2019 27th European Signal Processing Conference (EUSIPCO)

Blind Calibration of Sensor Arrays for Narrowband Signals with Asymptotically Optimal Weighting

Amir Weiss and Arie Yeredor
School of Electrical Engineering, Tel-Aviv University, Israel
amirwei2@mail.tau.ac.il, arie@eng.tau.ac.il

***Abstract*—We revisit the problem of blind calibration of uniform linear sensors arrays for narrowband signals and set the premises for the derivation of the optimal blind calibration scheme. In particular, instead of taking the direct (rather involved) Maximum Likelihood (ML) approach for joint estimation of all the unknown model parameters, we follow Paulraj and Kailath's classical approach in exploiting the special (Toeplitz) structure of the observed covariance. However, we offer a substantial improvement over Paulraj and Kailath's Least Squares (LS) estimate by using asymptotic approximations in order to obtain simple, (quasi-)linear Weighted LS (WLS) estimates of the sensors' gains and phases offsets with asymptotically optimal weighting. As we show in simulation experiments, our WLS estimates exhibit near-optimal performance, with a considerable improvement (reaching an order of magnitude and more) in the resulting mean squared errors, w.r.t. the corresponding ordinary LS estimates. We also briefly explain how the methodology derived in this work may be utilized in order to obtain (by certain modifications) the asymptotically optimal ML estimates w.r.t. the raw data via a (quasi)-linear WLS estimate.**

***Index Terms*—Sensor array processing, gain estimation, phase estimation, self-calibration, weighted least squares.**

I. Introduction

Array processing offers a host of signal processing tools serving to infer essential information related to signals impinging on the array. The accuracy of the associated estimates naturally depends not only on conventional, modeled error-sources (such as noise), but also on modeling-errors which reflect possible deviations (in practice) of the array parameters from their nominal values. To mitigate such modeling errors, e.g., in the gains and phases of the elements of the array, a calibration procedure is usually employed.

While "offline" calibration (i.e., prior to the "operational" period), using known calibration signals when possible, is relatively simple, self or blind calibration is typically a more desirable, yet a more challenging task. In this paper, we address the blind calibration of the gains and phases in a sensor array within the framework of narrowband signals. Naturally, this problem has already been widely addressed in the literature and is quite well-studied. A few important examples are Paulraj and Kailath's Least Squares (LS) (based) estimators for the unknown sensor gains and phases [1], Friedlander and Weiss' eigenstructure method [2], which jointly calibrates the array and estimate the sources' Direction-Of-Arrivals (DOAs), and the direct (rather involved) Maximum Likelihood (ML) approach, proposed in [3] by Chong and See, incorporating also mutual coupling as well as errors in the sensor positions, in which the ML Estimate (MLE) is pursued by an iterative algorithm. More recent examples are due to Liu *et al.*'s [4] and Wijnholds and Noorishad [5], where a diagonal Weighted LS (WLS) and the weighted alternating LS estimators are proposed, resp. Nevertheless, these weighting approaches are essentially heuristic and are not shown (nor claimed) to be optimal.

In this paper we revisit the problem of blind sensor gains and phases estimation, when the sources' DOAs and powers, as well as the noise level, are considered unknown. We derive closed-form (approximate) expressions for Optimally-Weighted LS (OWLS) estimates of the gains and phases. The provided (non-iterative) solutions are efficiently computed, and as we demonstrate in simulation, the resulting Mean Squared Errors (MSEs) are improved (in some scenarios) by more than an order of magnitude w.r.t. the MSEs attained by the Paulraj and Kailath's ordinary LS estimators, and approach the performance bounds (which are otherwise attained asymptotically only by *joint* ML estimation of *all* the unknown model parameters).

A. Notations

We use $x, \boldsymbol{x}$ and $\boldsymbol{X}$ for a scalar, column vector and matrix, resp. The superscripts $(\cdot)^{\mathrm{T}}$, $(\cdot)^{*}$, $(\cdot)^{\dagger}$ and $(\cdot)^{-1}$ denote the transposition, complex conjugation, conjugate transposition and inverse operators, resp. Further, $|z|$ and $\angle z$ denote the modulus and angle of a complex scalar $z \in \mathbb{C}$, resp., while $\Re\{z\}$ and $\Im\{z\}$ denote its real and imaginary parts, resp. We use $\boldsymbol{I}_K$ to denote the $K \times K$ identity matrix. $\mathbb{E}[\cdot]$ denotes expectation and the $\mathrm{Diag}(\cdot)$ operator forms an $M \times M$ diagonal matrix from its M-dimensional vector argument.

II. Problem Formulation

Consider a uniform linear array of M sensors, each with an *unknown* (deterministic) gain and phase response, and the presence of $D < M$ (unknown) narrowband sources, centered around some common carrier frequency with a wavelength λ, which are sufficiently far from the array to allow a planar wavefront ("far-field") approximation. Thus, let us denote the unknown gain and phase offset parameters as $\boldsymbol{\psi} \in \mathbb{R}_{+}^{M \times 1}$ and $\boldsymbol{\phi} \in \mathbb{R}^{M \times 1}$, resp., where ψ_m and ϕ_m are (resp.) the unknown gain and phase offsets of the m-th sensor.

