vector sensor arrays. However, due to the dissimilarities in the array manifold, the conditions for independently identifying the calibration errors and source DOAs are different and are not yet available.

In summary, we can observe that existing self-calibration approaches either require specific sensor placement to obtain an optimal solution or converge to a sub-optimal solution without geometry constrains. So in this work we try to address the issues associated with existing techniques by proposing two non-iterative *one-step* self-calibration algorithms that are array geometry independent, and applicable to both APS and AVS arrays. Furthermore, the proposed self-calibration technique using the co-array measurement model is also applicable to sparse arrays, whose preliminary results are presented in [35]. In our work, we leverage tools from sparse recovery techniques and draw inspiration from [21], which deals with the self-calibration problem for linear models, to decouple the calibration parameters from the other unknowns. However, the model we deal with is not linear anymore as the source directions are not known. In essence, the main problem of interest in this work is self-calibration with a non-linear measurement model, where we assume that the calibration errors are mainly originating from uncertainties in the analog

electronics and sensor elements. Additionally, we also derive the identifiablity conditions for a unique solution to exist while using AVS and sparse APS arrays.

B. Our contributions

This work introduces novel self-calibration methodologies combining traditional array processing theory with sparse recovery techniques. The validity of the proposed self-calibration algorithms is studied by considering the measurement model where the calibration errors affect both the signal-of-interest and noise. Further the adaptation of the proposed approaches to the measurement model where the calibration errors affect only the signal component of the data is also discussed. The main contributions can be summarized as follows:

- We develop non-iterative novel self-calibration algorithms that are applicable for both linear and non-linear arrays based on both the element-space and co-array data models, where the latter data model is even useful when there are more sources than sparsely placed sensors.
- We derive conditions to ensure a unique solution for estimating the DOAs and calibration errors for both AVS and sparse APS arrays. This important aspect is still not considered in the existing literature.
- We demonstrate the validity of the proposed approaches via numerical simulations as well as an experimental study. The latter shows the effectiveness of the introduced self-calibration techniques for an array of 4 AVSs measured in an anechoic chamber.

C. Notation and outline

Upper (lower) bold face letters are used for matrices (column vectors); $(\cdot)^*$ denotes conjugate, $(\cdot)^T$ denotes transpose and $(\cdot)^H$ denotes conjugate transpose; $\otimes$ denotes the Kronecker product, $\circ$ denotes the Khatri-Rao product and $\odot$ denotes the Schur-Hadamard (element-wise) product; $\mathbb{E}\{\cdot\}$ denotes the expectation operator; $\mathrm{tr}(\cdot)$ denotes the trace operator and $\mathbf{I}_n$ is the identity matrix of dimension n.

The detailed outline for this paper is as follows. In Section II, we present the measurement model with calibration errors, and the problem statement of estimating both the calibration errors and the DOAs. In Section III, we present the identifiability conditions for uniquely estimating the calibration errors and the source DOAs. In Section IV and Section V, the proposed calibration algorithms based on the element-space and the co-array domain measurement data model are presented, respectively. The simulation and experimental results of the proposed calibration algorithms are discussed in Section VI and Section VII, respectively.

II. PROBLEM STATEMENT

Consider a linear array of M sensors with Q channels, where $Q = M$ for APS arrays and $Q = 3M$ for AVS arrays. We are interested in estimating the azimuth directions of N narrow-band sources, denoted by $\boldsymbol{\theta} = [\theta_1, \theta_2, \dots, \theta_N]^T$ with $\theta_n \in [0, \pi]$ for $n = 1, \dots, N$, where the azimuth directions are measured with respect to the phase reference of the array.

This article has been accepted for publication in IEEE Transactions on Signal Processing. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TSP.2022.3214383

3

Each of the considered Q channels has a different receiver gain and phase response, which are not known. We refer to the unknown receiver gains and phases as calibration errors, and for the ith channel it is denoted as $g_i = \alpha_i e^{j\phi_i}$ with α_i and ϕ_i being the gain and phase mismatch, respectively. We collect the calibration errors in the diagonal matrix $\text{diag}(\mathbf{g})$ with $\mathbf{g} = [g_1, g_2, \ldots, g_Q]^T$. Let us also define the vectors $\boldsymbol{\alpha} = [\alpha_1, \alpha_2, \ldots, \alpha_Q]^T$ and $\boldsymbol{\phi} = [\phi_1, \phi_2, \ldots, \phi_Q]^T$.

Under the narrow-band assumption [5], the element-space signal, $\mathbf{x}(t)$, can be modeled as [13]

$$\mathbf{x}(t) = \text{diag}(\mathbf{g}) \left[\mathbf{A}(\boldsymbol{\theta}) \, \mathbf{s}(t) + \mathbf{n}(t) \right] \in \mathbb{C}^{Q \times 1}, \tag{1}$$

where

$$\mathbf{A}(\boldsymbol{\theta}) = [\mathbf{a}(\theta_1) \cdots \mathbf{a}(\theta_N)] \in \mathbb{C}^{Q \times N}$$

is the array manifold matrix, the source signals of wavelength λ are stacked in the vector $\mathbf{s}(t) \in \mathbb{C}^{N \times 1}$ and the receiver noise vector is given by $\mathbf{n}(t) \in \mathbb{C}^{Q \times 1}$. Here, we assume that both $\mathbf{s}(t)$ and $\mathbf{n}(t)$ are derived from an independent and identically distributed (i.i.d.) Gaussian distribution. Let us define $\mathbf{p} = [\mathbf{p}_1^T, \mathbf{p}_2^T, \ldots, \mathbf{p}_M^T]^T$, where $\mathbf{p}_m = [p_{m1}, p_{m2}]^T$ is the position of the mth sensor in the array defined in terms of wavelength (λ) of the observed signal. In particular for a linear array the position vector of the mth sensor is modified as $\mathbf{p}_m = [p_{m1}, 0]^T$. For the sake of simplicity, we refer the mth sensor position within a linear array just as p_m. Further Without loss of generality (w.l.o.g.), we consider the first sensor with $p_{11} = p_{12} = 0$ as the phase reference of the array. The spatial signature (or the array steering vector) for the nth source in the direction described by the vector $\mathbf{u}(\theta_n) = [\cos(\theta_n) \ \sin(\theta_n)]^T$ with respect to the first sensor of the APS array with M sensors is given by

$$\mathbf{a}_{\text{APS}}(\theta_n) = \left[e^{\text{j}2\pi\left(\mathbf{p}_1^T \mathbf{u}(\theta_1)\right)}, \ \ldots \ , e^{\text{j}2\pi\left(\mathbf{p}_M^T \mathbf{u}(\theta_n)\right)} \right]^T \in \mathbb{C}^{M \times 1}, \tag{2}$$

whereas the related array steering vector of the AVS array is given by

$$\begin{aligned} \mathbf{a}_{\text{AVS}}(\theta_n) &= \begin{bmatrix} 1 & \mathbf{u}^T(\theta_n) \end{bmatrix}^T \otimes \mathbf{a}_{\text{APS}}(\theta_n), \\ &= \mathbf{h}(\theta_n) \otimes \mathbf{a}_{\text{APS}}(\theta_n) \in \mathbb{C}^{3M \times 1}. \end{aligned} \tag{3}$$

For the APS array, we have $Q = M$ with $\mathbf{a}(\theta_n) = \mathbf{a}_{\text{APS}}(\theta_n)$ and for the AVS array we have $Q = 3M$ channels with $\mathbf{a}(\theta_n) = \mathbf{a}_{\text{AVS}}(\theta_n)$.

Usually the signal $\mathbf{x}(t)$ is uniformly sampled and L snapshots are collected in the data matrix $\mathbf{X} = [\mathbf{x}(1), \mathbf{x}(2), \ldots, \mathbf{x}(L)] \in \mathbb{C}^{Q \times L}$ to obtain

$$\mathbf{X} = \text{diag}(\mathbf{g}) \left[\mathbf{A}(\boldsymbol{\theta})\mathbf{S} + \mathbf{N} \right]. \tag{4}$$

Here, $\mathbf{S} = [\mathbf{s}(1), \mathbf{s}(2), \ldots, \mathbf{s}(L)] \in \mathbb{C}^{N \times L}$ and $\mathbf{N} = [\mathbf{n}(1), \mathbf{n}(2), \ldots, \mathbf{n}(L)] \in \mathbb{C}^{Q \times L}$. The covariance matrix of the signal $\mathbf{x}(t)$ is $\mathbf{R}_{\text{x}} = \mathbb{E}\{\mathbf{x}(t)\mathbf{x}^H(t)\} \in \mathbb{C}^{Q \times Q}$. We assume that the source signals $\mathbf{s}(t)$ are uncorrelated and have a diagonal covariance matrix $\mathbb{E}\{\mathbf{s}(t)\mathbf{s}^H(t)\} = \text{diag}(\boldsymbol{\sigma}_{\text{s}})$, which is not known. Similarly, the noise vector has a diagonal covariance matrix $\mathbb{E}\{\mathbf{n}(t)\mathbf{n}^H(t)\} = \text{diag}(\boldsymbol{\sigma}_{\text{n}})$, which is assumed to be

known or can be estimated. Then, the covariance domain model can be written as

$$\mathbf{R}_{\text{x}} = \text{diag}(\mathbf{g}) \left[\mathbf{A}(\boldsymbol{\theta})\text{diag}(\boldsymbol{\sigma}_{\text{s}})\mathbf{A}^H(\boldsymbol{\theta}) + \text{diag}(\boldsymbol{\sigma}_{\text{n}}) \right] \text{diag}^H(\mathbf{g}). \tag{5}$$

Here, it is assumed that $\mathbf{s}(t)$ and $\mathbf{n}(t)$ are mutually uncorrelated. It is also useful to express (5) in vectorized form as:

$$\mathbf{r}_{\text{x}} = \text{diag}(\mathbf{g}^* \otimes \mathbf{g}) \left[\mathbf{A}_{\text{co}}(\boldsymbol{\theta})\boldsymbol{\sigma}_{\text{s}} + \boldsymbol{\sigma}_{\text{n}} \right], \tag{6}$$

where $\text{vec}(\mathbf{R}_{\text{x}}) = \mathbf{r}_{\text{x}}$ and $\mathbf{A}_{\text{co}}(\boldsymbol{\theta}) = \mathbf{A}^*(\boldsymbol{\theta}) \circ \mathbf{A}(\boldsymbol{\theta})$ with the subscript "co" indicating the co-array manifold. In practice, the data matrix $\mathbf{X}$ is used to compute the sample data covariance matrix $\widehat{\mathbf{R}}_{\text{x}} = L^{-1}\mathbf{X}\mathbf{X}^H$. For the sake of convenience, henceforth, we use $\mathbf{R}_{\text{x}}$ instead of $\widehat{\mathbf{R}}_{\text{x}}$ with the knowledge that only an estimate of the covariance matrix is available.

Based on the co-array model in (6), the sensor elements can be smartly placed irregularly along the linear axis, such that $\mathbf{A}_{\text{co}}$ has full column rank. Usually such configuration of linear arrays leads to sparse array design [12] allowing one to resolve as many as $\mathcal{O}(M^2)$ sources using M sensors. As seen in (1) and (6), both $\mathbf{g}$ and $\boldsymbol{\theta}$ are unknowns, and additionally it is a non-linear estimation problem as $\boldsymbol{\theta}$ exists in the exponential terms of the array manifold matrix.

The main goal of this paper is to jointly estimate the Q complex (i.e., $2Q$ real) receiver gains $\mathbf{g}$ and N directions $\boldsymbol{\theta}$ given $\mathbf{X}$ or $\mathbf{r}_{\text{x}}$. To do so uniquely, as will be discussed in Section III, we will require a few reference sensors with known complex receiver gains in the array.

