Optimal Blind Calibration of Uniform Linear Arrays

Amogh Panchagatti⋆, Peter Gerstoft⋆ and Amir Weiss†

⋆University of California, San Diego, La Jolla, CA, USA, {apanchagatti,pgerstoft}@ucsd.edu
†Bar-Ilan University, Ramat Gan, Israel, amir.weiss@biu.ac.il

***Abstract*—Calibration of uniform linear arrays remains an important task that enables proper functionality of almost all spatial array processing downstream operations. We develop a computationally and statistically efficient method for *blind* calibration, i.e., without known calibration signals, that yields the maximum-likelihood estimate (MLE) of the gain and phase offset parameters. We use a computationally lean, but consistent initial estimator, and refine it via Fisher's scoring algorithm, for which we derive the associated Cramér–Rao Bound (CRB) of the model. Simulation results demonstrate that our method converges to the MLE, attains the CRB, and is superior to the recently proposed reduced-maximum-likelihood optimally-weighted least squares method.**

I. Introduction

Systems equipped with sensor arrays are employed in a variety of applications, from radar and sonar to wireless communication and surveillance systems. By using multiple sensors arranged in a well-designed geometry, these systems can capture spatial information that is essential for tasks such as target detection, source localization, and interference mitigation.

A critical procedure that strictly enables these capabilities is *calibration* of the elements in these arrays, a task that remains a critical challenge in such systems (e.g., [1], [2]). In practice, effects such as temperature fluctuations or frequency drift in the receivers can introduce errors in the model parameters, including relative gain and phase mismatches across sensors. These discrepancies can severely degrade system performance, making it necessary to recalibrate the sensor array on a regular basis, if not prior to each use.

Liu *et al.* [3] proposed a diagonal weighted least squares (LS) approach to solve the *blind* calibration problem, i.e., without known calibration signals. Later, Ramamohan et al. [4] addressed this problem by relaxing and reformulating it into a semi-definite programming one. In a different line of work [5]–[7], inspired by Paulraj and Kailath's approach in [8], Weiss and Yeredor developed a computationally attractive optimally-weighted LS (OWLS) estimator as a solution of a system of linear equations from nonlinear transformations of the measurements. However, these solutions are suboptimal for a model that is more accurate in certain settings [7, Sec. IV]. In other words, since the calibration procedure is not as precisely as possible, the accuracy of downstream tasks is not the highest possible. In order to achieve the best possible system performance—such as in direction-of-arrival (DOA) estimation—the calibration procedure must be carried out optimally, so as to enable subsequent spatial inference at the highest fidelity.

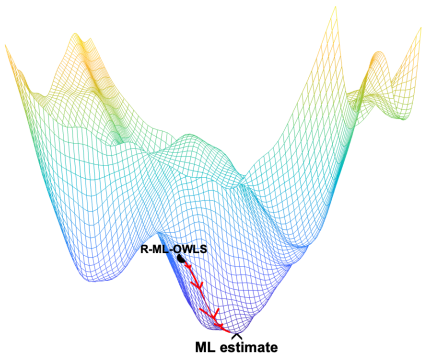

Fig. 1: Simplified illustration of our solution approach. By initializing with the consistent, though suboptimal R-ML-OWLS solution [7], our proposed iterative solution converges to the ML estimate, thus guaranteeing asymptotically optimal performance.

In this work, we address this challenge and develop—to the best of our knowledge—the first optimal blind calibration method for the signal model with the noise term unaffected from the array calibration errors. By utilizing the suboptimal solution proposed in [7]—referred to as the "reduced maximum likelihood OWLS" (R-ML-OWLS)—for initialization, we develop a statistically reliable iterative method, namely Fisher's scoring algorithm (FSA), for the computation of the maximum likelihood (ML) estimator of the unknown calibration gain and phase parameters, as well as of the additional nuisance model parameters. A simplified illustration of this approach is given in Fig. 1. We further derive the Cramér–Rao Bound (CRB) for this model, and validate in simulations our analytical derivations alongside the asymptotic statistical efficiency of our proposed method.

A. Notations

Lowercase letters in standard font (e.g., x) and sans-serif font (e.g., x) denote deterministic and random scalars, respectively; similarly, $\mathbf{x}$ and $\mathbf{x}$ represent deterministic and random vectors, and $\mathbf{X}$ and $\mathbf{X}$ denote deterministic and random matrices, respectively. The notation x_n represents the n-th entry of the random vector $\mathbf{x}$. We denote by $\mathsf{z} \sim \mathcal{CN}(0,1)$ a standard circularly symmetric complex normal random. The operators $(\cdot)^{\mathsf{T}}$, $(\cdot)^{\mathsf{H}}$, and $(\cdot)^{-1}$ denote the transpose,

conjugate transpose, and inverse, respectively, and $\mathbb{T}\mathrm{r}(\mathbf{X})$ is the trace of $\mathbf{X} \in \mathbb{C}^{N\times N}$. Expectation is denoted by $\mathbb{E}[\cdot]$. We use $\mathrm{sgn}(\cdot)$ for the standard sign function, and $\mathrm{Diag}(\mathbf{x})$ for the operator that takes the vector $\mathbf{x}$ and returns a diagonal matrix with $\mathbf{x}$ on its diagonal. Finally, $\nabla_{\boldsymbol{\theta}}$ is the gradient with respect to $\boldsymbol{\theta}$, and $\mathcal{I}\big(\boldsymbol{\theta}\big)$ is the Fisher information matrix (FIM) of $\boldsymbol{\theta}$.