More specifically, assuming the received signals are Low-Pass Filtered (LPF)[1] and sampled at (at least) the Nyquist

[1]The bandwidth of the LPF exceeds the bandwidth of the widest source.

rate, the vector of sampled (baseband) signals from all the M sensors is given by (for all $t \in \{1, \ldots, T\}$)

$$\boldsymbol{r}[t] = \boldsymbol{\Psi}\boldsymbol{\Phi}\left(\boldsymbol{A}(\boldsymbol{\theta})\boldsymbol{s}[t] + \boldsymbol{n}[t]\right) \triangleq \boldsymbol{\Psi}\boldsymbol{\Phi}\boldsymbol{x}[t] \in \mathbb{C}^{M\times 1}, \quad (1)$$

where

(i) $\boldsymbol{\Psi} \triangleq \mathrm{Diag}(\boldsymbol{\psi}) \in \mathbb{R}_{+}^{M\times M}$, $\boldsymbol{\Phi} \triangleq \mathrm{Diag}\left(e^{\jmath\boldsymbol{\phi}}\right) \in \mathbb{C}^{M\times M}$;

(ii) $\boldsymbol{s}[t] \triangleq [s_1[t] \,\cdots\, s_D[t]]^{\mathrm{T}} \in \mathbb{C}^{D\times 1}$ denotes the vector of sources with wavenumber $k = 2\pi/\lambda$, impinging on the array from azimuth angles $\boldsymbol{\theta} \triangleq [\theta_1 \,\cdots\, \theta_D]^{\mathrm{T}} \in \mathbb{R}^{D\times 1}$;

(iii) $\boldsymbol{A}(\boldsymbol{\theta}) \triangleq [\boldsymbol{a}(\theta_1) \,\cdots\, \boldsymbol{a}(\theta_D)] \in \mathbb{C}^{M\times D}$ denotes the array manifold matrix, with the steering vectors $\boldsymbol{a}(\theta_d) \triangleq \left[1 \; e^{\jmath k\gamma\cos(\theta_d)} \,\ldots\, e^{\jmath k(M-1)\gamma\cos(\theta_d)}\right]^{\mathrm{T}} \in \mathbb{C}^{M\times 1}$ as its columns (γ being the inter-element spacing);

(iv) $\boldsymbol{n}[t] \in \mathbb{C}^{M\times 1}$ denotes an additive noise vector, assumed to be (both spatially and temporally) independent, identically distributed (i.i.d.) zero-mean circular Complex Normal (CN) [6] with a covariance matrix $\boldsymbol{R}_n \triangleq \mathbb{E}\left[\boldsymbol{n}[t]\boldsymbol{n}[t]^{\dagger}\right] = \sigma_n^2\boldsymbol{I}_M$, where σ_n^2 is considered unknown; and

(v) $\boldsymbol{x}[t]$ denotes the signal that would have been received in the absence of gain or phase offsets, namely with $\boldsymbol{\Psi} = \boldsymbol{\Phi} = \boldsymbol{I}_M$.

We also assume that the sources may be modeled as mutually uncorrelated random processes. Therefore, in this work, $\boldsymbol{s}[t]$ is considered as an i.i.d. zero-mean circular CN vector process with an unknown diagonal covariance matrix $\boldsymbol{R}_s \triangleq \mathbb{E}\left[\boldsymbol{s}[t]\boldsymbol{s}[t]^{\dagger}\right]$. Furthermore, we assume $\boldsymbol{s}[t]$ and $\boldsymbol{n}[t]$ are also uncorrelated. As a consequence, it follows that

$$\boldsymbol{r}[t] \sim \mathcal{CN}\left(\boldsymbol{0}_M, \boldsymbol{R}\right), \forall t \in \{1, \ldots, T\}, \quad (2)$$

where $\boldsymbol{0}_M \in \mathbb{R}^{M\times 1}$ is the (M-dimensional) all-zeros vector and

$$\boldsymbol{R} \triangleq \mathbb{E}\left[\boldsymbol{r}[t]\boldsymbol{r}[t]^{\dagger}\right] = \boldsymbol{\Psi}\boldsymbol{\Phi}\boldsymbol{C}\boldsymbol{\Phi}^{*}\boldsymbol{\Psi} \in \mathbb{C}^{M\times M}, \quad (3)$$

$$\boldsymbol{C} \triangleq \mathbb{E}\left[\boldsymbol{x}[t]\boldsymbol{x}[t]^{\dagger}\right] = \boldsymbol{A}(\boldsymbol{\theta})\boldsymbol{R}_s\boldsymbol{A}(\boldsymbol{\theta})^{\dagger} + \sigma_n^2\boldsymbol{I}_M \in \mathbb{C}^{M\times M}, \quad (4)$$

where we have used $\boldsymbol{\Psi}^{\dagger} = \boldsymbol{\Psi}$ and $\boldsymbol{\Phi}^{\dagger} = \boldsymbol{\Phi}^{*}$.

The problem at hand can now be formulated as follows. *Given the statistically independent measurements* $\{\boldsymbol{r}[t]\}_{t=1}^{T}$ *whose (identical) distribution is prescribed by* (2)*, estimate the unknown (deterministic) parameters* $\{\boldsymbol{\psi}, \boldsymbol{\phi}\}$.

Notice that in this "blind" setup, for this formulation, $\boldsymbol{\theta}, \sigma_n^2$ and the diagonal elements of $\boldsymbol{R}_s$ are considered as nuisance parameters. However, for other problems described by the same model, the parameters of interest, and accordingly the nuisance parameters, may be defined differently. For example, in the DOAs estimation problem, $\boldsymbol{\theta}$ are the "goal" estimands, whereas $\boldsymbol{\psi}, \boldsymbol{\phi}, \sigma_n^2$ and the diagonal elements of $\boldsymbol{R}_s$ are considered as nuisance parameters. Nevertheless, our goal here is to provide an (asymptotically optimal) estimation scheme for $\boldsymbol{\psi}$ and $\boldsymbol{\phi}$, based on the understanding that the measurements $\{\boldsymbol{x}[t]\}$ of a perfectly calibrated sensor array would be preferable (in terms of the attainable performance) to $\{\boldsymbol{r}[t]\}$ in other estimation problems described by this model.

