## ENHANCED BLIND CALIBRATION OF UNIFORM LINEAR ARRAYS WITH ONE-BIT **OUANTIZATION BY KULLBACK-LEIBLER DIVERGENCE COVARIANCE FITTING**

Amir Weiss\* and Arie Yeredor<sup>†</sup>

\*Dept. of Computer Science and Applied Mathematics <sup>†</sup>School of Electrical Engineering Weizmann Institute of Science amir.weiss@weizmann.ac.il

# Tel-Aviv University arie@eng.tau.ac.il

## ABSTRACT

One-bit quantization has recently become an attractive option for data acquisition in cutting edge applications, due to the increasing demand for low power and higher sampling rates. Subsequently, the rejuvenated one-bit array processing field is now receiving more attention, as "classical" array processing techniques are adapted / modified accordingly. However, array calibration, often an instrumental preliminary stage in array processing, has so far received little attention in its one-bit form. In this paper, we present a novel solution approach for the *blind* calibration problem, namely, without using known calibration signals. In order to extract information within the second-order statistics of the quantized measurements, we propose to estimate the unknown sensors' gains and phases offsets according to a Kullback-Leibler Divergence (KLD) covariance fitting criterion. We then provide a quasi-Newton solution algorithm, with a consistent initial estimate, and demonstrate the improved accuracy of our KLD-based estimates in simulations.

Index Terms- Blind calibration, uniform linear arrays, one-bit quantization, Kullback-Leibler divergence.

## 1. INTRODUCTION

While high-resolution quantization of analog signals is obviously preferable to low-resolution in terms of accuracy (for signal recovery, feature extraction, etc.), other, practical considerations are often at a higher priority. For example, the power consumption of an Analog-to-Digital Converter (ADC) increases exponentially with the number of bits [1]. Other important considerations are the high cost of high-resolution devices [2], and their operational sampling rates, which are increasingly becoming insufficient for various applications, such as cognitive radio and radar [3–5].

As a remedy, one-bit quantization has been gradually receiving more attention in recent years [6-11]. Being essentially the most basic form of quantization, it offers the advantages of low cost and low complexity implementation, which allows for higher sampling rates. These benefits come at the cost of greater loss of information per sample, thus giving rise to the necessity of new, 1-bit-adapted digital signal processing algorithms.

In the context of passive array processing, one important problem, which nevertheless has so far received little attention in the literature in its 1-bit form [12], is calibration of the array. This inevitable problem is due to the fact that, in practice, the array sensors generally have different gain and phase responses for various possible reasons (e.g., imperfect manufacturing). Without proper compensation of these relative gain and phase offsets, the accuracy of most (if not all) array-based processing estimation tasks could severely deteriorate. This specific problem is the focus of this work.

While calibration using a known, user-controlled signal is sometimes a viable option, *blind* calibration is typically more desirable, yet more challenging. In this paper, we consider this challenging problem for Uniform Linear Arrays (ULAs), based on one-bit measurements of narrowband Gaussian signals. We propose a statistically enhanced blind calibration scheme, which exploits "hidden" Second-Order Statistics (SOS) information via the Kullback-Leibler Divergence (KLD) covariance fitting criterion. The resulting enhanced calibration leads to considerably higher accuracy in subsequent estimation tasks, as demonstrated in a simulation experiment.

#### 2. PROBLEM FORMULATION

Consider a ULA of N sensors, each with *unknown* deterministic gain and phase offsets, denoted  $\psi \in \mathbb{R}^{N \times 1}_+$  and  $\phi \in [-\pi, \pi)^{N \times 1}$ , where  $\psi_n$  and  $\phi_n$  are the unknown gain and phase offsets of the *n*-th sensor, resp. Further, consider the presence of an unknown number Mof unknown, "far field" narrowband sources, centered around some common carrier frequency with wavelength  $\lambda$ .