II. Problem Formulation

Consider the presence of N narrowband sources, whose corresponding emitted signals impinge on an array of M sensors. The received signal at each element of the array is modeled as a superposition of these N sources, which depends on the direction-of-arrival (DOA) parameters $\boldsymbol{\alpha} = [\alpha_1\, \alpha_2\, \cdots\, \alpha_N]^\top$, where α_n is the DOA of the n-th source, measured relative to the array's broadside. We focus in this work on uniform linear arrays (ULAs), and we assume that the number of sources satisfies $N < M-1$, ensuring sufficient degrees of freedom for (classical) DOA estimation [9].

Assuming that the received signals are low-pass filtered and sampled at (at least) their Nyquist rate, the resulting vector of baseband samples at time instance t from all M sensors is given by

$$\mathbf{r}[t] = \boldsymbol{\Psi}\boldsymbol{\Phi}\mathbf{A}\mathbf{s}[t] + \mathbf{w}[t] \in \mathbb{C}^{M\times 1},\ \forall\, t \in \{1,\ldots,T\}, \tag{1}$$

where the unknown gain and phase offset parameters, namely the goal estimands, are modeled here as deterministic and are denoted as $\boldsymbol{\psi} \in \mathbb{R}_+^{M\times 1}$ and $\boldsymbol{\phi} \in [-\pi,\pi)^{M\times 1}$, and accordingly,

$$\boldsymbol{\Psi} \triangleq \mathrm{Diag}(\boldsymbol{\psi}) \in \mathbb{R}^{M\times M}, \tag{2}$$

$$\boldsymbol{\Phi} \triangleq \mathrm{Diag}\left(e^{\jmath\boldsymbol{\phi}}\right) \in \mathbb{R}^{M\times M}. \tag{3}$$

The steering vector matrix in (1) is defined as $\mathbf{A} := \mathbf{A}(\boldsymbol{\alpha}) \triangleq [\mathbf{a}(\alpha_1)\ \mathbf{a}(\alpha_2)\ \cdots\ \mathbf{a}(\alpha_N)] \in \mathbb{C}^{M\times N}$, where for a ULA, the n-th steering vector is given by

$$\mathbf{a}(\alpha_n) \triangleq \left[1\ e^{-\jmath\frac{2\pi d}{\lambda}\sin(\alpha_n)}\ \cdots\ e^{-\jmath\frac{2\pi(M-1)d}{\lambda}\sin(\alpha_n)}\right]^\mathsf{T}, \tag{4}$$

and where $d \in \mathbb{R}_+$ and $\lambda \in \mathbb{R}_+$ are the array sensor spacing and the signals wavelength, respectively. We assume that $\mathbf{s}[t] \sim \mathcal{CN}(\mathbf{0}, \mathbf{R}_\mathbf{s})$ is a temporally independent, identically distributed (iid) process with an unknown diagonal covariance matrix $\mathbf{R}_\mathbf{s} \in \mathbb{R}^{M\times M}$. Similarly, we assume that $\mathbf{w}[t] \sim \mathcal{CN}(\mathbf{0}_M, \sigma_\mathsf{w}^2\mathbf{I}_M)$ is iid, and is statistically independent of $\mathbf{s}[t]$. As a result, the received signal is also circular complex normal,

$$\mathbf{r}[t] \sim \mathcal{CN}(\mathbf{0}_M, \mathbf{R}_{\boldsymbol{\theta}}), \quad \forall\, t \in \{1,\ldots,T\}, \tag{5}$$

where

$$\mathbf{R}_{\boldsymbol{\theta}} \triangleq \mathbb{E}[\mathbf{r}[t]\mathbf{r}[t]^\mathsf{H}] \tag{6}$$

$$= \boldsymbol{\Psi}\boldsymbol{\Phi}\mathbf{A}\mathbf{R}_\mathbf{s}\mathbf{A}^\mathsf{H}\boldsymbol{\Phi}^\mathsf{H}\boldsymbol{\Psi}^\mathsf{H} + \sigma_\mathsf{w}^2\mathbf{I}_M \tag{7}$$