III. Approximate Optimal Blind Calibration

We begin by recognizing that an (asymptotically) optimal solution to our problem would be obtained by joint ML estimation of $\boldsymbol{\psi}, \boldsymbol{\phi}, \boldsymbol{\theta}, \sigma_n^2$ and the diagonal elements of $\boldsymbol{R}_s$, which (asymptotically) yields efficient estimators ([7]) of $\boldsymbol{\psi}, \boldsymbol{\phi}$. However, since the derivation of the likelihood equations for this model is rather cumbersome, which, at any rate, leads to a highly nonlinear system of equations, and since the sufficient statistic of this model is the sample covariance matrix of the measurements $\widehat{\boldsymbol{R}} \triangleq \frac{1}{T}\sum_{t=1}^{T}\boldsymbol{r}[t]\boldsymbol{r}[t]^{\dagger} \in \mathbb{C}^{M\times M}$, we resort to (approximated) OWLS estimation of $\boldsymbol{\psi}, \boldsymbol{\phi}$ based (only) on $\widehat{\boldsymbol{R}}$.

A. Approximated OWLS Estimation of the Sensor Gains

The proposed estimators we shall present are in fact enhanced versions of the LS estimators proposed by Paulraj and Kailath [1] on the premises of the following observation. Since the array manifold matrix $\boldsymbol{A}(\boldsymbol{\theta})$ is a Vandermonde matrix (e.g., [8]) and all the signals involved are uncorrelated, the covariance matrix of a perfectly calibrated array $\boldsymbol{C}$ is a Toeplitz matrix (e.g., [9]). Therefore, using the fact that $|R_{ij}| = |C_{ij}|\psi_i\psi_j$, we can eliminate the dependence on the unknowns $\{C_{ij}\}$ by observing that

$$\log\left(\frac{|R_{ij}|}{|R_{k\ell}|}\right) = \log(\psi_i) + \log(\psi_j) - \log(\psi_k) - \log(\psi_\ell) \quad (5)$$

for any four indices satisfying $i-j = k-\ell$ (i.e., R_{ij} and $R_{k\ell}$ lie on the same diagonal, and therefore so do $C_{ij} = C_{k\ell}$).

Based on the relation (5), and due to the fact that, in practice, the true covariance matrix $\boldsymbol{R}$ is not available, it was proposed in [1] to use $\widehat{\boldsymbol{R}}$ instead of $\boldsymbol{R}$ and collect all the nonredundant relations for which (i,j) and (k,ℓ) pairs lie on the same main/super diagonals, to get a total of $k_\psi \triangleq \sum_{m=2}^{M} m(m-1)/2$ equations of the form (5). Denoting $\widetilde{\psi}_m \triangleq \log(\psi_m)$, collecting all k_ψ elements $\{\mu_{ijk\ell} \triangleq \log\left(|\widehat{R}_{ij}|/|\widehat{R}_{k\ell}|\right)\}$ (and adding the required reference equation[2]) into the vector $\boldsymbol{\mu} \in \mathbb{R}^{(k_\psi+1)\times 1}$, we get the LS estimate

$$\widehat{\widetilde{\boldsymbol{\psi}}}_{\mathrm{LS}} \triangleq \left(\boldsymbol{H}^{\mathrm{T}}\boldsymbol{H}\right)^{-1}\boldsymbol{H}^{\mathrm{T}}\boldsymbol{\mu} \;\Rightarrow\; \widehat{\boldsymbol{\psi}}_{\mathrm{LS}} \triangleq \exp\left(\widehat{\widetilde{\boldsymbol{\psi}}}_{\mathrm{LS}}\right), \quad (6)$$

which stems from $\boldsymbol{\mu} \approx \boldsymbol{H}\widetilde{\boldsymbol{\psi}}$, where $\boldsymbol{H} \in \mathbb{R}^{(k_\psi+1)\times M}$ is the matrix (consisting only of 0s, $\pm$1s and $\pm$2s) of the corresponding coefficients determined by (5) for different possible values of $\{i,j,k,\ell\}$ (see [1] or [10] for the explicit description of $\boldsymbol{H}$), and $\exp(\cdot)$ operates elementwise.

Indeed, theoretically, $\widehat{\boldsymbol{R}}$ can be made arbitrarily close to $\boldsymbol{R}$ by increasing (appropriately) the sample size T. However, in practice, the available sample size is always limited and is oftentimes fixed. Therefore, rather than relying on the coarse approximation $\widehat{\boldsymbol{R}} \approx \boldsymbol{R}$, which leads to the estimate (6), we propose a more refined analysis, which takes into account the estimation errors and exploits (some of) their (approximated) statistical properties for obtaining a more accurate estimator.

[2]For the exact details the reader is referred to [1] or [10].

More formally, for any finite sample size T, we have

$$\widehat{\boldsymbol{R}} \triangleq \boldsymbol{R}+\boldsymbol{\mathcal{E}} \;\Rightarrow\; \widehat{R}_{ij} = R_{ij}+\mathcal{E}_{ij},\; \forall i,j\in\{1,\ldots,M\}, \quad (7)$$

where $\{\mathcal{E}_{ij}\}$ denote the estimation errors in the estimation of $\{R_{ij}\}$. Hence, rewriting (5) with $\{\widehat{R}_{ij}\}$ replacing $\{R_{ij}\}$ yields

$$\mu_{ijk\ell} = \log\left(\frac{|\widehat{R}_{ij}|}{|\widehat{R}_{k\ell}|}\right) = \log\left(\frac{|R_{ij}+\mathcal{E}_{ij}|}{|R_{k\ell}+\mathcal{E}_{k\ell}|}\right) \quad (8)$$

$$= \log\left(\frac{|R_{ij}|}{|R_{k\ell}|}\right) + \log\left(\frac{|1+\mathcal{E}_{ij}/R_{ij}|}{|1+\mathcal{E}_{k\ell}/R_{k\ell}|}\right) \quad (9)$$

$$\triangleq \widetilde{\psi}_i+\widetilde{\psi}_j-\widetilde{\psi}_k-\widetilde{\psi}_\ell+\varepsilon_{ijk\ell},\;\; \forall i-j=k-\ell, \quad (10)$$

so that we now have the *exact* relation

$$\boldsymbol{\mu} = \boldsymbol{H}\widetilde{\boldsymbol{\psi}}+\boldsymbol{\varepsilon} \in \mathbb{R}^{(k_\psi+1)\times 1}, \quad (11)$$