The noisy vector of sampled, baseband-converted signals from all N sensors is given by

$$\boldsymbol{r}(t) = \boldsymbol{\Psi} \boldsymbol{\Phi} \left( \boldsymbol{A}(\boldsymbol{\alpha}) \boldsymbol{s}(t) + \boldsymbol{v}(t) \right) + \boldsymbol{w}(t) \triangleq \boldsymbol{\Psi} \boldsymbol{\Phi} \boldsymbol{x}(t) + \boldsymbol{w}(t) \in \mathbb{C}^{N \times 1},$$
(1)

for all  $t \in [T]$ , where  $[N] \triangleq \{1, \ldots, N\}$  for any  $N \in \mathbb{N}$ , and

(i) 
$$\Psi \triangleq \operatorname{Diag}(\psi) \in \mathbb{R}^{N \times N}_{+}, \Phi \triangleq \operatorname{Diag}\left(e^{j\phi}\right) \in \mathbb{C}^{N \times N};$$

- (ii)  $\boldsymbol{s}(t) \triangleq [s_1(t) \cdots s_M(t)]^{\mathrm{T}} \in \mathbb{C}^{M \times 1}$  denotes the sources, emitted from unknown azimuth angles  $\boldsymbol{\alpha} \triangleq [\alpha_1 \cdots \alpha_M]^{\mathrm{T}} \in$  $\mathbb{R}^{M \times 1}$ :
- (iii)  $\boldsymbol{A}(\boldsymbol{\alpha}) \triangleq [\boldsymbol{a}(\alpha_1) \cdots \boldsymbol{a}(\alpha_M)] \in \mathbb{C}^{N \times M}$  denotes the nominal array manifold matrix, with the steering vectors  $\boldsymbol{a}(\alpha_m) \triangleq \left[1 e^{j\frac{2\pi}{\lambda}d\cos(\alpha_m)} \cdots e^{j\frac{2\pi}{\lambda}(N-1)d\cos(\alpha_m)}\right]^{\mathrm{T}} \in \mathbb{C}^{N \times 1}$  as its columns, where  $d \in \mathbb{R}_+$  is the inter-element spacing;
- (iv)  $v(t) \in \mathbb{C}^{N \times 1}$  denotes additive, ambient noise or "interfering" signals, modeled as spatially and temporally independent, identically distributed (i.i.d.) zero-mean circular Complex Normal (CN) [13] with a covariance matrix  $\boldsymbol{R}_{v} \triangleq \mathbb{E}\left[\boldsymbol{v}(t)\boldsymbol{v}(t)^{\mathrm{H}}\right] = \sigma_{v}^{2}\boldsymbol{I}_{N}$ , where  $\sigma_{v}^{2}$  is unknown;
- (v)  $\boldsymbol{w}(t) \in \mathbb{C}^{N \times 1}$  denotes additive internal (e.g., thermal) receiver noise, unaffected by the gain and phase offsets, modeled as spatially and temporally i.i.d. zero-mean circular CN with a covariance matrix  $\mathbf{R}_w \triangleq \mathbb{E}\left[\boldsymbol{w}(t)\boldsymbol{w}(t)^{\mathrm{H}}\right] = \sigma_w^2 \boldsymbol{I}_N$ , where  $\sigma_w^2$  is known; and
- (vi)  $\boldsymbol{x}(t)$  denotes the signal that would have been received in the absence of gain or phase offsets and internal noise, namely with  $\Psi = \Phi = I_N$  and  $\sigma_w^2 = 0$ .