III. AMBIGUITY AND IDENTIFIABILITY

Before presenting the calibration algorithms, in this section, we discuss identifiability conditions under which a unique solution for both the calibration parameters and the source DOAs exists. The identifiability conditions for the APS arrays by considering the element-space model (4) is presented in [19]. We take inspiration from [19] and derive identifiability conditions for all the remaining measurement models relevant for both APS and AVS arrays. It should be immediately clear that, as both $\text{diag}(\mathbf{g})\mathbf{A}(\boldsymbol{\theta})$ and $\mathbf{S}$ (or $\boldsymbol{\sigma}_{\text{s}}$) are not known a priori, they cannot be computed uniquely as there will be a complex (or real) scaling ambiguity. Therefore, to fix the scaling ambiguity we perform calibration with respect to *sensor 1* at location $p_1 = 0$, i.e., we use $g_1 = 1$ for the element-space data model and $|g_1| = \alpha_1 = 1$ for the co-array data model.

After establishing the fact that the elements of $\mathbf{g}$ can only be estimated relative to the reference sensor, the next important question that needs to be addressed is to establish the well posedness of the self-calibration problem given the measurement data. From the element-space data model (4), we have $2QL$ nonlinear equations in N unknown DOAs, $2(Q-1)$ unknown calibration parameters, and $2NL$ unknown source signals. Hence, for well posedness of the calibration problem, we require

$$2QL \geq N + 2Q - 2 + 2NL \quad \Rightarrow \quad \frac{N + 2(Q-1)}{2(Q-N)} \leq L,$$

which is meaningful only for $Q > N$. Furthermore, from the co-array data model (6), we have $2QN - N^2 + 1$ nonlinear

equations[1] in N unknown DOAs, $2(Q-1)$ unknown calibration parameters, and N unknown source powers. Hence, for well posedness, we require

$$2QN - N^2 + 1 \geq 2N + 2Q - 2 \quad \Rightarrow \quad Q \geq \frac{N^2 + 2N - 3}{2(N-1)}.$$

Finally, we study under which conditions we can uniquely estimate $\mathbf{g}$ and $\boldsymbol{\theta}$ given the measurement data. However due to the non-linear nature of the estimation problem, it is not straightforward to derive the identifiability conditions based on the element-space data model (4) or co-array data model (6). Therefore, to begin with, for linear arrays we derive sufficient conditions for uniquely estimating $\mathbf{g}$ and $\boldsymbol{\theta}$ based on the assumption that $\operatorname{diag}(\mathbf{g})\mathbf{A}(\boldsymbol{\theta})$ $(\operatorname{diag}(\mathbf{g}^* \otimes \mathbf{g})\mathbf{A}_{\mathrm{co}}(\boldsymbol{\theta}))$ is given, with the knowledge that in practice only the column span of it is available from the measurement data.

1) The element-space data model: For deriving the sufficient conditions, let us define the phase of $\operatorname{diag}(\mathbf{g})\mathbf{A}(\boldsymbol{\theta})$ as

$$\rho_q(n) = \frac{1}{2\pi} \operatorname{angle}\left(g_q \left[\mathbf{A}(\boldsymbol{\theta})\right]_{qn}\right) = p_q \cos(\theta_n) + \phi_q, \quad (7)$$

for $q = 1, \ldots, Q$ and $n = 1, \ldots, N$. Introducing $\boldsymbol{\rho}_n = [\rho_1(n), \ldots, \rho_Q(n)]^T$ and defining $\mathbf{p}_{\mathrm{ext}} := \mathbf{p}$ for the APS array and $\mathbf{p}_{\mathrm{ext}} := \mathbf{1}_3 \otimes \mathbf{p}$ for the AVS array, we can write the above equation compactly as

$$\boldsymbol{\rho}_n = \mathbf{p}_{\mathrm{ext}} \cos(\theta_n) + \boldsymbol{\phi} = \begin{bmatrix} \mathbf{p}_{\mathrm{ext}} & \mathbf{I}_Q \end{bmatrix} \begin{bmatrix} \cos(\theta_n) \\ \boldsymbol{\phi} \end{bmatrix} \quad (8)$$

for $n = 1, 2, \ldots, N$. This is an under-determined system of Q equations, which has rank $Q - 1$ (with $\phi_1 = 0$), and Q unknowns. When $N = 1$, it is possible to solve (8), if another sensor/channel's phase error is known in the array (say w.l.o.g. $\phi_2 = 0$ in addition to $\phi_1 = 0$). However, when $N \geq 2$, we can eliminate $\boldsymbol{\phi}$ by considering

$$\boldsymbol{\rho}_n - \boldsymbol{\rho}_1 = \mathbf{p}_{\mathrm{ext}} \left[\cos(\theta_n) - \cos(\theta_1)\right],$$

to obtain $N-1$ linearly independent equations in N unknown DOAs of the form

$$\mathbf{p}_{\mathrm{ext}}^\dagger \ (\boldsymbol{\rho}_n - \boldsymbol{\rho}_1) = \cos(\theta_n) - \cos(\theta_1); \quad n = 2, \cdots, N. \quad (9)$$

The system in (9) is still underdetermined. Nonetheless, if one of the DOAs is known (say w.l.o.g. θ_1 is known) then we can identify the remaining DOAs. This result for a scalar sensor array ($Q = M$) was presented in [19].

However, for an AVS array ($Q = 3M$), the need of knowing the direction of one calibrator source θ_1 can be relaxed as the direction information is available in the magnitude of the element-space data model. This is a novel observation that is not presented in the existing literature. It can be seen by explicitly considering only the magnitude of $\operatorname{diag}(\mathbf{g})\mathbf{A}(\boldsymbol{\theta})$, resulting in

$$\nu_q(n) = \left| g_q \left[\mathbf{A}(\boldsymbol{\theta})\right]_{qn} \right| = \alpha_q \left| h_q(\theta_n) \right| \quad (10)$$

for $q = 1, \ldots, 3M$. Here,

$$h_q(\theta_n) = \begin{cases} 1, & 1 \leq q \leq M. \\ \cos(\theta_n), & M+1 \leq q \leq 2M. \\ \sin(\theta_n), & 2M+1 \leq q \leq 3M. \end{cases} \quad (11)$$

Let us consider the equations related to $q = M+1$, which are given by

$$\nu_{M+1}(n) = \alpha_{q+1} \cos(\theta_n).$$

If we assume $N \geq 2$, we can eliminate the unknown α_{q+1} to obtain

$$\cos(\theta_1) = \frac{\nu_{M+1}(1)}{\nu_{M+1}(n)} \cos(\theta_n).$$

Thus we can compute θ_1 as

$$\theta_1 = \arccos\left(\frac{\nu_{M+1}(n)}{\nu_{M+1}(n)} \cos(\theta_n)\right).$$

This value of θ_1 can be used in (9), which eliminates the need of knowing one of the DOAs for uniquely identifying all the N DOAs for the AVS linear array. The array manifold matrix $\mathbf{A}(\boldsymbol{\theta})$ is known once all the N DOAs are computed. Then using (7) and (10), respectively, the phase and gain errors can be computed.

Now to check if the derived sufficient condition for the APS linear array is also necessary, we need to show that the solution of $\mathbf{g}$ and $\boldsymbol{\theta}$ is not unique if we do not consider the calibrator source. To do so, assume an M-element APS array, and N far-field sources. For such configuration, due to the structure of $\mathbf{A}(\boldsymbol{\theta})$ and the nature of $\mathbf{g}$, we can have $\operatorname{diag}(\mathbf{g})\mathbf{A}(\boldsymbol{\theta}) = \operatorname{diag}(\mathbf{g} \odot \mathbf{a}(\theta_0))(\mathbf{A}(\boldsymbol{\theta}) \odot \mathbf{a}^*(\theta_0)) = \operatorname{diag}(\tilde{\mathbf{g}})\mathbf{A}(\tilde{\boldsymbol{\theta}})$, where generally $\mathbf{g} \neq \tilde{\mathbf{g}}$ and $\boldsymbol{\theta} \neq \tilde{\boldsymbol{\theta}}$ indicating the non-uniqueness of the solution.

In summary, with the element-space model for a linear APS array, irrespective of the array geometry, given $N \geq 2$ and $\operatorname{diag}(\mathbf{g})\mathbf{A}(\boldsymbol{\theta})$, the requirement of a calibrator source is a sufficient and necessary condition for a unique solution of $\mathbf{g}$ and $\boldsymbol{\theta}$ to exist. In contrast, a calibrator source is not needed for a linear AVS array.

2) The co-array data model: In contrast to the element-space formulation, there are many self-calibration approaches that are developed using the co-array model [13], [19], [20]. However, the conditions for the solution to exist are still not explored. So in this section, we derive the conditions using the co-array data model in (6) for uniquely estimating $\mathbf{g}$ and $\boldsymbol{\theta}$, given $\operatorname{diag}(\mathbf{g}^* \otimes \mathbf{g})\mathbf{A}_{\mathrm{co}}(\boldsymbol{\theta})$. To do so, consider the phase of $\operatorname{diag}(\mathbf{g}^* \otimes \mathbf{g})\mathbf{A}_{\mathrm{co}}(\boldsymbol{\theta})$ that is given by

$$\begin{aligned} \rho_{pq}(n) &= \frac{1}{2\pi} \operatorname{angle}\left(g_p^* g_q \left(\left[\mathbf{A}^*(\boldsymbol{\theta})\right]_{pn} \circ \left[\mathbf{A}(\boldsymbol{\theta})\right]_{qn}\right)\right) \\ &= (p_p - p_q) \cos(\theta_n) - (\phi_p - \phi_q), \end{aligned} \quad (12)$$

for $p, q = 1, \cdots, Q$ with $p \neq q$ and $n = 1, \cdots, N$.

If $N = 1$, we require two sensors/channels with known phase errors. Suppose w.l.o.g. that $\phi_1 = \phi_2 = 0$, then we can compute the DOA as

$$\theta_1 = \arccos\left(\frac{\rho_{12}(1)}{p_1 - p_2}\right), \quad (13)$$

---

[1] The covariance matrix $\mathbf{R}_\mathrm{x}$ is completely characterized by $N + 1$ real eigenvalues and $2QN - N^2 - N$ real parameters related to the orthonormal eigenvectors associated to the signal subspace.

with no specific requirements for $p_1 \neq 0$ or $p_2 \neq 0$. Defining $\boldsymbol{\rho}_n = [\rho_{11}(n), \rho_{12}(n), \ldots, \rho_{QQ}(n)]^T$, we can compactly write (12) as

$$\boldsymbol{\rho}_n = \mathbf{Dp}_{\text{ext}}\cos(\theta_n) - \mathbf{D}\boldsymbol{\phi} = \begin{bmatrix}\mathbf{Dp}_{\text{ext}} & -\mathbf{D}\end{bmatrix}\begin{bmatrix}\cos(\theta_n)\\ \boldsymbol{\phi}\end{bmatrix},\tag{14}$$

where $\mathbf{D} \in \mathbb{R}^{Q^2-Q}$ is the difference matrix that we use to compute the pairwise differences in (12). If $N \geq 2$, irrespective of the array geometry, the phase errors $\boldsymbol{\phi}$ can be eliminated by considering

$$\boldsymbol{\rho}_n - \boldsymbol{\rho}_1 = \mathbf{Dp}_{\text{ext}}\left[\cos(\theta_n) - \cos(\theta_1)\right]; \quad \forall n = 2, \ldots, N,$$

which can be equivalently expressed as

$$\theta_n = \arccos\left((\mathbf{Dp}_{\text{ext}})^\dagger\left[\boldsymbol{\rho}_n - \boldsymbol{\rho}_1\right] + \cos(\theta_1)\right).\tag{15}$$

This is similar to the element-space version as seen in (9) and it is underdetermined. Nonetheless for APS linear arrays, similar to the element-space model, if one of the source DOAs is known (say w.l.o.g. θ_1 is known) then we can identify the remaining DOAs.