where $\boldsymbol{\varepsilon}$ is the transformed "measurement noise" in the resulting system of linear equations (11), collecting $\{\varepsilon_{ijk\ell}\}$ with the respective corresponding indices. Now, from the Gauss-Markov theorem [11], the Best Linear Unbiased Estimator (BLUE) of $\widetilde{\boldsymbol{\psi}}$ from $\boldsymbol{\mu}$ is given by the OWLS estimator

$$\widehat{\widetilde{\boldsymbol{\psi}}}_{\text{OWLS}} \triangleq \left(\boldsymbol{H}^{\mathrm{T}}\boldsymbol{\Lambda}_\varepsilon^{-1}\boldsymbol{H}\right)^{-1}\boldsymbol{H}^{\mathrm{T}}\boldsymbol{\Lambda}_\varepsilon^{-1}\left(\boldsymbol{\mu}-\boldsymbol{\eta}_\varepsilon\right), \quad (12)$$

where $\boldsymbol{\eta}_\varepsilon \triangleq \mathbb{E}[\boldsymbol{\varepsilon}]$ and $\boldsymbol{\Lambda}_\varepsilon \triangleq \mathbb{E}\left[(\boldsymbol{\varepsilon}-\boldsymbol{\eta}_\varepsilon)(\boldsymbol{\varepsilon}-\boldsymbol{\eta}_\varepsilon)^{\mathrm{T}}\right]$ are the mean and covariance matrix of $\boldsymbol{\varepsilon}$, resp. The BLUE attains the minimal attainable MSE matrix out of all linear unbiased estimators.

Thus, our goal now is to obtain closed-form expressions (even if approximated ones) for $\boldsymbol{\eta}_\varepsilon$ and $\boldsymbol{\Lambda}_\varepsilon$ in terms of the available or estimable quantities, in order to eventually obtain the estimator (12), or at least a well-approximated version thereof. To this end, assume T is sufficiently large such that $|\mathcal{E}_{ij}| \ll |R_{ij}|$ for all possible (i,j). With this, using $\log(|z|)=\Re\{\log(z)\}$ which holds for any $z\in\mathbb{C}$, and the first-order Taylor expansion approximation $\log(1+z)\approx z$ (which holds for all $z\in\mathbb{C}$ satisfying $|z|\ll 1$), the equivalent "measurement noise" $\varepsilon_{ijk\ell}$ reads

$$\varepsilon_{ijk\ell} = \log\left(|1+\mathcal{E}_{ij}/R_{ij}|\right)-\log\left(|1+\mathcal{E}_{k\ell}/R_{k\ell}|\right)$$
$$\approx \Re\left\{\mathcal{E}_{ij}/R_{ij}-\mathcal{E}_{k\ell}/R_{k\ell}\right\}, \forall i,j,k,\ell\in\{1,\ldots,M\}. \quad (13)$$

Clearly, since $\widehat{\boldsymbol{R}}$ is unbiased, it follows that

$$\mathbb{E}[\mathcal{E}_{ij}]=0 \;\Rightarrow\; \mathbb{E}[\varepsilon_{ijk\ell}]\approx 0,\;\; \forall i,j,k,\ell\in\{1,\ldots,M\}, \quad (14)$$

so that $\boldsymbol{\eta}_\varepsilon=\mathbb{E}[\boldsymbol{\varepsilon}]\approx \boldsymbol{0}_{(k_\psi+1)}$. As for the covariance matrix of $\boldsymbol{\varepsilon}$, which now reads $\boldsymbol{\Lambda}_\varepsilon\approx\mathbb{E}\left[\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}^{\mathrm{T}}\right]$, based on the assumption (2) that $\{\boldsymbol{r}[t]\}$ are all circular CN, and in particular using Isserlis' theorem [12], we show in Appendix A that the elements of $\boldsymbol{\Lambda}_\varepsilon$ are (approximately) given by

$$\mathbb{E}\left[\varepsilon_{i_1j_1k_1\ell_1}\cdot\varepsilon_{i_2j_2k_2\ell_2}\right]\approx\frac{1}{T}\cdot 0.5\cdot$$
$$\Re\left\{\frac{R_{i_1j_2}R^*_{j_1i_2}}{R_{i_1j_1}R_{i_2j_2}}-\frac{R_{i_1\ell_2}R^*_{j_1k_2}}{R_{i_1j_1}R_{k_2\ell_2}}-\frac{R_{k_1j_2}R^*_{\ell_1i_2}}{R_{k_1\ell_1}R_{i_2j_2}}+\frac{R_{k_1\ell_2}R^*_{\ell_1k_2}}{R_{k_1\ell_1}R_{k_2\ell_2}}\right.$$
$$\left.+\frac{R_{i_1i_2}R^*_{j_1j_2}}{R_{i_1j_1}R^*_{i_2j_2}}-\frac{R_{i_1k_2}R^*_{j_1\ell_2}}{R_{i_1j_1}R^*_{k_2\ell_2}}-\frac{R_{k_1i_2}R^*_{\ell_1j_2}}{R_{k_1\ell_1}R^*_{i_2j_2}}+\frac{R_{k_1k_2}R^*_{\ell_1\ell_2}}{R_{k_1\ell_1}R^*_{k_2\ell_2}}\right\}, \quad (15)$$

so that $\boldsymbol{\Lambda}_\varepsilon$ is (approximately) a function of $\boldsymbol{R}$ *only*.