$$\triangleq \boldsymbol{\Psi}\boldsymbol{\Phi}\mathbf{C}\boldsymbol{\Phi}^*\boldsymbol{\Psi} + \sigma_\mathsf{w}^2\mathbf{I}_M \in \mathbb{C}^{M\times M}, \tag{8}$$

and $\mathbf{C} = \mathbf{A}\mathbf{R}_\mathbf{s}\mathbf{A}^\mathsf{H} \in \mathbb{C}^{M\times M}$ is a Hermitian Toeplitz matrix since $\mathbf{A}$ is a Vandermonde matrix. Note that since $\mathbf{C}$ is Hermitian Toeplitz, it is completely characterized by its first row, denoted as $\mathbf{c} \triangleq [C_{11}\, C_{12}\, \cdots\, C_{1M}]^\mathsf{T} \in \mathbb{C}^{M\times 1}$, for which we define

$$\mathbf{c} \triangleq \boldsymbol{\rho} + \jmath\,\boldsymbol{\iota} \in \mathbb{C}^{M\times 1}, \tag{9}$$

where $\boldsymbol{\rho} \in \mathbb{R}^{M\times 1}$ and $\boldsymbol{\iota} \in \mathbb{R}^{M\times 1}$ denote its real and imaginary parts, respectively. The covariance matrix (6) is denoted as $\mathbf{R}_{\boldsymbol{\theta}}$, with subindex $\boldsymbol{\theta}$, to emphasize its parametric structure with respect to the unknown model parameters, collected into the vector

$$\boldsymbol{\theta} \triangleq \left[\boldsymbol{\psi}^\mathsf{T}\boldsymbol{\phi}^\mathsf{T}\boldsymbol{\rho}^\mathsf{T}\boldsymbol{\iota}^\mathsf{T}\sigma_\mathsf{w}^2\right]^\mathsf{T} \in \mathbb{R}^{(4M-3)\times 1}. \tag{10}$$

The total number of real-valued parameters is $4M-3$, which are:

- The real part vector $\boldsymbol{\rho} = [\rho_1 \cdots \rho_M]^\mathsf{T}$, contributing M unknown parameters;
- The imaginary part vector $\boldsymbol{\iota} = [\iota_1 \cdots \iota_M]^\mathsf{T}$, with ι_1 fixed to zero since C_{11} is real, contributing $M-1$ unknown parameters;
- The phases vector $\boldsymbol{\phi} = [\phi_1 \cdots \phi_M]^\mathsf{T}$, where ϕ_1 and ϕ_2 are fixed to zero, without loss of generality, due to the rotational ambiguity of the DOAs, contributing $M-2$ unknown parameters;
- The gains vector $\boldsymbol{\psi} = [\psi_1 \cdots \psi_M]^\mathsf{T}$, where without loss of generality $\psi_1 = 1$ as a reference, contributing $M-1$ unknown parameters; and
- A single (nonnegative) real-valued parameter σ_w^2, namely to the noise variance.

The problem we address in this work is as follows: Given the statistically independent measurements $\{\mathbf{r}[t]\}_{t=1}^T$ whose (identical) distribution is prescribed by (5), the task is to estimate the unknown calibration parameters $\{\boldsymbol{\psi}, \boldsymbol{\phi}\}$. Note that $\mathbf{R}_{\boldsymbol{\theta}}$, which depends on these parameters, fully characterizes the second-order statistics—and hence the distribution—of $\{\mathbf{r}[t]\}_{t=1}^T$. Thus, estimating $\boldsymbol{\theta}$ yields $\mathbf{R}_{\boldsymbol{\theta}}$, which can then be used for subsequent inference tasks.

A. Maximum Likelihood Blind Calibration

Given the signal model (5), we now formulate the ML estimation problem for blind calibration. For T independent snapshots $\{\mathbf{r}_t := \mathbf{r}[t]\}_{t=1}^T$, the likelihood of all T snapshots is the product of the individual likelihoods, given by

$$L(\boldsymbol{\theta} \mid \mathbf{r}_1,\ldots,\mathbf{r}_T) = \prod_{t=1}^T \frac{1}{\sqrt{(2\pi)^M \det(\mathbf{R}_{\boldsymbol{\theta}})}}\, \exp\!\left(-\tfrac{1}{2}\,\mathbf{r}_t^\mathsf{H}\mathbf{R}_{\boldsymbol{\theta}}^{-1}\mathbf{r}_t\right). \tag{11}$$