For an AVS array, similar to the element-space model, the magnitude of $\text{diag}(\mathbf{g}^* \otimes \mathbf{g})\mathbf{A}_{\text{co}}(\boldsymbol{\theta})$ also contains the direction information. Specifically,

$$\begin{aligned}\nu_{pq}(n) &= \left|g_p g_q\left([\mathbf{A}^*(\boldsymbol{\theta})]_{pn} \circ [\mathbf{A}(\boldsymbol{\theta})]_{qn}\right)\right|,\\ &= \psi_p\psi_q h_p(\theta_n)h_q(\theta_n)\end{aligned}\tag{16}$$

for $p, q = 1, \cdots, 3M$ and $n = 1, \ldots, N$, where we recall that $h_p(\theta_n)$ is as in (11). Consider w.l.o.g. the equation related to $p = M+1$ and $q = M+2$, i.e.,

$$\nu_{M+1\,M+2}(n) = \psi_{M+1}\psi_{M+2}\cos^2(\theta_n).$$

When $N \geq 2$, we can eliminate the unknown gain errors ψ_{M+1} and ψ_{M+2} in the above equation by computing

$$\cos(\theta_1) = \left[\frac{\nu_{M+1\,M+2}(1)}{\nu_{M+1\,M+2}(n)}\cos^2(\theta_n)\right]^{1/2},$$

which can now be used in (15) to compute the DOAs. Once the DOAs are computed, the phase errors can be computed from (14), with respect to one of the reference sensors/channels in the array as the rank of $\mathbf{D}$ is always $Q-1$. The gain errors can be computed from the amplitude relations in (16).

Thus it can be concluded that irrespective of the array geometry of the linear array with the co-array data model, it is sufficient to have one phase reference sensor and one (no) calibrator source for a linear APS (respectively, AVS) array, for uniquely estimating $\mathbf{g}$ and $\boldsymbol{\theta}$ when $N \geq 2$ and $\text{diag}(\mathbf{g}^* \otimes \mathbf{g})\mathbf{A}_{\text{co}}(\boldsymbol{\theta})$ is given.

Unlike for the element-space approach, the derived sufficient condition for the APS linear array using the co-array model is not necessary. This aspect is showcased in the subsequent discussion, with certain assumptions on the array geometry, where we will see that using the co-array model the solution of $\mathbf{g}$ and $\boldsymbol{\theta}$ can be unique even if we do not consider the calibrator source.

3) Sparse APS array based on co-array data model: In comparison to (8), which is an under-determined system, it can be observed that (14) is a tall system with (Q^2-Q) equations and $(Q+1)$ unknowns. APS linear arrays with a particular structure in the array geometry, such as specific sparse arrays or uniform linear arrays (ULAs) result in redundant relations that are part of (14). Those redundancies in the structured APS linear array allow for estimating $\mathbf{g}$ and subsequently $\boldsymbol{\theta}$ without the knowledge of a known calibrator source leading to another set of sufficient conditions. This is discussed in the following part.

From the co-array perspective of scalar sensor arrays, the distinct elements of $\mathbf{Dp}_{\text{ext}}$, as seen in (14), behave like virtual sensor locations given by the difference set $\{p_i - p_j, 1 \leq i, j \leq M\}$. Those virtual sensor locations increase the degrees-of-freedom (DOF) of the array allowing for estimating more sources than physical sensors, if they are placed strategically. In order to look at the self-calibration problem for such array configurations, let us reuse some definitions from [11].

Definition 1. *(Difference co-array)* For an M-element sensor array, with p_i denoting the position of the ith sensor, define the set

$$\mathcal{D} = \{p_i - p_j\}, \ \forall i, j = 1, 2, \ldots, M,$$

which allows for a repetition of its elements. We also define the set $\mathcal{D}_\mathcal{U}$, which consists of the distinct elements of the set $\mathcal{D}$. Then, the difference co-array of the given array is defined as the array which has sensors located at positions given by the set $\mathcal{D}_\mathcal{U}$.

Definition 2. *(Weight function)* An integer valued weight function $w : \mathcal{D}_\mathcal{U} \rightarrow \mathbb{N}^+$ is defined as

$$w(p) = \text{no. of occurances of } p \text{ in } \mathcal{D}, p \in \mathcal{D}_\mathcal{U},$$

where $\mathbb{N}^+$ is the set of positive integers. The weight function $w(p)$ denotes the number of times p occurs in $\mathcal{D}$.

The cardinality of the set $\mathcal{D}_\mathcal{U}$ for a given array gives the degrees of freedom (DOF) that can be obtained from the difference co-array associated with that array. The motivation of sparse array design, such as the minimum redundancy array (MRA), sparse ruler array or nested array, is to maximize the number of DOF of the co-array for a fixed M, which in other words means the value of the weight function $w(p), \forall p \in \mathcal{D}_\mathcal{U} \setminus \{0\}$ has to be minimized. However, from the self-calibration perspective, a value of the weight function $w(p), \forall p \in \mathcal{D}_\mathcal{U} \setminus \{0\}$ greater than 1 is beneficial as this results in redundancies in (14). By exploiting redundancies in those relations for the nth source and each p, the directional terms can be eliminated resulting in an equation with only phase terms, i.e.,

$$\rho_{pq}(n) - \rho_{kl}(n) = \rho_{pqkl}(n) = \phi_p - \phi_q - \phi_k + \phi_l,\tag{17}$$

where $p_p - p_q = p_k - p_l$ for $p, q, k, l = 1, \ldots, M$ and $n = 1, \ldots, N$. Such relations for all p, q, k, l can be expressed as a system of equations, i.e.,

$$\begin{bmatrix}\ldots & \rho_{pqkl}(n) & \ldots\end{bmatrix}^T = \mathbf{T}\begin{bmatrix}\phi_1 & \ldots & \phi_M\end{bmatrix}^T,\tag{18}$$

This article has been accepted for publication in IEEE Transactions on Signal Processing. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TSP.2022.3214383

6

where $\mathbf{T}$ is a deterministic matrix, which depends on the chosen array geometry and the phase errors can be estimated by inverting it. The maximum amount of redundancies can be found in a uniform linear array (ULA), where for M elements, $w(\pm d) = M - d$, for $d = 0, 1, \ldots, M - 1$. The rank of $\mathbf{T}$ is then always $M - 2$, indicating that the phase errors can be estimated with respect to an arbitrary reference and within an arbitrary progressive phase factor [13], [17]. We now look into the rank of the $\mathbf{T}$ matrix for different structured sparse linear arrays and summarize how the phase errors can be estimated for each of those scenarios,

- To design an M-element sparse array, taking self calibration into consideration, there is a trade-off between DOF and redundancies. The maximum rank of $\mathbf{T}$ for an M-element APS array is upper bounded by $M - 2$. The rank of $\mathbf{T}$ for structured sparse arrays including the nested array [11] and super nested array [36] is always $M - 3$, whereas for the co-prime arrays [12], which enjoy more redundancies, it is $M - 2$.
- If there is a provision to introduce additional sensors within a sparse array to allow for sufficient redundancies, then the rank of $\mathbf{T}$ can be increased to $M - 2$. For example, for an MRA [10], with $M = 5$ and $\mathbf{p} = [0, 1, 4, 7, 9]^T$, the rank of $\mathbf{T}$ is 1. However, if we introduce two phase reference sensors with $\mathbf{p} = [0, 1, \mathbf{2}, \mathbf{3}, 4, 7, 9]^T$, then the rank of $\mathbf{T}$ is 5.

On the other hand, the gain errors can be estimated by considering the redundancies in the amplitude relations of $\mathrm{diag}(\mathbf{g}^* \otimes \mathbf{g})\mathbf{A}_{\mathrm{co}}(\boldsymbol{\theta})$. Irrespective of the array geometry, the rank of the equivalent $\mathbf{T}$ matrix obtained by considering $|\mathrm{diag}(\mathbf{g}^* \otimes \mathbf{g})\mathbf{A}_{\mathrm{co}}(\boldsymbol{\theta})|$ is always $M - 1$, indicating that the gain errors can be estimated with respect to the chosen reference sensor.

Using redundancies present in the co-array data model of an APS array with $N \geq 1$ and $\mathrm{diag}(\mathbf{g}^* \otimes \mathbf{g})\mathbf{A}_{\mathrm{co}}(\boldsymbol{\theta})$ is given, we can conclude that for a ULA, with two phase reference sensors in the array, while for sparse arrays, with at least two or more phase reference sensors in the array, it is also possible for uniquely estimating the calibration errors and source DOAs.

Remark 1 (On the identifiability analysis of non-linear arrays)**.** *The identifiability conditions for non-linear AVS arrays can be derived along similar lines of non-linear APS arrays as presented in [19]. It can be shown that for both APS and AVS arrays with $N \geq 2$, it is sufficient to have one reference sensor with a known gain and phase error for uniquely estimating both the calibration parameters and the source DOAs. In particular for APS non-linear arrays, the need for a reference source can be relaxed for the purpose of self-calibration due to the presence of extra degrees-of-freedom in its spatial frequencies.*

Remark 2 (On the choices to resolve the identifiability issues)**.** *To derive the sufficient conditions based on the redundancy-based calibration technique, we choose to have reference sensors with known phase errors in the array to improve the rank of the $\mathbf{G}$ matrix such that the phase errors can be*

estimated. However, we can also have other a priori conditions on the phase errors, such as $\sum_{q=1}^{Q} \phi_q = 0$, that improve the rank of $\mathbf{G}$, leading to another set of sufficient conditions to estimate the phase errors and subsequently the source DOAs uniquely.

IV. SELF CALIBRATION WITH THE ELEMENT-SPACE MODEL

In this section, we focus on estimating the complex-valued receiver gains and the source DOAs, when only a few snapshots are available. In such cases, the sample data covariance matrix will be a very poor estimate of $\mathbf{R}_x$ and hence we focus on the element-space data model. The algorithms provided in this section, do not make any assumptions on the array geometry or on the structure of the covariance matrix $\mathbf{R}_x$.

Assuming that the true directions are from a uniform grid of $D \gg N$ points, i.e., assuming that $\theta_n \in \left\{ 0, \frac{\pi}{D} \cdots, \frac{\pi(D-1)}{D} \right\}$, for $n = 1, 2, \ldots, N$, we can approximate (1) as

$$\mathbf{x}(t) = \mathrm{diag}(\mathbf{g})[\mathbf{A}_{\mathbb{D}}\,\mathbf{z}(t) + \mathbf{n}(t)], \quad (19)$$

where $\mathbf{A}_{\mathbb{D}}$ is a $Q \times D$ dictionary matrix with column vectors of the form $\mathbf{a}(\bar{\theta}_d)$, where $\bar{\theta}_d$ is the dth point of the uniform grid of directions, i.e., $\bar{\theta}_d = \frac{\pi d}{D}$, $d = 0, 1, \ldots, D - 1$, and $\mathbf{z}(t)$ is a length-Q vector containing the source signal related to the corresponding discretized directions. We emphasize here that finding the columns of $\mathbf{A}_{\mathbb{D}}$ that correspond to non-zero elements of $\mathbf{z}(t)$ amounts to finding the DOAs. As seen in (19), by assuming that the source DOAs lie on a pre-defined uniform grid, we transform a non-linear estimation problem into a bilinear estimation problem with $\mathbf{c}$ and $\mathbf{z}(t)$ being the unknowns (from which we can derive $\mathbf{g}$ and $\boldsymbol{\theta}$, respectively).