Of course, the true $\boldsymbol{R}$ is in fact unknown. However, since $\widehat{\boldsymbol{R}}$ is the MLE of $\boldsymbol{R}$, by virtue of the invariance property of the MLE [13], it follows that $\widehat{\boldsymbol{\Lambda}}_\varepsilon$, a matrix whose elements are defined as in (15), but with $\{\widehat{R}_{ij}\}$ replacing $\{R_{ij}\}$, is (approximately) the MLE of $\boldsymbol{\Lambda}_\varepsilon$. Therefore, we propose the following "ML-based OWLS" estimate of the sensor gains

$$\widehat{\widetilde{\boldsymbol{\psi}}}_{\text{ML-OWLS}} \triangleq \left(\boldsymbol{H}^{\mathrm{T}}\widehat{\boldsymbol{\Lambda}}_\varepsilon^{-1}\boldsymbol{H}\right)^{-1}\boldsymbol{H}^{\mathrm{T}}\widehat{\boldsymbol{\Lambda}}_\varepsilon^{-1}\boldsymbol{\mu} \quad (16)$$

$$\Rightarrow\; \widehat{\boldsymbol{\psi}}_{\text{ML-OWLS}} \triangleq \exp\left(\widehat{\widetilde{\boldsymbol{\psi}}}_{\text{ML-OWLS}}\right). \quad (17)$$

Note that for a sufficiently large T: $\widehat{\boldsymbol{\psi}}_{\text{ML-OWLS}}\approx\widehat{\boldsymbol{\psi}}_{\text{OWLS}}$ by virtue of the continuous mapping theorem [14] and the consistency of the MLE ([15]) $\widehat{\boldsymbol{\Lambda}}_\varepsilon$.

Furthermore, notice that since the ML estimation errors $\{\mathcal{E}_{ij}\}$ are asymptotically (non-circular) jointly CN, according to (13), the transformed estimation errors $\{\varepsilon_{ijk\ell}\}$ are approximately asymptotically jointly Normal. Thus, it follows that $\widehat{\widetilde{\boldsymbol{\psi}}}_{\text{OWLS}}$ is (approximately) also the MLE of $\widetilde{\boldsymbol{\psi}}$ based on $\boldsymbol{\mu}$. Once again, using the invariance property of the MLE, it follows that $\widehat{\boldsymbol{\psi}}_{\text{OWLS}}$ is the MLE of $\boldsymbol{\psi}$ based on $\boldsymbol{\mu}$. Therefore, we conclude that $\widehat{\boldsymbol{\psi}}_{\text{ML-OWLS}}$ is asymptotically the MLE of $\boldsymbol{\psi}$ based on $\boldsymbol{\mu}$ (but not based on the sufficient statistic $\widehat{\boldsymbol{R}}$, and thus not on the raw data $\{\boldsymbol{r}[t]\}_{t=1}^T$, since $\boldsymbol{\mu}$ is not an invertible function of $\widehat{\boldsymbol{R}}$).

B. *Approximated OWLS Estimation of the Sensor Phases*

Following Paulraj and Kailath [1], from (3) we have

$$\angle R_{ij}-\angle R_{k\ell}=\phi_i-\phi_j-\phi_k+\phi_\ell,\;\; \forall i-j=k-\ell. \quad (18)$$

Based on the same logic presented in the previous subsection, taking all the nonredundant relations from the super diagonals (excepting the singleton R_{1M}), we get a total of $k_\phi \triangleq \sum_{m=2}^{M-1}m(m-1)/2$ equations of the form (18) (the diagonal is excluded since it is real-valued and yields trivial equations in (18)). Collecting all such k_ϕ elements $\{\nu_{ijk\ell}\triangleq\angle\widehat{R}_{ij}-\angle\widehat{R}_{k\ell}\}$ into the vector $\boldsymbol{\nu}\in\mathbb{R}^{(k_\phi+2)\times 1}$ (and adding the required reference equations[3]), we get the LS estimate

$$\boldsymbol{\nu}\approx\boldsymbol{G}\boldsymbol{\phi}\;\Rightarrow\;\widehat{\boldsymbol{\phi}}_{\text{LS}}\triangleq\left(\boldsymbol{G}^{\mathrm{T}}\boldsymbol{G}\right)^{-1}\boldsymbol{G}^{\mathrm{T}}\boldsymbol{\nu}, \quad (19)$$

where $\boldsymbol{G}\in\mathbb{R}^{(k_\phi+2)\times M}$ is the coefficients matrix (consisting only of 0s, ± 1s and -2s) as prescribed by (18) for the different possible values of $\{i,j,k,\ell\}$ (see [1] or [10] for the explicit description of $\boldsymbol{G}$).

Fortunately, since $\angle z=\Im\{\log(z)\}$ for any $z\in\mathbb{C}$, writing (18) in terms of $\{\widehat{R}_{ij}\}$ according to (7) yields

$$\nu_{ijk\ell}=\angle\widehat{R}_{ij}-\angle\widehat{R}_{k\ell}=\Im\left\{\log\left(\frac{R_{ij}+\mathcal{E}_{ij}}{R_{k\ell}+\mathcal{E}_{k\ell}}\right)\right\} \quad (20)$$

$$=\underbrace{\Im\left\{\log\left(\frac{R_{ij}}{R_{k\ell}}\right)\right\}}_{\angle R_{ij}-\angle R_{k\ell}}+\Im\left\{\log\left(\frac{1+\mathcal{E}_{ij}/R_{ij}}{1+\mathcal{E}_{k\ell}/R_{k\ell}}\right)\right\} \quad (21)$$

$$\triangleq\phi_i-\phi_j-\phi_k+\phi_\ell+\epsilon_{ijk\ell},\;\; \forall i-j=k-\ell, \quad (22)$$

[3] For the exact details the reader is referred to [1] or [10].

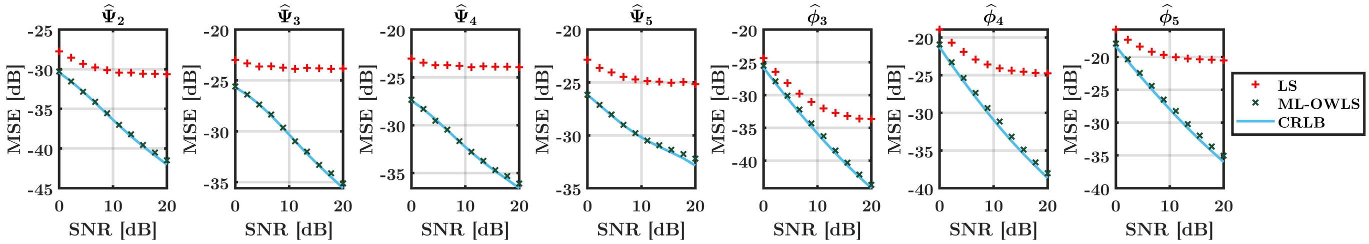

Fig. 1: MSE vs. SNR for $T = 500$. Empirical results were obtained by averaging 10^4 independent trials. Evidently, the improvement w.r.t. the "naïve" equally-weighted LS approach can reach more than an order of magnitude in the high SNR regime.