Taking the natural logarithm of (11), multiplying by $1/T$, omitting terms independent of $\boldsymbol{\theta}$, and using

$$\sum_{t=1}^T \mathbf{r}_t^\mathsf{H}\mathbf{R}_{\boldsymbol{\theta}}^{-1}\mathbf{r}_t \triangleq T\cdot\mathbb{T}\mathrm{r}\Big(\mathbf{R}_{\boldsymbol{\theta}}^{-1}\mathbf{S}\Big), \tag{12}$$

where $\mathbf{S} \triangleq \frac{1}{T}\sum_{t=1}^T \mathbf{r}_t\mathbf{r}_t^\mathsf{H}$ is the sample covariance matrix, we obtain the (constants-free) normalized log-likelihood

$$\mathcal{L}(\boldsymbol{\theta} \mid \mathbf{r}_1,\ldots,\mathbf{r}_T) \triangleq -\tfrac{1}{T}\log\big(\det(\mathbf{R}_{\boldsymbol{\theta}})\big) - \mathbb{T}\mathrm{r}\left(\mathbf{R}_{\boldsymbol{\theta}}^{-1}\mathbf{S}\right). \tag{13}$$

Maximizing (13) with respect to $\boldsymbol{\theta}$ yields the ML estimate (MLE),

$$\widehat{\boldsymbol{\theta}} = \arg\max_{\boldsymbol{\theta}}\ \mathcal{L}\big(\boldsymbol{\theta} \mid \mathbf{r}_1,\ldots,\mathbf{r}_T\big),$$

which in particular provides asymptotically efficient estimators of ψ, ϕ for blind calibration of the array.[1]

III. Proposed Calibration Algorithm

In Subsection III-A below, we describe FSA, which we use in our proposed method to iteratively increase the (normalized) log-likelihood function. Of course, as in most nonlinear, nonconvex optimization problems, a key challenge with using a gradient-based method is determining an appropriate initial estimate that (at least typically) lies in the basin of attraction of the MLE. In Subsection III-B, we explain why the R-ML-OWLS [7] serves as such an initial estimate. Our complete method is presented in Algorithm 1. Once $\widehat{\boldsymbol{\theta}}$, the MLE of $\boldsymbol{\theta}$ is obtained, $\widehat{\psi}$ and $\widehat{\phi}$ can be used for optimal calibration of the array.

A. Fisher's Scoring Algorithm

In order to find the estimate $\widehat{\boldsymbol{\theta}}$ that maximizes $\mathcal{L} := \mathcal{L}(\boldsymbol{\theta} \mid \mathbf{r}_1, \ldots, \mathbf{r}_T)$, namely the MLE, one can employ a Newton–Raphson-type iterative scheme. FSA is such a convenient and theoretically justified option [10], where the exact Hessian is replaced by the FIM. Specifically, at the k-th iteration, the update rule of FSA is given by (e.g., [11, Eq. (7.5)])

$$\boldsymbol{\theta}^{(k+1)} = \boldsymbol{\theta}^{(k)} + \left[\mathcal{I}\left(\boldsymbol{\theta}^{(k)}\right)\right]^{-1} \nabla_{\boldsymbol{\theta}} \mathcal{L}(\boldsymbol{\theta})\big|_{\boldsymbol{\theta}=\boldsymbol{\theta}^{(k)}}, \tag{14}$$

where $\nabla_{\boldsymbol{\theta}} \mathcal{L}(\boldsymbol{\theta})$ is the gradient of the normalized log-likelihood with respect to $\boldsymbol{\theta}$ and $\mathcal{I}(\boldsymbol{\theta})$ is the FIM, which are derived in the Appendix, both evaluated in (14) at the current estimate $\boldsymbol{\theta}^{(k)}$. By iteratively applying this update, the algorithm converges to a stationary point of the likelihood. However, without proper initialization, it may (and most probably) converge to a *local* maximum that is not the global optimizer, which does not correspond to the MLE. To mitigate this risk, we propose the following initialization, designed to facilitate convergence to the MLE.

B. Statistically Consistent Initialization

Although the R-ML-OWLS estimator proposed in [7] is not exploiting all the available data, and is therefore suboptimal—and in particular, not the MLE—for the signal model (1), it is nevertheless a *consistent* estimator. Moreover, it has a simple analytical closed-form expression [7, Eq. (45)]. Hence, it can serve as an initial solution for FSA that is guaranteed to approach the true value of $\boldsymbol{\theta}$ as T increases. Consequently, as T increases, it is likely that this initial solution will reside in the basin of attraction of the sought after MLE. We therefore argue that the R-ML-OWLS suboptimal estimator is a "good" initial solution, as empirically demonstrated in Section V.

IV. Validation of Analytical Derivations

We now empirically validate our derivations of the CRB and the gradient of the log-likelihood, given in the Appendix, using FSA. Specifically, for each

[1]We assume that all regularity conditions regarding standard ML estimation are met, and in particular that the model is identifiable.