Defining the calibration matrix $\mathrm{diag}(\mathbf{c}) = \mathrm{diag}^{-1}(\mathbf{g})$, we can express the "calibrated" signal $\mathbf{y}(t)$ as

$$\mathbf{y}(t) = \mathrm{diag}(\mathbf{c})\mathbf{x}(t) = \mathrm{diag}(\mathbf{x}(t))\mathbf{c} = \mathbf{A}_{\mathbb{D}}\mathbf{z}(t) + \mathbf{n}(t). \quad (20)$$

Exploiting the nature of the calibration errors, which combined with simple algebraic manipulation, the bilinear estimation problem in (19) is further transformed into a linear estimation problem in (20).

Leveraging the fact that the calibration parameters remain unchanged during an observation window where we collect L snapshots, we can obtain more equations, i.e.,

$$\underbrace{\begin{bmatrix} \mathrm{diag}(\mathbf{x}(1)) & -\mathbf{A}_{\mathbb{D}} & & \\ \vdots & & \ddots & \\ \mathrm{diag}(\mathbf{x}(L)) & & & -\mathbf{A}_{\mathbb{D}} \end{bmatrix}}_{\mathbf{G}} \underbrace{\begin{bmatrix} \mathbf{c} \\ \mathbf{z} \end{bmatrix}}_{\boldsymbol{\gamma}} = \underbrace{\begin{bmatrix} \mathbf{n}(1) \\ \vdots \\ \mathbf{n}(L) \end{bmatrix}}_{\mathbf{n}},$$

(21)

where $\mathbf{z} = \mathrm{vec}(\mathbf{Z}) \in \mathbb{C}^{DL}$ with $\mathbf{Z} = [\mathbf{z}(1), \mathbf{z}(2), \cdots, \mathbf{z}(L)] = [\mathbf{z}_1, \mathbf{z}_2, \ldots, \mathbf{z}_D]^T$. Here, $\mathbf{z}(l) \in \mathbb{C}^D$ and $\mathbf{z}_d \in \mathbb{C}^L$.

Although at the outset, it seems as if there are $Q + DL$ unknowns in (21), the vector $\mathbf{z}$ is structured. Specifically, the vectors $\mathbf{z}(l)$, $l = 1, \ldots, L$ are sparse, and more importantly, they have the *same* sparsity pattern with the indices of the nonzero pattern indicating the source directions. The prior knowledge of having sparsity along the spatial domain can be incorporated by initially considering the l_2 norm of all

the time samples corresponding to a particular spatial index of $\mathbf{Z}$, i.e., by defining $z_d^{(\ell_2)} = \|\mathbf{z}_d\|_2$ for $d = 1, 2, \ldots, D$, and then by using the sparsity promoting l_1 norm penalty on the vector $\mathbf{z}^{(\ell_2)} = \left[z_1^{(\ell_2)}, z_2^{(\ell_2)}, \ldots, z_D^{(\ell_2)}\right]^T$ as $f(\mathbf{z}) = \|\mathbf{z}^{(l_2)}\|_{\ell_1} = \sum_{d=1}^{D} z_d^{(\ell_2)}$.

The optimization problem to jointly estimate the calibration parameters and DOAs with a sparsity constraint along the spatial domain of the matrix $\mathbf{Z}$ can then be expressed as:

$$\min_{\mathbf{c}, \mathbf{z}} \quad \|\mathbf{G}\boldsymbol{\gamma}\|_2^2 + \eta\, f(\mathbf{z}) \quad \text{s.t.} \quad (\mathbf{c}, \mathbf{z}) \in \mathcal{C} \qquad (22)$$

where $\boldsymbol{\gamma} = [\mathbf{c}^T \; \mathbf{z}^T]^T$ and η is the regularization parameter that allows for a trade off between the goodness of fit of the solution to the given data and the sparsity prior on $\mathbf{z}$. The constraint set for APS linear arrays is $\mathcal{C} := \{(\mathbf{c}, \mathbf{z}) \mid c_1 = 1, \mathbf{z}_1 = 1\}$ while for AVS linear arrays it is $\mathcal{C} := \{(\mathbf{c}, \mathbf{z}) \mid c_1 = 1\}$. Recall that for APS arrays, we need one reference sensor and we need to know one of the DOAs to avoid ambiguities. This is done by setting $c_1 = 1$ and $\mathbf{z}_1 = 1$, which is equivalent to having a calibrator source at $\bar{\theta}_1$ (w.l.o.g.). Since for AVS arrays, we do not need any calibrator source, we only need a reference sensor in that case. Furthermore, the constraint set for both the APS and AVS non-linear arrays is $\mathcal{C} = \{(\mathbf{c}, \mathbf{z}) \mid c_1 = 1\}$. Recall that for non-linear arrays, we only need a reference sensor to avoid ambiguities. The optimization problem (22) is a convex optimization problem, which can be solved using any off-the-shelf solver. For large L, if the number of sources can be estimated, the complexity of the formulation in (22) can be reduced by using the ℓ_1-SVD technique [8] on the measurement data matrix $\mathbf{X}$. Furthermore, for the choice of the regularization parameter η, we follow the discrepancy principle discussed in [8].

Remark 3. *In contrast to the considered measurement model, the calibration errors affect only the signal component of the data, when the errors originate due to the perturbation of the sensors, gain and phase patterns or due to the position or orientation errors of the sensors in the array [34], [37]. In such case,* (1) *and* (20) *can be modified, respectively as,*

$$\mathbf{x}(t) = \operatorname{diag}(\mathbf{g})\mathbf{A}(\boldsymbol{\theta})\,\mathbf{s}(t) + \mathbf{n}(t).$$

$$\mathbf{y}(t) = \operatorname{diag}(\mathbf{c})\mathbf{x}(t) = \operatorname{diag}(\mathbf{x}(t))\mathbf{c} = \mathbf{A}_{\mathbb{D}}\mathbf{z}(t) + \operatorname{diag}(\mathbf{c})\mathbf{n}(t).$$

The proposed calibration approach in (22) *is still applicable here with the additive noise term being modified as* $\hat{\mathbf{n}}(t) = \operatorname{diag}(\mathbf{c})\mathbf{n}(t)$.

V. SELF CALIBRATION WITH THE CO-ARRAY DATA MODEL

In this approach both the calibration errors and the source DOAs will be estimated jointly based on the covariance matrix of the measurement data. Similar to (19), the directions are assumed to be on a uniform grid of $D \gg N$ points. Then (6) can be approximated as

$$\mathbf{r}_{\mathrm{x}} = \operatorname{diag}(\mathbf{g}^* \otimes \mathbf{g})[\mathbf{A}_{\mathrm{co}\mathbb{D}}\boldsymbol{\sigma}_{\mathrm{z}} + \boldsymbol{\sigma}_{\mathrm{n}}], \qquad (23)$$

where $\mathbf{A}_{\mathrm{co}\mathbb{D}}$ is a $Q^2 \times D$ dictionary matrix that consists of column vectors of the form $\mathbf{a}^*(\bar{\theta}_d) \otimes \mathbf{a}(\bar{\theta}_d)$, with $\bar{\theta}_d$ as defined

before. Again similar to (20), defining the calibration matrix $\operatorname{diag}(\mathbf{c}^* \otimes \mathbf{c}) = \operatorname{diag}^{-1}(\mathbf{g}^* \otimes \mathbf{g})$, we can express (23) as

$$\operatorname{diag}(\mathbf{c}^* \otimes \mathbf{c})\mathbf{r}_{\mathrm{x}} = \operatorname{diag}(\mathbf{r}_{\mathrm{x}})(\mathbf{c}^* \otimes \mathbf{c}) = \mathbf{A}_{\mathrm{co}}(\boldsymbol{\theta})\boldsymbol{\sigma}_{\mathrm{s}} + \boldsymbol{\sigma}_{\mathrm{n}}. \quad (24)$$

Since $(\mathbf{c}^* \otimes \mathbf{c}) = \operatorname{vec}(\mathbf{C})$, with $\mathbf{C} = \mathbf{c}\mathbf{c}^H$, (23) can be compactly rewritten as

$$\underbrace{\begin{bmatrix} \operatorname{diag}(\mathbf{r}_{\mathrm{x}}) & -\mathbf{A}_{\mathrm{co}\mathbb{D}} \end{bmatrix}}_{\mathbf{G}_{\mathrm{co}}} \underbrace{\begin{bmatrix} \operatorname{vec}(\mathbf{C}) \\ \boldsymbol{\sigma}_{\mathrm{z}} \end{bmatrix}}_{\boldsymbol{\gamma}_{\mathrm{co}}} = \boldsymbol{\sigma}_{\mathrm{n}}. \qquad (25)$$

Similar to the element-space formulation, we have transformed the non-linear estimation problem in (6) to a linear estimation problem in (25). The above system is underdetermined with $Q^2 + D$ unknowns in Q^2 equations (note that some equations might even be redundant). However, as $\operatorname{vec}(\mathbf{C})$ has a Kronecker structure, the actual number of unknowns reduces to Q and $\boldsymbol{\sigma}_{\mathrm{z}}$ is a sparse vector with non-zero elements at the location of the source DOAs. By considering the aforementioned constraints on the calibration errors and source DOAs, the estimation problem can be cast as

$$\min_{\mathbf{C}, \boldsymbol{\sigma}_{\mathrm{z}}} \quad \|\mathbf{G}_{\mathrm{co}}\boldsymbol{\gamma}_{\mathrm{co}} - \boldsymbol{\sigma}_{\mathrm{n}}\|_2^2 + \eta\, \|\boldsymbol{\sigma}_{\mathrm{z}}\|_0 \quad \text{s.t.} \quad (\mathbf{C}, \boldsymbol{\sigma}_{\mathrm{z}}) \in \mathcal{C}_{\mathrm{co}} \tag{26}$$

where $\boldsymbol{\gamma}_{\mathrm{co}} = [\operatorname{vec}^T(\mathbf{C}), \boldsymbol{\sigma}_{\mathrm{z}}^T]^T$, η is the regularization parameter, for $N \geq 2$ the constraint set $\mathcal{C}_{\mathrm{co}} = \{(\mathbf{C}, \boldsymbol{\sigma}_{\mathrm{z}}) \mid \boldsymbol{\sigma}_{\mathrm{z}} \succeq \mathbf{0}, \mathbf{C} = \mathbf{c}\mathbf{c}^H, c_1 = 1, \boldsymbol{\sigma}_{\mathrm{z}}(1) = 1\}$ for APS linear arrays and $\mathcal{C}_{\mathrm{co}} = \{(\mathbf{C}, \boldsymbol{\sigma}_{\mathrm{z}}) \mid \boldsymbol{\sigma}_{\mathrm{z}} \succeq \mathbf{0}, \mathbf{C} = \mathbf{c}\mathbf{c}^H, c_1 = 1\}$ for AVS linear arrays. For APS linear arrays, the requirement of knowing one of the DOAs is expressed as $\boldsymbol{\sigma}_{\mathrm{z}}(1) = 1$ (w.l.o.g.). Further for APS ULAs and some APS sparse linear arrays, the redundancies in the co-array measurements can be used for the estimation of the source DOAs and the calibration errors with two phase reference sensors in the array. In such cases the constraint set even with $N \geq 1$ is $\mathcal{C}_{\mathrm{co}} = \{(\mathbf{C}, \boldsymbol{\sigma}_{\mathrm{z}}) \mid \boldsymbol{\sigma}_{\mathrm{z}} \succeq \mathbf{0}, \mathbf{C} = \mathbf{c}\mathbf{c}^H, c_1 = c_2 = 1\}$. Furthermore, the constraint set for both the APS and AVS non-linear arrays with $N \geq 2$ is $\mathcal{C}_{\mathrm{co}} = \{(\mathbf{C}, \boldsymbol{\sigma}_{\mathrm{z}}) \mid \boldsymbol{\sigma}_{\mathrm{z}} \succeq \mathbf{0}, \mathbf{C} = \mathbf{c}\mathbf{c}^H, c_1 = 1\}$. Again we can recall that for non-linear arrays, we only need a reference sensor to avoid ambiguities. The optimization problem in (26) is non-convex due to the l_0 norm (cardinality) constraint and the rank-one equality constraint on $\mathbf{C}$. We can relax (26) by replacing the cardinality constraint with its convex approximation $\|\boldsymbol{\sigma}_{\mathrm{z}}\|_1$ and by replacing the rank-one equality constraint (i.e., $\mathbf{C} = \mathbf{c}\mathbf{c}^H$) in the set $\mathcal{C}_{\mathrm{co}}$ with a convex inequality constraint (i.e., $\mathbf{C} \succeq \mathbf{c}\mathbf{c}^H$). The new set which is same as $\mathcal{C}_{\mathrm{co}}$ except for the rank-one convex inequality constraint is denoted as $\tilde{\mathcal{C}}_{\mathrm{co}}$. The relaxed optimization problem can be expressed as,

$$\min_{\mathbf{C}, \boldsymbol{\sigma}_{\mathrm{z}}} \quad \|\mathbf{G}_{\mathrm{co}}\boldsymbol{\gamma}_{\mathrm{co}} - \boldsymbol{\sigma}_{\mathrm{n}}\|_2^2 + \eta\|\boldsymbol{\sigma}_{\mathrm{z}}\|_1 \quad \text{s.t.} \quad (\mathbf{C}, \boldsymbol{\sigma}_{\mathrm{z}}) \in \tilde{\mathcal{C}}_{\mathrm{co}}. \tag{27}$$