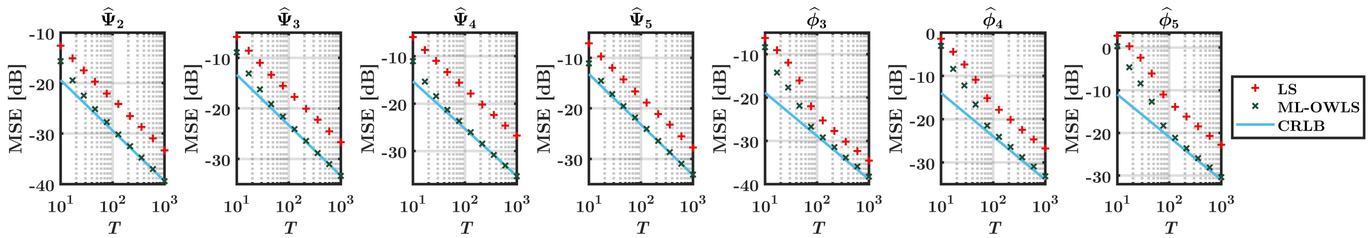

Fig. 2: MSE vs. T for SNR= 10[dB]. Empirical results were obtained by averaging 10^4 independent trials. As seen, the achieved gain by the proposed estimates w.r.t. ordinary LS is substantial even in the large sample size regime.

so that we now have the *exact* relation

$$\boldsymbol{\nu} = \boldsymbol{G}\boldsymbol{\phi} + \boldsymbol{\epsilon} \in \mathbb{R}^{(k_\phi+2)\times 1}, \tag{23}$$

where $\boldsymbol{\epsilon}$ is the transformed "measurement noise" in the resulting system of linear equations (23), collecting $\{\epsilon_{ijk\ell}\}$ with the respective corresponding indices. Notice that if we define

$$\xi_{ijk\ell} \triangleq \log\left(\frac{1+\mathcal{E}_{ij}/R_{ij}}{1+\mathcal{E}_{k\ell}/R_{k\ell}}\right), \quad \forall i-j=k-\ell, \tag{24}$$

then we also have $\xi_{ijk\ell} = \varepsilon_{ijk\ell} + \jmath \cdot \epsilon_{ijk\ell}$. Hence, based on the arguments stated before (13), we have

$$\epsilon_{ijk\ell} \approx \Im\left\{\mathcal{E}_{ij}/R_{ij} - \mathcal{E}_{k\ell}/R_{k\ell}\right\}, \forall i,j,k,\ell \in \{1,\dots,M\}. \tag{25}$$

From (25), it follows immediately that $\boldsymbol{\eta}_\epsilon \triangleq \mathbb{E}[\boldsymbol{\epsilon}] \approx \mathbf{0}_{(k_\phi+2)}$. Additionally, as shown in Appendix A, using the expressions already obtained for the previous computation of $\boldsymbol{\Lambda}_\varepsilon$, we easily obtain the elements of $\boldsymbol{\Lambda}_\epsilon \triangleq \mathbb{E}\left[(\boldsymbol{\epsilon}-\boldsymbol{\eta}_\epsilon)(\boldsymbol{\epsilon}-\boldsymbol{\eta}_\epsilon)^{\mathrm{T}}\right]$, which now reads $\boldsymbol{\Lambda}_\epsilon \approx \mathbb{E}\left[\boldsymbol{\epsilon}\boldsymbol{\epsilon}^{\mathrm{T}}\right]$, given (approximately) by

$$\begin{aligned}&\mathbb{E}\left[\epsilon_{i_1j_1k_1\ell_1}\cdot\epsilon_{i_2j_2k_2\ell_2}\right] \approx \frac{1}{T}\cdot 0.5\cdot\\ &\Re\left\{\frac{R_{i_1j_2}R^*_{j_1i_2}}{R_{i_1j_1}R_{i_2j_2}} - \frac{R_{i_1\ell_2}R^*_{j_1k_2}}{R_{i_1j_1}R_{k_2\ell_2}} - \frac{R_{k_1j_2}R^*_{\ell_1i_2}}{R_{k_1\ell_1}R_{i_2j_2}} + \frac{R_{k_1\ell_2}R^*_{\ell_1k_2}}{R_{k_1\ell_1}R_{k_2\ell_2}}\right.\\ &\left.-\frac{R_{i_1i_2}R^*_{j_1j_2}}{R_{i_1j_1}R^*_{i_2j_2}} + \frac{R_{i_1k_2}R^*_{j_1\ell_2}}{R_{i_1j_1}R^*_{k_2\ell_2}} + \frac{R_{k_1i_2}R^*_{\ell_1j_2}}{R_{k_1\ell_1}R^*_{i_2j_2}} - \frac{R_{k_1k_2}R^*_{\ell_1\ell_2}}{R_{k_1\ell_1}R^*_{k_2\ell_2}}\right\},\end{aligned} \tag{26}$$

so that $\boldsymbol{\Lambda}_\epsilon$ is also (approximately) a function of $\boldsymbol{R}$ *only*.