Algorithm 1: R-ML-OWLS-initialized FSA

Input: $\{\mathbf{r}_t\}_{t=1}^T$, convergence tolerance $\epsilon \in \mathbb{R}_+$, maximum number of iterations $K_{\mathrm{m}} \in \mathbb{N}$

Output: The estimate $\widehat{\boldsymbol{\theta}}$, which includes $\widehat{\psi}, \widehat{\phi}$

1) Compute the sample covariance matrix:
$$\mathbf{S} \leftarrow \frac{1}{T} \sum_{t=1}^{T} \mathbf{r}_t \mathbf{r}_t^{\mathsf{H}}$$
2) Obtain the R-ML-OWLS [7] estimate of $\boldsymbol{\theta}$, set it as the initial estimate $\boldsymbol{\theta}^{(0)}$, and set $k \leftarrow 0$.
3) **While** $\|\boldsymbol{\theta}^{(k)} - \boldsymbol{\theta}^{(k-1)}\| \geq \epsilon$ and $k \leq K_{\mathrm{m}}$ **do**:
   a) Compute $\nabla_{\boldsymbol{\theta}} \mathcal{L}(\boldsymbol{\theta})\big|_{\boldsymbol{\theta}=\boldsymbol{\theta}^{(k)}}$ and $\mathcal{I}\left(\boldsymbol{\theta}^{(k)}\right)$ (see (20) and (27), respectively, in the Appendix).
   b) Update:
   $$\boldsymbol{\theta}^{(k+1)} \leftarrow \boldsymbol{\theta}^{(k)} + \left[\mathcal{I}\left(\boldsymbol{\theta}^{(k)}\right)\right]^{-1} \nabla_{\boldsymbol{\theta}} \mathcal{L}(\boldsymbol{\theta})\big|_{\boldsymbol{\theta}=\boldsymbol{\theta}^{(k)}}$$
   c) Increment $k \leftarrow k + 1$.
4) **Return** $\widehat{\boldsymbol{\theta}} \leftarrow \boldsymbol{\theta}^{(k)}$.

unknown θ_i, the CRB is given by the (i,i)-th entry of $\mathcal{I}^{-1}(\boldsymbol{\theta})$. We conduct 10^3 Monte Carlo (MC) trials, compute the parameter-wise empirical square errors, and compare them with the analytic expressions of their respective CRBs.

For validation of our derivations only, we initialize FSA at the *true parameter values*, leveraging the fact that for any *finite* sample size T, the MLE is almost surely *not* equal to the true parameter vector.[2] Let $L \in \mathbb{N}$ denote the number of MC trials. For each parameter θ_i, we define its estimated squared error (ESE) as

$$\mathrm{ESE}(\theta_i) \triangleq \frac{1}{L} \sum_{\ell=1}^{L} \left(\theta_i - \widehat{\theta}_i^{[\ell]}\right)^2, \tag{15}$$

where $\widehat{\theta}_i^{[\ell]}$ denotes the estimate of θ_i in the ℓ-th MC trial. We then aggregate these ESEs into an average squared error (ASE),

$$\mathrm{ASE}(\boldsymbol{\theta}) \triangleq \frac{1}{4M-3} \sum_{i=1}^{4M-3} \mathrm{ESE}(\theta_i), \tag{16}$$

and compare it with the average of the diagonal elements of the inverse FIM, termed here as the average CRB (ACRB),

$$\mathrm{ACRB}(\boldsymbol{\theta}) \triangleq \frac{1}{4M-3} \sum_{i=1}^{4M-3} \left[\mathcal{I}^{-1}(\boldsymbol{\theta})\right]_{ii}. \tag{17}$$

As shown in Figs. 2 and 3, the ASE closely matches the ACRB,[3] confirming both the validity of our gradient and FIM computation, and the asymptotic efficiency of the estimator, when properly initialized.

We have thus established the CRB for model (1), i.e., a lower bound on the mean squared error (MSE) of any unbiased estimator of $\boldsymbol{\theta}$. This bound serves as a benchmark against which all subsequent blind

[2]When the MLE is consistent, it converges in probability to the true parameter vector in the limit $T \to \infty$. However, for our signal model (and specifically with complex normal noise), for any finite T we have $\mathbb{P}\left(\widehat{\boldsymbol{\theta}} \neq \boldsymbol{\theta}\right) = 1$.

[3]Except at extremely low SNR, as expected.

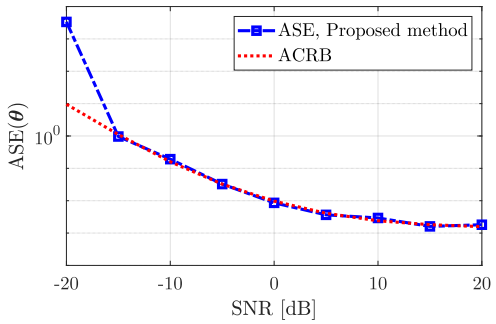

Fig. 2: The ASE and ACRB versus SNR for $T = 10^3$. Already from -15 dB the average of estimated MSEs of our proposed estimators coincides with the average of the corresponding CRBs.