The convex inequality constraint, $\mathbf{C} \succeq \mathbf{c}\mathbf{c}^H$, is equivalent to $\begin{bmatrix} \mathbf{C} & \mathbf{c} \\ \mathbf{c}^H & 1 \end{bmatrix} \succeq 0$ from Schur's lemma. The resulting problem is a semi-definite programming problem that can be solved with any off-the-shelf solver. For the choice of the regularization parameter η, we can use any standard method adopted in sparse signal recovery [38]. In practice, for the finite snapshot

scenario, $\mathbf{C}$ obtained after solving (27) might not be rank one and the closest estimates of the calibration parameters can be obtained from the first dominant singular vector of $\mathbf{C}$. The formulation in (27) is also applicable to sparse arrays for estimating DOAs (when there are more sources than sensors) and calibration parameters jointly as presented in [35].

Remark 4. *If calibration errors affect only the signal component of the data, then* (6) *can be modified as*

$$\mathbf{r}_{\mathrm{x}} = \mathrm{diag}(\mathbf{g}^* \otimes \mathbf{g})\mathbf{A}_{\mathrm{co}}(\boldsymbol{\theta})\boldsymbol{\sigma}_{\mathrm{s}} + \boldsymbol{\sigma}_{\mathrm{n}}.$$

The proposed calibration approach in (27) *is still applicable here with a slight modification. More specifically,* (25) *and* (26) *can then be modified, respectively as,*

$$\underbrace{\begin{bmatrix} \mathrm{diag}(\mathbf{r}_{\mathrm{x}} - \boldsymbol{\sigma}_{\mathrm{n}}) & -\mathbf{A}_{\mathrm{co}\mathbb{D}} \end{bmatrix}}_{\mathbf{G}_{\mathrm{co}}} \underbrace{\begin{bmatrix} \mathrm{vec}(\mathbf{C}) \\ \boldsymbol{\sigma}_{\mathrm{z}} \end{bmatrix}}_{\boldsymbol{\gamma}_{\mathrm{co}}} = 0,$$

$$\min_{\mathbf{C}, \boldsymbol{\sigma}_{\mathrm{z}}} \quad \|\mathbf{G}_{\mathrm{co}}\boldsymbol{\gamma}_{\mathrm{co}}\|_2^2 + \eta\|\boldsymbol{\sigma}_{\mathrm{z}}\|_0 \quad \mathrm{s.\,t.} \quad (\mathbf{C}, \boldsymbol{\sigma}_{\mathrm{z}}) \in \mathcal{C}_{\mathrm{co}}.$$

VI. Numerical experiments

In this section, we present numerical simulations to illustrate the performance of all the proposed solvers for the joint estimation of the source DOAs and calibration parameters. Firstly we consider the element-space model based solver in (22) only for AVS linear arrays. Recall that AVS arrays do not require the presence of a reference source (see Section III). Then the covariance model in (27) is considered for both the APS and AVS linear array. Finally, we analyze the root mean square error (RMSE) of the DOA estimates obtained from the presented algorithms and compare them with existing calibration methods. The RMSE results for scenarios with more sensors than sources are also compared with the Cramér-Rao lower bound (CRLB) on the DOA estimates.
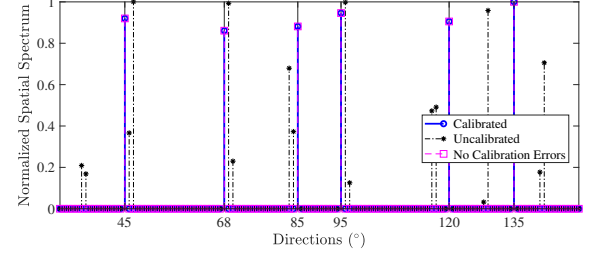
A. Element-space model

We consider a scenario with $M = 8$ AVSs arranged in a uniform linear array (ULA) configuration where the spacing between the consecutive sensors is half a wavelength of the considered narrowband source signals. Further, we consider a scenario with $N = 6$ narrowband far-field signals impinging on the array from distinct DOAs with an observation period consisting of $L = 50$ snapshots. The grid is chosen to be uniform between $[0° \ 180°]$ with $1°$ resolution. Without loss of generality, we assume the first channel of the first AVS in the array as the reference channel whose gain is 1 and phase is $0°$. The gain and phase errors are chosen from a uniform distribution over the interval [-3; 3] dB and $[-20°; 20°]$, respectively.

Based on the optimization problem in (22), the results of DOA estimation post calibration are presented in Fig. 1. In order to verify the correctness of the formulation in (22), we initially considered an ideal scenario without measurement noise. The DOA spectra based on (22) are presented in Fig. 1(a). It is seen in Fig. 1(a), that we recover the
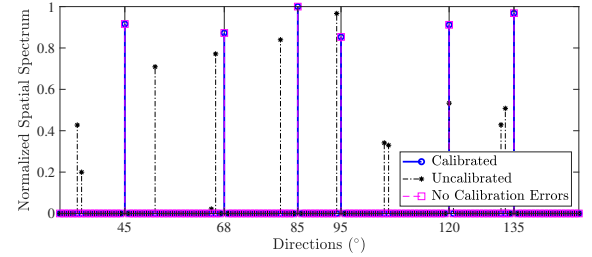
(a) l_1-SVD spectra without measurement noise.

(b) l_1-SVD spectra with SNR = 10 dB.
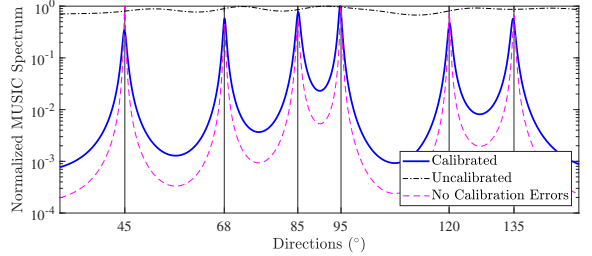
(c) MUSIC spectra with an SNR = 10 dB.

Fig. 1. The l_1-SVD and MUSIC spectra using the element-space data model based solver in (22) for an AVS ULA with $M = 8$, $N = 6$ and $L = 50$. The true DOAs are indicated by the black solid lines.

exact source DOAs after solving (22), where as for the uncalibrated data, the source DOA estimates based on the l_1-SVD algorithm [8] are very poor. Further, we considered the measurement data with a signal-to-noise ratio (SNR) of 10 dB and the corresponding DOA spectra obtained from solving (22) are presented in Fig. 1(b), where we draw a similar inference as in Fig. 1(a).

On the other hand, the issues of a pre-defined grid on the DOA estimates obtained after solving (22) can be minimized by applying the MUSIC algorithm on the gain and phase compensated covariance matrix. The gain and phase errors are estimated from (22), and the corresponding MUSIC spectra are presented in Fig. 1(c). It can be inferred that for the measurement data with an SNR of 10 dB, MUSIC with the uncalibrated data results in poor estimates, whereas the DOA estimates after calibration in Fig. 1(c) provides similar results as in Fig. 1(b). The two-step procedure to obtain DOA estimates from MUSIC spectrum significantly improves results when the sources do not lie on a pre-defined grid.

B. Co-array data model

To illustrate the effectiveness of the covariance domain formulation provided in (27), we consider both a conventional ULA with less sources than sensors and a sparse linear array

with more sources than sensors, where the smallest spacing between the consecutive sensors is half a wavelength of the considered narrowband source signals. Here, all the far-field source DOAs are chosen to be on the grid. In both the scenarios, without loss of generality, for the APS arrays we considered the first two sensors as references whereas for the AVS arrays the first channel is considered as a reference with gain of 1 and phase of $0°$.

1) ULA with less sources than sensors: Consider a ULA with $M = 8$, $N = 4$ far-field sources and SNR = 10 dB. Firstly, we consider a finite sample scenario with the observation period consisting of $L = 1000$ snapshots whose l_1 norm based DOA spectra upon solving (27) are plotted in Fig. 2(a) for the APS ULA and in Fig. 2(b) for the AVS ULA. The uncalibrated data in all the plots results in low resolution DOA spectra and very poor DOA estimates. In Fig. 2(a), the DOA spectra upon solving (27) show an improvement compared to the DOA spectra computed with the uncalibrated data. However, the resulting DOA spectra still have low resolution, as the model considered in (27) is not exact due to the finite sample approximation of the covariance matrix estimation. On the other hand, in Fig. 2(b), the DOA spectra based on (27) are significantly superior with high resolution compared to the DOA spectra computed with the uncalibrated data. However upon closer observation, we can notice that the DOA estimates are slightly biased for a couple of sources and also there are some spurious peaks in the DOA spectra. It is observed that the model mismatches due to the finite sample approximation of the covariance matrix estimation, has higher impact on reducing the sparsity of the DOA spectra for the APS ULA in comparison to an equivalent AVS ULA.
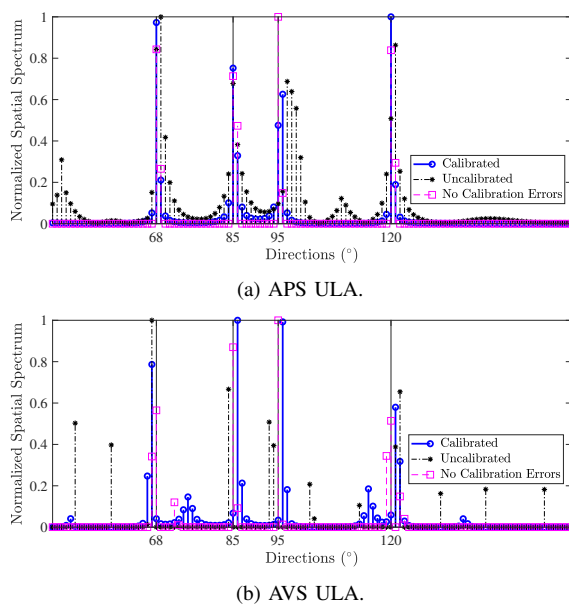

(a) APS ULA.

(b) AVS ULA.

Fig. 2. The l_1 norm based DOA spectra for both the APS and AVS ULA based on the co-array data model solver in (27) with $L = 1000$, SNR = 10 dB, $M = 8$ and $N = 4$ far-field sources. The true DOAs are indicated by the black solid lines.