Recalling (again) that $\boldsymbol{R}$ is unknown, we propose the following ML-based OWLS estimate of the sensor phases

$$\widehat{\boldsymbol{\phi}}_{\text{ML-OWLS}} \triangleq \left(\boldsymbol{G}^{\mathrm{T}}\widehat{\boldsymbol{\Lambda}}_\epsilon^{-1}\boldsymbol{G}\right)^{-1}\boldsymbol{G}^{\mathrm{T}}\widehat{\boldsymbol{\Lambda}}_\epsilon^{-1}\boldsymbol{\nu}, \tag{27}$$

where $\widehat{\boldsymbol{\Lambda}}_\epsilon$ is a matrix whose elements are defined as in (26), but with $\{\widehat{R}_{ij}\}$ replacing $\{R_{ij}\}$. Based on the same arguments presented in the end of the preceding subsection, we conclude that for a sufficiently large T, the proposed estimator (27) approximately coincides with the MLE of $\boldsymbol{\phi}$ based on $\boldsymbol{\nu}$ (but not based on $\widehat{\boldsymbol{R}}$, and accordingly not on the raw data $\{\boldsymbol{r}[t]\}_{t=1}^T$, since $\boldsymbol{\nu}$ is not an invertible function of $\widehat{\boldsymbol{R}}$).

IV. Simulation Results

We consider model (1) in a scenario which consists of a $M = 5$ elements array with half wavelength inter-element spacing (i.e., $\gamma = \lambda/2$), and $D = 3$ equal power, zero-mean unit variance sources arriving at angles $\boldsymbol{\theta} = -[35^\circ\ 63^\circ\ 25^\circ]^{\mathrm{T}}$. The sensors' gains and phases were set to $\boldsymbol{\psi} = [1\ 1.3\ 2\ 1.7\ 2.1]^{\mathrm{T}}$ and $\boldsymbol{\phi} = [0^\circ\ 0^\circ\ 5^\circ\ 10^\circ\ -7^\circ]^{\mathrm{T}}$, resp., where w.l.o.g. we assume ψ_1, ϕ_1 and ϕ_2 are known (and serve as references). Empirical results were obtained by averaging 10^4 independent trials.

Fig. 1 presents the MSEs obtained by $\widehat{\boldsymbol{\psi}}_{\text{ML-OWLS}}$ and $\widehat{\boldsymbol{\phi}}_{\text{ML-OWLS}}$ vs. the SNR, where $T = 500$ is fixed. For comparison, we also show the MSEs obtained by Paulraj and Kailath's LS estimators, (6) and (19), and the Cramér-Rao Lower Bound (CRLB) on the corresponding MSEs obtained in any unbiased joint estimation of all the unknown parameters. Similarly, Fig. 2 presents the same quantities, however now vs. the sample size T, where the SNR is fixed at 10[dB]. As seen, the proposed estimates exhibit nearly optimal performance, approximately attaining the CRLB, i.e., the asymptotic performance of the MLE based on the raw data. Notice that although this (near) optimality is theoretically obtained only asymptotically, in practice, this asymptotic state may be reached within (only) a few dozens of samples. Moreover, the improvement w.r.t. ordinary LS estimation can reach more than an order of magnitude in the high SNR regime.

V. Conclusion and Future Work

In the context of array processing of narrowband signals, we presented an approximately optimal blind calibration technique. Based on the Toeplitz structure of the covariance matrix of the perfectly calibrated array, and using a first-order approximation for the computation of the transformed "measurement noise" covariance matrix, we derived the ML-OWLS estimators of the sensors' gains and phases. These estimators were shown empirically to be nearly optimal asymptotically, approximately attaining the CRLB.

We note that $\widehat{\boldsymbol{R}}$ is consistent even when the vector samples $\{\boldsymbol{r}[t]\}_{t=1}^{T}$ are not CN, and by virtue of the central limit theorem [16], all elements of $\widehat{\boldsymbol{R}}$ are jointly CN, asymptotically. Therefore, the derived estimators (16) and (27) are still the approximate BLUEs of $\boldsymbol{\psi}$ and $\boldsymbol{\phi}$, resp., based on $\boldsymbol{\mu}$ and $\boldsymbol{\nu}$ (but not necessarily based on $\widehat{\boldsymbol{R}}$). However, it can be shown that the resulting "measurements noise" vectors $\boldsymbol{\varepsilon}$ and $\boldsymbol{\epsilon}$ are correlated, which means that the proposed estimation scheme may be further improved, exploiting this statistical coupling. Moreover, with certain modifications in their definitions, $\boldsymbol{\mu}$ and $\boldsymbol{\nu}$ may serve together as an invertible function of the sufficient statistic $\widehat{\boldsymbol{R}}$. Accordingly, the resulting estimators, which are modified version of (16) and (27), are approximately the MLEs w.r.t. $\widehat{\boldsymbol{R}}$, hence they are also approximately the MLEs based on the raw data $\{\boldsymbol{r}[t]\}_{t=1}^{T}$. On top of that, considering the above, these modified estimators also retain their asymptotic optimality, w.r.t. the data $\widehat{\boldsymbol{R}}$, for a more general model, and are applicable for various types of signal models. An elaborated discussion on this issue (and more) will be presented in a future paper in which we consider an extended, more general model.

Appendix A
Computation of the Noise Covariance Matrices

As seen from (13) and (26), $\varepsilon_{ijk\ell}$ and $\epsilon_{ijk\ell}$ are the real and imaginary parts, resp., of the same complex number (24), $\xi_{ijk\ell}$. Thus, define

$$z_{ijk\ell} \triangleq \mathcal{E}_{ij}/R_{ij} - \mathcal{E}_{k\ell}/R_{k\ell}, \;\; \forall i,j,k,\ell \in \{1,\ldots,M\}, \quad (28)$$

and for simplicity denote $z_1 := z_{i_1 j_1 k_1 \ell_1}$ and $z_2 := z_{i_2 j_2 k_2 \ell_2}$. Starting with $\boldsymbol{\Lambda}_{\varepsilon}$, it may be easily shown that

$$\mathbb{E}\left[\Re\{z_1\}\Re\{z_2\}\right] = 0.5 \cdot \Re\left\{\mathbb{E}\left[z_1 z_2^*\right] + \mathbb{E}\left[z_1 z_2\right]\right\}, \quad (29)$$