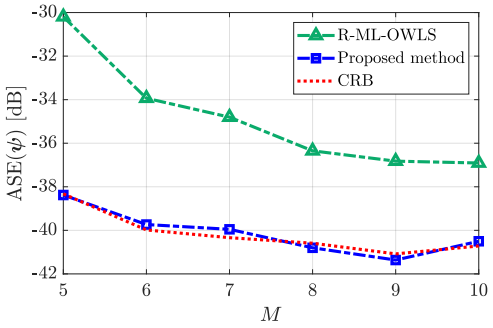

Fig. 4: ASE($\boldsymbol{\psi}$) versus M for $T = 3000$ and SNR $= 10$ dB. A significant gain in the MSE relative to the R-ML-OWLS is observed.

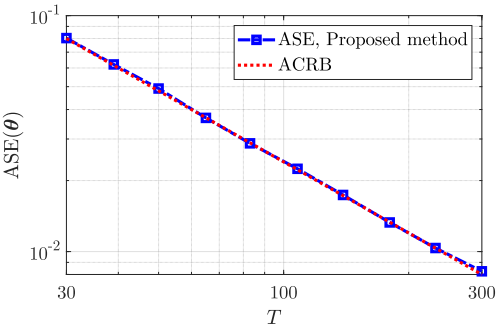

Fig. 3: The ASE and ACRB versus T for an SNR level of 10dB. Our method reliably converges to the MLE, and already for a moderate sample size of $T = 30$ achieves the CRB.

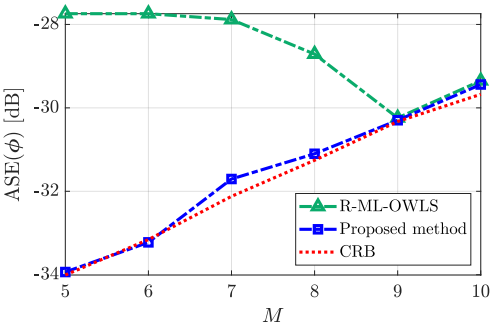

Fig. 5: ASE($\boldsymbol{\phi}$) versus M for $T = 3000$ and SNR $= 10$ dB. While R-ML-OWLS approaches the CRB only for $M \geq 9$, our method yields optimal calibration performance even for $M = 5$.

calibration methods can be compared, and indicates the best achievable estimation accuracy.

V. Simulation Results

We now evaluate the proposed blind calibration method, where we consider model (1), with (5), in a finite-snapshot representative scenario with:

- $N = 3$ sources;
- Sensor array sizes $M \in \{5, 6, 7, 8, 9, 10\}$;
- $T = 3000$ snapshots; and
- SNR $= 10$ dB.

The base gains and phases for the smallest array size ($M = 5$) are $\boldsymbol{\psi}_{\text{base}} = [1\ 1.3\ 1.1\ 0.7\ 2.2]$ and $\boldsymbol{\phi}_{\text{base}} = \frac{\pi}{180}[\,0\ 0\ 5\ 11\ -8]$, respectively. For arrays with more than $M = 5$ elements, these vectors are repeated and truncated to match the actual sensor dimension. We fix the source angles at $\boldsymbol{\alpha} = [-35^\circ\ -73^\circ\ -28^\circ]^{\mathsf{T}}$.

In our simulation, we vary the number of sensors M from 5 to 10 and compute the average MSE for both the gain and phase parameters over $L = 500$ MC independent trials. Estimates are obtained using both the suboptimal R-ML-OWLS and our proposed method. The ASEs for $\boldsymbol{\psi}$ and $\boldsymbol{\phi}$ are defined as

$$\text{ASE}(\boldsymbol{\psi}) \triangleq \frac{1}{L(M-1)} \sum_{\ell=1}^{L} \sum_{i=2}^{M} \left(\widehat{\psi}_i^{[\ell]} - \psi_i\right)^2,$$

$$\text{ASE}(\boldsymbol{\phi}) \triangleq \frac{1}{L(M-2)} \sum_{\ell=1}^{L} \sum_{i=3}^{M} \left(\widehat{\phi}_i^{[\ell]} - \phi_i\right)^2.$$

Figure 4, presenting ASE($\boldsymbol{\psi}$) vs. M, shows that our proposed method attains the CRB, and significantly improves upon R-ML-OWLS. Additionally, R-ML-OWLS gradually approaches the CRB as M increases, which aligns with the discussion on the regime where $M \to \infty$ in [7, Section IV]. Likewise, in Fig. 5 our method outperforms R-ML-OWLS in estimation of the phase parameters, with R-ML-OWLS converging to the CRB only for $M \geq 9$. Moreover, our method essentially reaches the CRB for both gains and phases, confirming its optimal performance (i.e., asymptotic efficientcy), and its superiority over R-ML-OWLS.

VI. Concluding Remarks

We derive a statistically reliable algorithm for the computation of the MLE for blind calibration of (commonly used) uniform linear arrays. Our method utilizes the previously proposed, suboptimal R-ML-OWLS estimator, and successively refines it via FSA, which converges to the MLE. Our derivation include the corresponding CRB, whose benefit in our context is twofold: for FSA, and as a benchmark for the best attainable performance in terms of MSE among all unbiased estimators. Simulations corroborate that the proposed method essentially attains this bound.