In order to overcome the discussed issues with DOA estimates and the effects of a predefined grid, similar to the

element-space approach, a grid-free approach such as MUSIC algorithm can be applied on the measurement data in (5), which is compensated for the gain and phase errors obtained from (27). Those MUSIC spectra based on the calibrated data are presented in Fig. 3. The results in Fig. 3(b) for the AVS ULA is compared with [19] (referred to as Weiss-Friedlander approach). The results in Fig. 3(a) for the APS ULA is compared with [13] (referred to as the Paulraj-Kailath approach[2]), as the Weiss-Friedlander approach is not effective for linear scalar sensor arrays.

In Figs. 3(a) and (b), we see that the MUSIC spectra have a higher resolution and improved estimates compared to the equivalent l_1 norm based DOA spectra. On contrary, the spectra based on the uncalibrated data is not able to resolve all the sources and the resolution of the spectra is also degraded. Further, for the APS ULA in Fig. 3(a), the proposed approach outperforms [13], and for the AVS ULA in Fig. 3(b), it can be observed that although [19] results in a sharper peaks compared to the proposed approach, the estimates are highly biased.
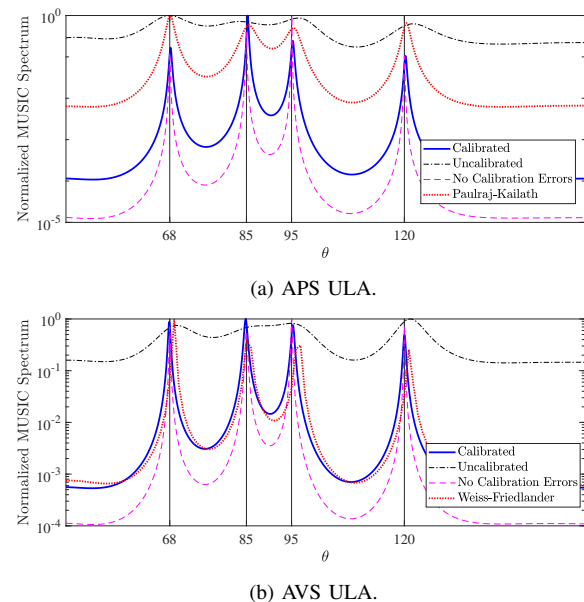

(a) APS ULA.

(b) AVS ULA.

Fig. 3. The MUSIC spectra for both the APS and AVS ULA based on the co-array data model solver in (27) with $L = 1000$, SNR = 10 dB, $M = 8$ and $N = 4$ far-field sources. The true DOAs are indicated by the black solid lines.

It can be summarized that based on the formulation in (27), it is possible to jointly estimate both the calibration errors as well as the source DOAs and the estimation results are good when the number of time snapshots are higher and the grid-mismatches are minimal. However, when the number of time snapshots are limited and we have a pre-defined grid, solving (27) can be used as a pre-conditioning step to estimate the calibration errors. Then a grid-free approach such as MUSIC can be applied on the gain and phase errors compensated measurement data to obtain improved and reliable DOA estimates.

---

[2]During the submission of this manuscript it came to the authors' attention that an improved version of [13] for scalar sensor arrays that considers an optimally-weighted least squares (OWLS) approach was proposed in [27].

2) Sparse array with more sources than sensors: Consider a hole-free sparse linear array with $M = 6$, $\mathbf{p} = [0\,1\,2\,3\,6\,9]^T$, $N = 8$ far-field sources and SNR = 10 dB. The rank of the $\mathbf{T}$ matrix [cf. (18)] for the considered sparse array is 4 (i.e., $M - 2$). For this scenario, we present spatial smoothing MUSIC (SS MUSIC) spectra [11] based on the gain and phase compensated measurement data, where the calibration errors are estimated by evaluating the proposed formulation in (27). We consider a finite sample scenario with the observation period consisting of $L = 500$ snapshots whose SS MUSIC spectra are shown in Fig. 4(a) for the APS array and Fig. 4(b) for the AVS array. The results of SS MUSIC for both the APS and AVS array are compared with the sparse total least squares (STLS) calibration approach [29].

(a) APS sparse array.
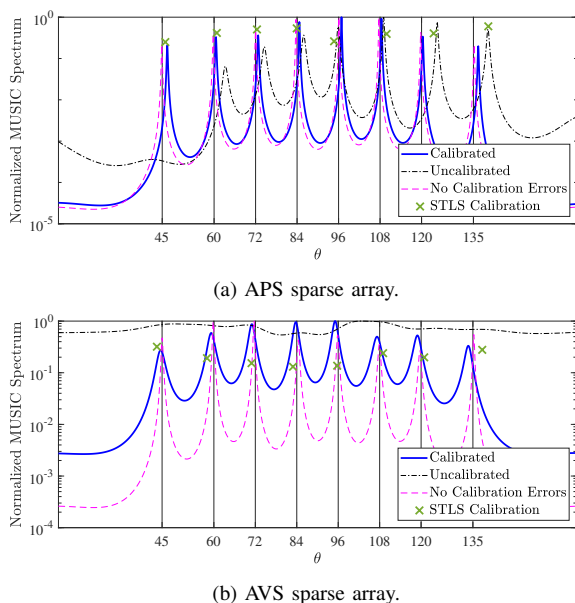

(b) AVS sparse array.

Fig. 4. The spatial smoothing MUSIC (SS MUSIC) spectra for both the APS and AVS sparse linear array based on the co-array data model solver in (27) with $L = 500$, SNR = 10 dB, $M = 6$, $\mathbf{p} = [0\,1\,2\,3\,6\,9]^T$ and $N = 8$ far-field sources. The true DOAs are indicated by the black solid lines.

In both Figs. 4(a) and (b), we see that post calibration, the SS MUSIC spectra have a higher resolution and are comparable to the scenario with no calibration errors, whereas the spectra based on the uncalibrated data are not able to resolve all the sources and the resolution of the spectra is also degraded. Furthermore, for both the APS and AVS sparse array with 500 snapshots, the performance of our proposed method is better than the STLS calibration approach [29].

The simulation setup for the AVS sparse linear array considered in Fig. 4(b), consists of less sources ($N = 8$) than the number of channels of the AVS array, ($3M = 18$). The proposed calibration approach in (27) is still applicable to an AVS sparse linear array with more sources than channels. However, because of the aperture limitation, when many sources are closely spaced it will be hard to discriminate them. To solve this issue, we can further boost the aperture by spatially undersampling the AVS array as in [39]. Such a setup is considered in Fig. 5, where the aperture is doubled and the smallest spacing between consecutive sensors is unit wavelength (instead of half a wavelength) of the considered

narrowband source signals. The SS-MUSIC spectra for an ideal scenario with $M = 6$, $N = 19$ ($> 3M$), SNR = 10 dB and $L = \infty$ are shown. Similar inferences as from Fig. 4(b),
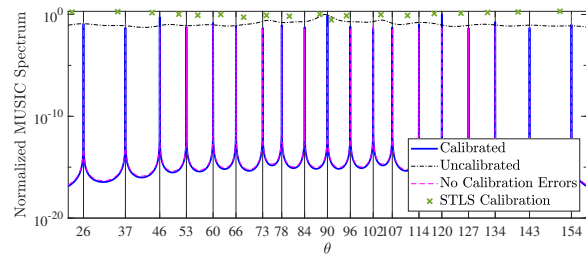

Fig. 5. The spatial smoothing MUSIC (SS MUSIC) spectra for the AVS sparse linear array based on the co-array data model solver in (27) with $M = 6$, $\mathbf{p} = [0\,1\,2\,3\,6\,9]^T$ (with smallest inter-sensor spacing equals to λ of the considered narrowband source signals), SNR = 10 dB, $L = \infty$ and $N = 19$ far-field sources. The true DOAs are indicated by the black solid lines.

can be made in Fig. 5, which showcases the applicability of the proposed calibration approach in (27) with a spatially undersampled AVS array with more sources than channels.

C. Monte-Carlo experiments

In this section we study the statistical behavior through the root mean square error (RMSE) of the DOA estimator based on the proposed calibration procedure for different scenarios. We consider both AVSs and APSs arranged in a ULA and sparse linear array configurations.

1) Uniform linear array with less sources than sensors: Firstly, we consider $M = 8$ sensors arranged in a ULA configuration and three far-field sources, i.e., $N = 3$ with $\boldsymbol{\theta} = [78°, 90°, 102°]$. The gain and phase perturbations follow a uniform distribution over the interval of [-2, 2] dB and $[-40°, 40°]$, respectively. For both the element-space formulation (22) and covariance domain formulation (27), we have chosen the pre-defined grid between $0°$ and $180°$ with $1°$ resolution. The RMSE of the DOA estimates based on the l_1 norm spectra (either by solving (22) or (27)) as well as the MUSIC spectra are presented for the considered scenarios.

Fixed SNR and varying snapshots: The RMSE of the DOA estimates for the source present at $90°$ based on 500 Monte-Carlo trials for both the APS and AVS ULA are presented in Fig. 6. Here the calibration errors and SNR of 10 dB were fixed for all the trials while the number of snapshots are varying. The RMSE of the DOA estimates in Fig. 6 based on the l_1 norm spectra by solving (22) is referred to as "Calibrated - Element Space" and by solving (27) is referred to as "Calibrated - Coarray". Further, the RMSE in the DOA estimates in Fig. 6 based on the MUSIC spectra by solving (22) is referred to as "Calibrated - Element Space - MUSIC" and by solving (27) is referred to as "Calibrated - Coarray - MUSIC".

In Fig. 6(a), we considered the AVS ULA with an SNR of 10 dB. It is seen that as the number of snapshots increases, the RMSE of the DOA estimates for the uncalibrated case does not decrease, whereas after calibration based on both the l_1 norm spectra and the MUSIC spectra, the results
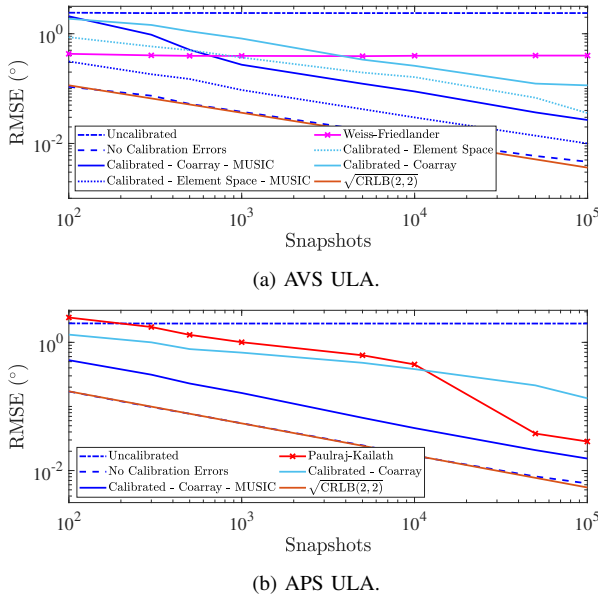
(a) AVS ULA.

(b) APS ULA.

Fig. 6. RMSE variation of the DOA estimates for the source at $90°$ using both the APS and AVS ULA with $M = 8$, $N = 3$ and $\boldsymbol{\theta} = [78°, 90°, 102°]$ for a fixed SNR of 10 dB as the number of snapshots are varying.

approach the ideal scenario with no calibration errors and its CRLB. For a given number of snapshots, MUSIC based DOA estimates result in lower RMSE values when compared with the equivalent l_1 norm based DOA estimates, further emphasizing the fact that the calibration estimates are robust to the model mismatches while solving either (22) or (27). On the other hand, the RMSE of the DOA estimates based on the Weiss-Friedlander approach [19] is also presented in Fig. 6, where the calibration parameters were initialized with a gain of 1 and a phase of $0°$. It is seen that the RMSE of the DOA estimates decreases initially, however it tends to saturate as the number of snapshots increases as it leads to a sub-optimal solution depending on the initialization. Also it can be observed that the DOA estimates based on the MUSIC spectra with calibration parameters estimated from (27) require more snapshots to obtain better DOA estimates with low RMSE as the finite sample errors in the estimation of the covariance matrix are high for a low number of snapshots and those are not modeled in the formulation of (27). Furthermore, based on the MUSIC spectra in Fig. 6(a), it can be observed that the performance of the element-space approach is far superior than the covariance domain approach.