so we may concentrate on the computation of $\mathbb{E}\left[z_1 z_2^*\right]$ and $\mathbb{E}\left[z_1 z_2\right]$. It follows from (28) that both of these expectations are easily computed given the covariances and pseudo-covariances of $\{\mathcal{E}_{ij}\}$. Thus,

$$\mathbb{E}\left[\mathcal{E}_{ij}\mathcal{E}_{k\ell}^*\right] = \mathbb{E}\left[\widehat{R}_{ij}\widehat{R}_{k\ell}^*\right] - R_{ij}R_{k\ell}^*$$
$$= \frac{1}{T^2}\sum_{t_1,t_2=1}^{T}\mathbb{E}\left[r_i[t_1]r_j^*[t_1]r_k^*[t_2]r_\ell[t_2]\right] - R_{ij}R_{k\ell}^*. \quad (30)$$

Using the fact the $\{\boldsymbol{r}[t]\}_{t=1}^{T}$ are all i.i.d. circular CN, applying Isserlis' theorem [12] gives

$$\mathbb{E}\left[r_i[t_1]r_j^*[t_1]r_k^*[t_2]r_\ell[t_2]\right] = \underbrace{\mathbb{E}\left[r_i[t_1]r_j^*[t_1]\right]}_{=R_{ij}}\underbrace{\mathbb{E}\left[r_k^*[t_2]r_\ell[t_2]\right]}_{=R_{k\ell}^*} +$$
$$\underbrace{\mathbb{E}\left[r_i[t_1]r_k^*[t_2]\right]}_{=\delta_{t_1t_2}\cdot R_{ik}}\underbrace{\mathbb{E}\left[r_j^*[t_1]r_\ell[t_2]\right]}_{=\delta_{t_1t_2}\cdot R_{j\ell}^*} + \underbrace{\mathbb{E}\left[r_i[t_1]r_\ell[t_2]\right]}_{=0\text{ (circularity)}}\underbrace{\mathbb{E}\left[r_j^*[t_1]r_k^*[t_2]\right]}_{=0\text{ (circularity)}}, \quad (31)$$

where $\delta_{t_1t_2}$ denotes the Kronecker delta of $t_1, t_2 \in \mathbb{Z}$. Substituting (31) into (30), and repeating for $\mathbb{E}\left[\mathcal{E}_{ij}\mathcal{E}_{k\ell}\right]$ with exactly the same technique, we obtain after simplification

$$\mathbb{E}\left[\mathcal{E}_{ij}\mathcal{E}_{k\ell}^*\right] = \frac{1}{T}R_{ik}R_{j\ell}^*, \quad \mathbb{E}\left[\mathcal{E}_{ij}\mathcal{E}_{k\ell}\right] = \frac{1}{T}R_{i\ell}R_{jk}^*, \quad (32)$$

for all $i,j,k,\ell \in \{1,\ldots,M\}$. Now, substituting (28) in $\mathbb{E}\left[z_1 z_2^*\right]$ and $\mathbb{E}\left[z_1 z_2\right]$, and using linearity, with the covariances and pseudo-covariances (32) we obtain (15). Similarly, since

$$\mathbb{E}\left[\Im\{z_1\}\Im\{z_2\}\right] = 0.5 \cdot \Re\left\{\mathbb{E}\left[z_1 z_2^*\right] - \mathbb{E}\left[z_1 z_2\right]\right\}, \quad (33)$$

(26) is given by exactly the same expressions already obtained.

References

[1] A. Paulraj and T. Kailath, "Direction of arrival estimation by eigenstructure methods with unknown sensor gain and phase," in *Proc. of ICASSP*, vol. 10, 1985, pp. 640–643.

[2] 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.

[3] B. C. Ng and C. M. S. See, "Sensor-array calibration using a maximum-likelihood approach," *IEEE Trans. on Antennas and Propagation*, vol. 44, no. 6, pp. 827–835, 1996.

[4] 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.

[5] S. J. Wijnholds and P. Noorishad, "Statistically optimal self-calibration of regular imaging arrays," in *Signal Processing Conference (EUSIPCO), 2012 Proceedings of the 20th European*, 2012, pp. 1304–1308.

[6] B. Loesch and B. Yang, "Cramér-Rao bound for circular and noncircular complex independent component analysis," *IEEE Trans. on Signal Processing*, vol. 61, no. 2, pp. 365–379, 2013.

[7] R. A. Fisher, "On the mathematical foundations of theoretical statistics," *Phil. Trans. R. Soc. Lond. A*, vol. 222, no. 594-604, pp. 309–368, 1922.

[8] K. B. Petersen, M. S. Pedersen *et al.*, "The matrix cookbook," *Technical University of Denmark*, vol. 7, no. 15, p. 510, 2008.

[9] R. M. Gray, "Toeplitz and circulant matrices: A review," *Foundations and Trends® in Communications and Information Theory*, vol. 2, no. 3, pp. 155–239, 2006.

[10] K. N. Ramamohan, S. P. Chepuri, D. F. Comesana, G. C. Pousa, and G. Leus, "Blind calibration for acoustic vector sensor arrays," in *Proc. of ICASSP*, 2018, pp. 3544–3548.

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

[12] L. Isserlis, "On a formula for the product-moment coefficient of any order of a normal frequency distribution in any number of variables," *Biometrika*, vol. 12, no. 1/2, pp. 134–139, 1918.

[13] N. Mukhopadhyay, *Probability and statistical inference*. CRC Press, 2000.

[14] H. B. Mann and A. Wald, "On stochastic limit and order relationships," *The Annals of Mathematical Statistics*, vol. 14, no. 3, pp. 217–226, 1943.

[15] H. Cramér, *Mathematical methods of statistics (PMS-9)*. Princeton university press, 2016, vol. 9.

[16] S. Ross, *A First Course in Probability 8th Edition*. Pearson, 2009.