Extensions to more general array geometries and broader signal models are considerably more challenging—at least analytically—and constitute an attractive path for future research. As modern array processing is increasingly embracing novel configurations, there is a growing need for innovative methods to effectively address the resulting complexity.

VII. Appendix

For the sake of brevity, let us denote $\mathbf{R}_{\boldsymbol{\theta}}$ as $\mathbf{R}$. The core idea of our derivation is based on the following relation: for $i,j \in \{1,\ldots,M\}$,

$$\mathbf{R}_{ij} = \begin{cases} \mathbf{C}_{ii}\psi_i^2 + \sigma_{\mathsf{w}}^2 & \text{if } i=j \\ \mathbf{C}_{ij}\psi_i\psi_j e^{\jmath(\phi_i-\phi_j)} & \text{otherwise} \end{cases}, \tag{18}$$

where we recall

$$\mathcal{L} = -\tfrac{1}{T}\log\big(\det(\mathbf{R})\big) - \mathbb{T}\mathrm{r}\left(\mathbf{R}^{-1}\mathbf{S}\right), \tag{19}$$

namely $\mathcal{L}$ is a function of $\mathbf{R}$ and $\mathbf{S}$ only. Taking the derivative w.r.t θ_k and using the chain rule,

$$\frac{\partial \mathcal{L}}{\partial \theta_k} = \sum_{\ell_1,\ell_2=1}^{M} \frac{\partial \mathcal{L}}{\partial R_{\ell_1\ell_2}} \frac{\partial R_{\ell_1\ell_2}}{\partial \theta_k}, \tag{20}$$

Using the following relations (e.g., [12]–[14])

$$\frac{\partial \log\left(\det(\mathbf{R})\right)}{\partial R_{ij}} = \left(\mathbf{R}^{-1}\right)_{ji}, \tag{21}$$

$$\frac{\partial\, \mathbb{T}\mathrm{r}\left(\mathbf{R}^{-1}\mathbf{S}\right)}{\partial R_{ij}} = -\left(\mathbf{R}^{-1}\mathbf{S}\mathbf{R}^{-1}\right)_{ji}, \tag{22}$$

the derivative of $\mathcal{L}$ with respect to R_{ij} is:

$$\frac{\partial \mathcal{L}}{\partial R_{ij}} = -\frac{1}{T}\left(\mathbf{R}^{-1}\right)_{ji} + \left(\mathbf{R}^{-1}\mathbf{S}\mathbf{R}^{-1}\right)_{ji}.$$

We now move on to compute the partial derivatives of the covariance matrix $\mathbf{R} \in \mathbb{C}^{M\times M}$ with respect to five groups of parameters, in the order described in (10). Recall that

$$\mathbf{R} = \boldsymbol{\Psi}\boldsymbol{\Phi}\mathbf{C}\boldsymbol{\Phi}^*\boldsymbol{\Psi} + \sigma_{\mathsf{w}}^2\mathbf{I}_M, \tag{23}$$

where $\mathbf{C}$ is Hermitian Toeplitz with its first row is determined by $\{\rho_k\}$ and $\{\iota_k\}$.

1) Gradient with respect to the gain parameters: For any $m \in \{2,\ldots,M\}$,

$$\frac{\partial R_{ij}}{\partial \psi_m} = \begin{cases} C_{ij}[\delta_{mi}\psi_j + \delta_{mj}\psi_i]e^{\jmath(\phi_i-\phi_j)} & i \neq j \\ 2\,C_{ii}\psi_i & i=j=m \\ 0 & \text{otherwise.} \end{cases} \tag{24}$$

Here, δ_{mi} denotes the Kronecker delta.

2) Gradient with respect to the phase parameters: For any $m \in \{3,\ldots,M\}$,

$$\frac{\partial R_{ij}}{\partial \phi_m} = \begin{cases} \jmath C_{ij}\psi_i\psi_j e^{\jmath(\phi_i-\phi_j)} & m=i\neq j \\ -\jmath C_{ij}\psi_i\psi_j e^{\jmath(\phi_i-\phi_j)} & m=j\neq i \\ 0 & \text{otherwise} \end{cases} \tag{25}$$

3) Gradient with respect to ρ_k:

$$\frac{\partial R_{ij}}{\partial \rho_k} = \begin{cases} \psi_i\,\psi_j\,e^{\jmath(\phi_i-\phi_j)} & |i-j| = |k-1| \\ 0, & \text{otherwise.} \end{cases}$$

4) Gradient with respect to ι_k: Similarly, ι_k is the imaginary part of C_{1k}, whose effect on $\mathbf{R}$ resembles ρ_k but includes a factor of $\pm\jmath$. Concretely, for $k \in \{2,\ldots,M\}$:

$$\frac{\partial R_{ij}}{\partial \iota_k} = \begin{cases} \jmath\,\mathrm{sgn}(j-i)\psi_i\psi_j e^{\jmath(\phi_i-\phi_j)} & |i-j|=|k-1| \\ 0 & \text{otherwise.} \end{cases}$$

The sgn function appears due to the conjugate symmetry property of the matrix $\mathbf{R}$.