Similarly in Fig. 6(b), we considered the APS ULA with an SNR of 10 dB. For the APS ULA, only formulation in (27) is considered and the results of the proposed methodology are compared with the Paulraj-Kailath approach [13]. The RMSE of the DOA estimates of the proposed methodology follows same trend as seen for the AVS ULA in Fig. 6(a). On the other hand, although the calibration approach in [13] achieves the optimal solution, it requires more snapshots to achieve similar performance as the proposed methodology.

Fixed number of snapshots and varying SNR: The variation of the RMSE in the DOA estimates with respect to a change in SNR for a fixed number of snapshots is considered

in Fig. 7. The same setup as in Fig. 6 is considered with $N = 3$ ($\boldsymbol{\theta} = [78°, 90°, 102°]$) where the RMSE of the source at $90°$ is presented. In Figs. 7(a) and (b), we consider the
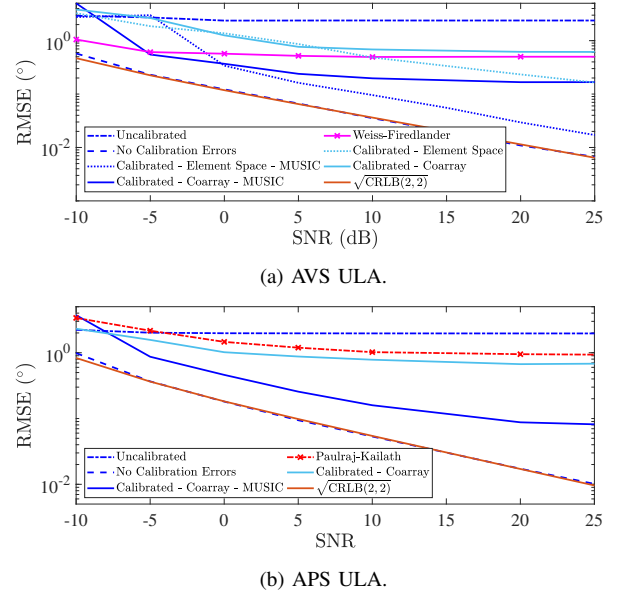

(a) AVS ULA.

(b) APS ULA.

Fig. 7. RMSE variation of the DOA estimates for the source at $90°$ using both the APS and AVS ULA with $M = 8$, $N = 3$ and $\boldsymbol{\theta} = [78°, 90°, 102°]$ as the SNR varies for a fixed number of snapshots of 1000.

AVS and the APS ULA, respectively, with 1000 snapshots and varying SNR. Similar to Fig. 6, it is seen that after calibration using the formulation in (22) as well as in (27) the RMSE of the DOA estimates decreases as the SNR increases for both the l_1 based spectra and the MUSIC spectra. Also as expected we can observe that the MUSIC spectra based DOA estimates outperform the l_1 based DOA estimates for a given SNR. Further, it can be inferred that the RMSE of the DOA estimates based on the proposed element-space model calibration technique asymptotically approaches the ideal scenario with no calibration errors and its CRLB. On the other hand, we can observe that the RMSE in the DOA estimates using the Weiss-Friedlander approach in Fig. 7(a) for the AVS ULA and the Paulraj-Kailath approach in Fig. 7(b) for the APS ULA, initially decreases as the SNR increases. However for an SNR greater than 5 dB the RMSE of the DOA estimates tends to saturate due to the finite sample errors in the covariance matrix estimation.

Gain and Phase RMSE estimates: Finally, the RMSE in the gain and phase error estimates for the setup considered in Figs. 6 and 7 is considered. In Fig. 8 the norm of the difference between the estimates and the actual values of the gain and phase errors is presented. In Figs. 8(a) and 8(b), RMSE related to the phase and gain error estimates with varying snapshots is considered, with the SNR being 10 dB. It can be observed in Figs. 8(a) and 8(c) that the RMSE related to both the phase and gain errors tends to approach zero as the number of snapshots increases except for the Weiss-Friedlander approach [19] as it produces a sub-optimal solution. This trend is consistent for the proposed calibration approach based on both the element-space and co-array formulation. Further, in Figs. 8(a) and 8(b)

(a) RMSE of phase errors.
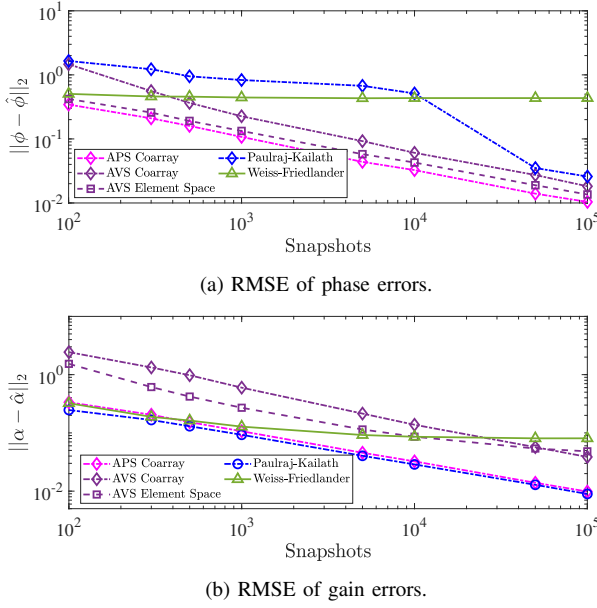

(b) RMSE of gain errors.

Fig. 8. RMSE variation of the gain and phase error estimates for both the APS and AVS ULA setup considered in Figs. 6 and 7 as the snapshots varies for a fixed SNR of 10 dB.

(a) APS sparse array.

(b) AVS sparse array.

Fig. 9. RMSE variation in the DOA estimates for the source at $90°$ based on the SS-MUSIC spectra using both APS and AVS sparse arrays with $M = 6$, $\mathbf{p} = [0\,1\,2\,3\,6\,9]^T$, $N = 2$, SNR = 10 dB and $\boldsymbol{\theta} = [70°, 90°]$ as the number of snapshots are varying. Here the scenario with "No Calibration Errors" is considered as the baseline reference to compare the performance of the proposed self-calibration solver.

it can be observed that the RMSE related to the phase errors based on the proposed calibration approach outperforms the Paulraj-Kailath approach.

2) Sparse linear array with more sources than sensors: In Fig. 9, the RMSE of the DOA estimates for APS and AVS sparse linear array based on the SS MUSIC spectra obtained using gain and phase compensated measurement data for different SNRs and for different numbers of data snapshots is presented. Here, we use $M = 6$, $\mathbf{p} = [0\,1\,2\,3\,6\,9]^T$ and $N = 2$ with $\boldsymbol{\theta} = [70°, 90°]$. The RMSE is computed for the source at $90°$ using 500 independent Monte-Carlo trials, but with fixed gain and phase errors. In Figs. 9(a) and (b), we can observe that as the number of snapshots increases, the RMSE of the DOA estimate after calibration approaches the ideal scenario without any sensor errors. Furthermore, the RMSE for the STLS calibration saturates both when increasing the number of snapshots, as it converges to a sub-optimal solution.

VII. EXPERIMENTAL RESULTS

An experimental study was conducted in order to demonstrate the proposed joint DOA and calibration algorithm for AVS arrays. As discussed, each AVS consists of a pressure microphone and several orthogonal particle velocity transducers. A particle velocity transducer is commonly referred to as a Microflown [1]. A reliable calibration procedure is crucial for relating the sensor output to the physical quantity perceived. Unlike microphone calibration, there are no standardized procedures yet defined for characterizing the broadband response of particle velocity sensors.

Microflown sensors were originally calibrated using a sound pressure microphone as a reference in a standing wave tube [15], where the ratio between sound pressure and particle velocity (i.e., acoustic impedance) is well understood. Novel methods were later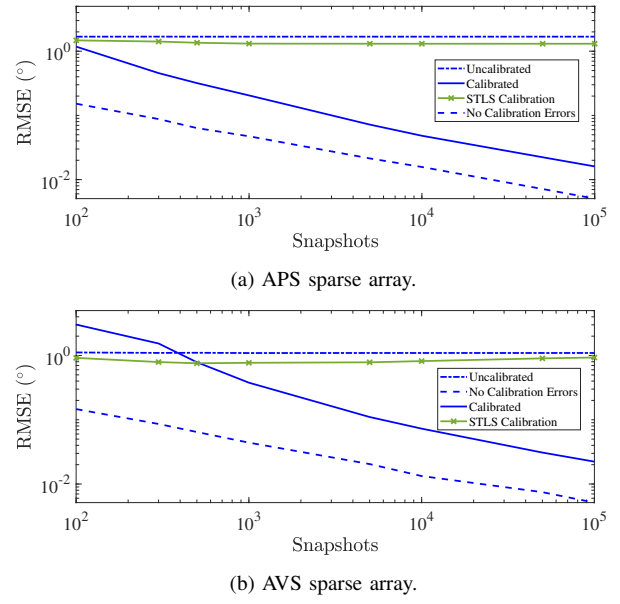 proposed for covering a wider frequency range, such as the "Piston-On-a-Sphere" technique (POS) [14]. This approach relies on a sound source of known impedance measured in free field conditions and it achieves good results at mid and high frequencies. Thereafter, the POS technique was extended to lower frequencies by also measuring the acoustic pressure inside the sound source [16]. As a result, a full-bandwidth calibration procedure is now available by combining two measurement steps. In this section, the DOA estimation results based on the calibrated data using the POS technique (referred to as POS calibration), the Weiss-Friedlander approach [19] and the proposed calibration techniques (both the element-space and co-array approaches) are presented.

A picture of the experimental setup is shown in Fig. 10, where five AVSs are seen arranged in a linear array configuration along with three speakers. The smallest inter-sensor spacing was 0.05 m with sensors located at positions $\mathbf{p} = [0, 1, 2, 4, 6]^T$ and the speakers were located along the circumference of a circle of radius $r = 3.6$ m with respect to the reference AVS in the array (the distance to the sources is more than 20 times the aperture of the array and therefore satisfying the far-field condition). The measurements were carried out in a fully anechoic chamber of the Faculty of Applied Physics of TU Delft (Netherlands) using uncorrelated white Gaussian excitations driving multiple 3 inch loudspeakers (resulting in high SNRs of approximately 30 dB). An Heim DATaRec 24 channels acquisition device with a sampling frequency of 25 kHz was used to record the data. The acoustic pressure and particle velocity information at a given frequency were obtained by computing a short time Fourier transform (STFT). Each recording was fragmented into segments of 1024 samples with 50% overlap. A Hanning window was applied to each

This article has been accepted for publication in IEEE Transactions on Signal Processing. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TSP.2022.3214383

13

Fig. 10. Picture of the experimental setup considering five AVSs and three speakers, located at a radius of $r = 3.6$ m.

data segment prior to the STFT.

The raw output signals from all the five AVSs at a time instant t for a particular frequency bin were collected in a vector $\mathbf{x}(t)$, similar to (1). Without loss of generality, we have considered the first channel of the first AVS in the array as the reference channel with known gain and phase response which is sufficient to obtain a unique solution as seen in the identifiability conditions for AVS arrays. The joint DOA and calibration algorithm based on (22) and (27) were applied on the captured measurement data $\mathbf{x}(t)$ consisting of $L = 1000$ snapshots at a frequency of $f = 2000$ Hz. The corresponding grid-free MUSIC spectra based on the post-calibration measurement data are presented in Fig. 11.