5) Gradient computation with respect to σ_{w}^2: For the noise variance σ_{w}^2,

$$\frac{\partial R_{ij}}{\partial \sigma_{\mathsf{w}}^2} = \delta_{ij} \;\Rightarrow\; \frac{\partial \mathbf{R}}{\partial \sigma_{\mathsf{w}}^2} = \mathbf{I}_M. \tag{26}$$

A. Derivation of the FIM

Since $\mathbf{r}[t]$ is an iid zero-mean circular complex normal process with covariance $\mathbf{R}_{\boldsymbol{\theta}}$, i.e., $\mathbf{r}[t] \sim \mathcal{CN}\big(\mathbf{0}_M, \mathbf{R}_{\boldsymbol{\theta}}\big)$, the (k,ℓ)-th entry of the FIM takes the form [15, Eq. (5.22)]:

$$[\mathcal{I}(\boldsymbol{\theta})]_{k\ell} = \mathbb{T}\mathrm{r}\left(\mathbf{R}_{\boldsymbol{\theta}}^{-1}\frac{\partial \mathbf{R}_{\boldsymbol{\theta}}}{\partial \theta_k}\mathbf{R}_{\boldsymbol{\theta}}^{-1}\frac{\partial \mathbf{R}_{\boldsymbol{\theta}}}{\partial \theta_\ell}\right). \tag{27}$$

Since we have computed all the partial derivatives of the covariance matrix $\mathbf{R}_{\boldsymbol{\theta}}$ with respect to each of the unknowns, namely θ_i for each $1 \le i \le 4M-3$, each of the elements of the FIM is readily given based on these expressions.

References

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

[2] R. Pöhlmann, S. Zhang, E. Staudinger, S. Caizzone, A. Dammann, and P. A. Hoeher, "Bayesian in-situ calibration of multiport antennas for doa estimation: Theory and measurements," *IEEE Access*, vol. 10, pp. 37 967–37 983, 2022.

[3] A. Liu, M. Tegmark, S. Morrison, A. Lutomirski, and M. Zaldarriaga, "Precision calibration of radio interferometers using redundant baselines," *Monthly Notices of the Royal Astronomical Society*, vol. 408, no. 2, pp. 1029–1050, 2010.

[4] K. N. Ramamohan, S. P. Chepuri, D. F. Comesaña, and G. Leus, "Blind sensor array calibration and DOA estimation of broadband sources," in *2019 53rd Asilomar Conference on Signals, Systems, and Computers*, 2019, pp. 1304–1308.

[5] A. Weiss and A. Yeredor, "Blind calibration of sensor arrays for narrowband signals with asymptotically optimal weighting," in *2019 27th European Signal Processing Conference (EUSIPCO)*, 2019, pp. 1–5.

[6] A. Weiss and A. Yeredor, "Non-iterative blind calibration of nested arrays with asymptotically optimal weighting," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP)*, 2021, pp. 4630–4634.

[7] A. Weiss and A. Yeredor, "Asymptotically optimal blind calibration of uniform linear sensor arrays for narrowband Gaussian signals," *IEEE Trans. Signal Process.*, vol. 68, pp. 5322–5333, 2020.

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

[9] D. Astély, A. L. Swindlehurst, and B. Ottersten, "Spatial signature estimation for uniform linear arrays with unknown receiver gains and phases," *IEEE Trans. Signal Process.*, vol. 47, no. 8, pp. 2128–2138, 1999.

[10] M. R. Osborne, "Fisher's method of scoring," *International Statistical Review / Revue Internationale de Statistique*, vol. 60, no. 1, pp. 99–117, 1992.

[11] S. M. Kay, *Fundamentals of statistical signal processing: estimation theory*. Prentice-Hall, Inc., 1993.

[12] K. Petersen and M. Pedersen, "The matrix cookbook," *Technical University of Denmark, Kongens Lyngby*, 2012.

[13] P. Gerstoft, C. F. Mecklenbräuker, A. Xenaki, and S. Nannuru, "Multisnapshot sparse Bayesian learning for doa," *IEEE Signal Processing Letters*, vol. 23, no. 10, pp. 1469–1473, 2016.

[14] Y. Park, P. Gerstoft, and C. F. Mecklenbräuker, "Atom-constrained gridless doa refinement with wirtinger gradients," *IEEE Open Journal of Signal Processing*, 2024.

[15] H. L. Van Trees, *Optimum array processing: Part IV of detection, estimation, and modulation theory*. John Wiley & Sons, 2002.