(a) N = 2, $\boldsymbol{\theta} = [-45°, -90°]^T$.

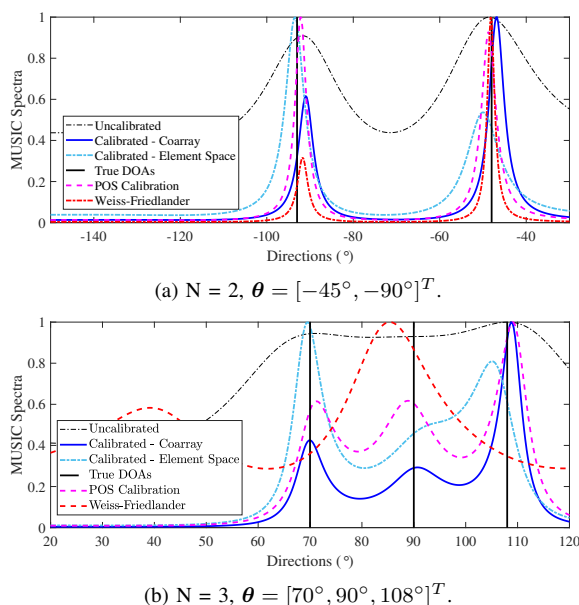(b) N = 3, $\boldsymbol{\theta} = [70°, 90°, 108°]^T$.

Fig. 11. MUSIC spectra based DOA estimates using an AVS array with $M = 5$, $N = 2$ and $f = 2000$ Hz. The true DOAs are indicated by the black solid lines.

In Fig. 11(a) and (b), we considered two of the three speakers with $\boldsymbol{\theta} = [-45°, -90°]^T$ and three speakers that are closely spaced with $\boldsymbol{\theta} = [70°, 90°, 108°]^T$, respectively.

We can observe that for the uncalibrated data, the resolution of MUSIC is poor. However, improved spectra with higher resolution can be seen after compensating with the estimated calibration parameters. The MUSIC spectrum obtained from (27), results in a high resolution comparable to the results that are obtained with the reference POS calibration approach. However, the spectrum obtained from (22), has a lower resolution (especially in the three source case) and shows a small bias compared to the co-array domain based solver. The, Weiss-Friedlander approach results in degraded estimates compared to the proposed approach, specifically in Fig.11(b) it can be observed that none of the sources are resolved.

VIII. CONCLUDING REMARKS

In this paper, we proposed a self calibration technique for both the element-space and co-array data models that is applicable to both acoustic pressure and vector sensor arrays. Also, we derived and discussed a number of identifiability conditions for all the considered cases under which a unique solution for both the calibration parameters and the source DOAs can be obtained. It is interesting to note that for the AVS array, irrespective of the considered geometry, it is possible to calibrate all the sensors with respect to only one of the channels in the array.

Based on the proposed approach, we showed that it is indeed possible to jointly estimate calibration errors and source directions using a *one-step* approach by exploiting the underlying algebraic structure and convex optimization techniques. It is shown that for infinite data records, we can in fact obtain the optimal solution suggesting the feasibility of the convex relaxations for both the element-space and co-array data models. However, when the number of time snapshots are limited and we have a pre-defined grid, we stated that the proposed methodology can be used as a pre-conditioning step to estimate the calibration errors. Then a grid-free approach such as MUSIC/SS-MUSIC can be applied on the gain and phase errors compensated measurement data to obtain improved and reliable DOA estimates. Furthermore, through simulations, we showed that even for finite data records we are able to recover all the source DOAs and we perform better than the existing calibration techniques for all the considered scenarios. Finally, experimental results based on real measurement data with an AVS linear array that are collected in an anechoic chamber are presented to showcase the effectiveness of the proposed calibration techniques using both the element-space and co-array data model.

REFERENCES

[1] H.-E. De Bree, "The microflow e-book," *Microflown Technologies, Arnhem*, 2007.

[2] J. P. Kitchens, "Acoustic vector-sensor array processing," Ph.D. dissertation, Massachusetts Institute of Technology, 2010.

[3] A. Nehorai and E. Paldi, "Acoustic vector-sensor array processing," *IEEE Transactions on Signal Processing*, 1994.

[4] A. L. Swindlehurst and T. Kailath, "A performance analysis of subspace-based methods in the presence of model errors. i. the music algorithm," *IEEE Transactions on signal processing*, 1992.

[5] H. L. Van Trees, *Detection, estimation, and modulation theory. Part IV. , Optimum array processing*. New York: Wiley-Interscience, 2002.

This article has been accepted for publication in IEEE Transactions on Signal Processing. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TSP.2022.3214383

14

[6] J. Capon, "High-resolution frequency-wavenumber spectrum analysis," *Proceedings of the IEEE*, vol. 57, no. 8, pp. 1408–1418, 1969.

[7] R. Roy and T. Kailath, "Esprit-estimation of signal parameters via rotational invariance techniques," *IEEE Transactions on acoustics, speech, and signal processing*, vol. 37, no. 7, pp. 984–995, 1989.

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

[9] S. P. Chepuri, G. Leus *et al.*, "Sparse sensing for statistical inference," *Foundations and Trends® in Signal Processing*, 2016.

[10] A. Moffet, "Minimum-redundancy linear arrays," *IEEE Transactions on antennas and propagation*, vol. 16, no. 2, pp. 172–175, 1968.

[11] P. Pal and P. Vaidyanathan, "Nested arrays: a novel approach to array processing with enhanced degrees of freedom," *Signal Processing, IEEE Transactions on*, vol. 58, no. 8, pp. 4167–4181, 2010.

[12] P. Pal and P. P. Vaidyanathan, "Coprime sampling and the music algorithm," in *Digital Signal Processing Workshop and IEEE Signal Processing Education Workshop (DSP/SPE)*. IEEE, 2011.

[13] A. Paulraj and T. Kailath, "Direction of arrival estimation by eigenstructure methods with unknown sensor gain and phase," in *Acoustics, Speech, and Signal Processing, IEEE International Conference on ICASSP'85.*, vol. 10. IEEE, 1985, pp. 640–643.

[14] F. Jacobsen and V. Jaud, "A note on the calibration of pressure-velocity sound intensity probes," *The Journal of the Acoustical Society of America*, vol. 120, no. 2, pp. 830–837, 2006.

[15] H.-E. d. Bree, W. Druyvesteyn, and M. Elwenspoek, "Realisation and calibration of a novel half inch pu sound intensity probe," in *Audio Engineering Society Convention 106*. Audio Engineering Society, 1999.

[16] T. G. Basten and H.-E. de Bree, "Full bandwidth calibration procedure for acoustic probes containing a pressure and particle velocity sensor," *The Journal of the Acoustical Society of America*, 2010.

[17] D. Astély, A. L. Swindlehurst, and B. Ottersten, "Spatial signature estimation for uniform linear arrays with unknown receiver gains and phases," *IEEE Transactions on Signal Processing*, 1999.

[18] K. N. Ramamohan, S. P. Chepuri, D. F. Comesana, G. C. Pousa, and G. Leus, "Blind calibration for acoustic vector sensor arrays," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2018.

[19] A. J. Weiss and B. Friedlander, "Eigenstructure methods for direction finding with sensor gain and phase uncertainties," *Circuits, Systems and Signal Processing*, vol. 9, no. 3, pp. 271–300, 1990.

[20] S. J. Wijnholds and A.-J. Van Der Veen, "Multisource self-calibration for sensor arrays," *IEEE Transactions on Signal Processing*, vol. 57, no. 9, pp. 3512–3522, 2009.

[21] S. Ling and T. Strohmer, "Self-calibration and bilinear inverse problems via linear least squares," *arXiv preprint arXiv:1611.04196*, 2016.

[22] E. K. Hung, "A critical study of a self-calibrating direction-finding method for arrays," *IEEE Transactions on Signal Processing*, vol. 42, no. 2, pp. 471–474, 1994.

[23] Q. Cheng, Y. Hua, and P. Stoica, "Asymptotic performance of optimal gain-and-phase estimators of sensor arrays," *IEEE Transactions on Signal Processing*, vol. 48, no. 12, pp. 3587–3590, 2000.

[24] M. Viberg and A. L. Swindlehurst, "A bayesian approach to auto-calibration for parametric array signal processing," *IEEE Transactions on Signal Processing*, vol. 42, no. 12, pp. 3495–3507, 1994.

[25] M. P. Wylie, S. Roy, and H. Messer, "Joint doa estimation and phase calibration of linear equispaced (les) arrays," *IEEE Transactions on Signal Processing*, vol. 42, no. 12, pp. 3449–3459, 1994.

[26] Y. Li and M. Er, "Theoretical analyses of gain and phase error calibration with optimal implementation for linear equispaced array," *IEEE Transactions on Signal Processing*, vol. 54, no. 2, pp. 712–723, 2006.

[27] A. Weiss, B. Nadler, and A. Yeredor, "Asymptotically optimal blind calibration of acoustic vector sensor uniform linear arrays," in *ICASSP 2020-2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2020, pp. 4677–4681.

[28] V. C. Soon, L. Tong, Y.-F. Huang, and R. Liu, "A subspace method for estimating sensor gains and phases," *IEEE Transactions on signal processing*, vol. 42, no. 4, pp. 973–976, 1994.

[29] K. Han, P. Yang, and A. Nehorai, "Calibrating nested sensor arrays with model errors," *IEEE Transactions on Antennas and Propagation*, vol. 63, no. 11, pp. 4739–4748, 2015.

[30] P. K. Tam and K. T. Wong, "Cramer-rao bounds for direction finding by an acoustic vector sensor under nonideal gain-phase responses, noncollocation, or nonorthogonal orientation," *IEEE Sensors Journal*, vol. 9, no. 8, pp. 969–982, 2009.

[31] Y. Song and K. T. Wong, "A lower bound of direction-of-arrival estimation for an acoustic vector sensor subject to sensor breakdown," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 48, no. 4, pp. 3703–3708, 2012.

[32] P. K. Tam, K. T. Wong, and Y. Song, "An hybrid cramer-rao bound in closed form for direction-of-arrival estimation by an "acoustic vector sensor" with gain-phase uncertainties," *IEEE transactions on signal processing*, vol. 62, no. 10, pp. 2504–2516, 2014.

[33] X. Yuan, "Direction-finding with a misoriented acoustic vector sensor," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 48, no. 2, pp. 1809–1815, 2012.

[34] Y. Song, K. T. Wong, and F. Chen, ""quasi-blind" calibration of an array of acoustic vector-sensors that are subject to gain errors/mis-location/mis-orientation," *IEEE transactions on signal processing*, vol. 62, no. 9, pp. 2330–2344, 2014.

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

[36] C.-L. Liu and P. Vaidyanathan, "Super nested arrays: Linear sparse arrays with reduced mutual coupling—part i: Fundamentals," *IEEE Transactions on Signal Processing*, vol. 64, no. 15, pp. 3997–4012, 2016.

[37] Z. Ma and K. Ho, "A study on the effects of sensor position error and the placement of calibration emitter for source localization," *IEEE Transactions on Wireless Communications*, vol. 13, no. 10, pp. 5440–5452, 2014.

[38] P. Pal and P. P. Vaidyanathan, "On application of lasso for sparse support recovery with imperfect correlation awareness," in *2012 Conference Record of the Forty Sixth Asilomar Conference on Signals, Systems and Computers (ASILOMAR)*. IEEE, 2012, pp. 958–962.

[39] K. Nambur Ramamohan, M. Contino, S. P. Chepuri, D. F. Comesaña, and G. Leus, "Doa estimation and beamforming using spatially under-sampled avs arrays," in *IEEE 7th International Workshop on CAMSAP*. IEEE, 2017, pp. 1–5.