We also assume that the sources are mutually uncorrelated. In particular, we assume that s(t) is a (temporally) i.i.d. zero-mean circular CN vector process with an unknown diagonal covariance matrix  $\mathbf{R}^s \triangleq \mathbb{E}[s(t)s(t)^{\mathrm{H}}]$ . In addition, we assume s(t), v(t) and w(t) are all mutually statistically independent. As a result we get

$$\boldsymbol{r}(t) \sim \mathcal{CN}\left(\boldsymbol{0}_{N}, \boldsymbol{R}\right), \forall t \in [T],$$
 (2)

where

$$\boldsymbol{R} \triangleq \mathbb{E}\left[\boldsymbol{r}(t)\boldsymbol{r}(t)^{\mathrm{H}}\right] = \boldsymbol{\Psi}\boldsymbol{\Phi}\boldsymbol{C}\boldsymbol{\Phi}^{*}\boldsymbol{\Psi} + \sigma_{w}^{2}\boldsymbol{I}_{N} \in \mathbb{C}^{N \times N}, \qquad (3)$$

$$\boldsymbol{C} \triangleq \mathbb{E}\left[\boldsymbol{x}(t)\boldsymbol{x}(t)^{\mathrm{H}}\right] = \boldsymbol{A}(\boldsymbol{\alpha})\boldsymbol{R}^{s}\boldsymbol{A}^{\mathrm{H}}(\boldsymbol{\alpha}) + \sigma_{v}^{2}\boldsymbol{I}_{N} \in \mathbb{C}^{N \times N}, \quad (4)$$

and we have used  $\Psi^{\mathrm{H}} = \Psi$  and  $\Phi^{\mathrm{H}} = \Phi^{*}$ .

In this work, rather than assuming access to the discrete-time signal (1) measured by an ideal receiver with an  $\infty$ -bit ADC, we assume access only to a "coarse" quantized version thereof, obtained by a 1-bit ADC low-complexity receiver. Specifically, the 1-bit quantized vector of signals from all N sensors at time t is given by

$$\boldsymbol{y}(t) \triangleq \mathcal{Q}(\boldsymbol{r}(t)) \in \left\{ e^{\jmath\left(\frac{\pi}{4} + \frac{\pi k}{2}\right)} : 0 \le k \le 3 \right\}^{N \times 1}, \quad (5)$$

where the complex-valued 1-bit quantization operator is defined as

$$\mathcal{Q}(z) \triangleq \frac{1}{\sqrt{2}} \cdot \Big[ \operatorname{sign} \left( \Re\{z\} \right) + \jmath \cdot \operatorname{sign} \left( \Im\{z\} \right) \Big], \forall z \in \mathbb{C}$$

and, with slight abuse of notations,  $\mathcal{Q}(\cdot)$  operates elementwise in (5).

The one-bit blind calibration problem can now be formulated as follows. Given the quantized measurements  $\{\boldsymbol{y}(t)\}_{t=1}^{T}$  as in (5), estimate the unknown gain and phase offsets  $\{\boldsymbol{\psi}, \boldsymbol{\phi}\}$ .

## 3. THE ARCSINE LAW AND THE IMPLIED SOS-IDENTIFIABILITY

Before we present our proposed solution approach, it is instructive to study the SOS of the one-bit measurements (5), and the implied identifiability conditions. To this end, following [9,12,14], which all rely on the extension [15] of Van Vleck and Middleton's pioneering work [16], due to (2), by the *arcsine law* we have

$$\boldsymbol{R}^{y} \triangleq \mathbb{E}\left[\boldsymbol{y}(t)\boldsymbol{y}(t)^{\mathrm{H}}\right] = \frac{2}{\pi}\mathrm{sine}^{-1}\left(\overline{\boldsymbol{R}}\right) \in \mathbb{C}^{N \times N},$$
 (6)

where  $\overline{R}$  is the normalized covariance matrix of r(t), defined as

$$\overline{\mathbf{R}} \triangleq \mathrm{Ddiag}(\mathbf{R})^{-\frac{1}{2}} \mathbf{R} \mathrm{Ddiag}(\mathbf{R})^{-\frac{1}{2}} \triangleq \mathbf{D}^{-\frac{1}{2}} \mathbf{R} \mathbf{D}^{-\frac{1}{2}},$$
 (7)

where  $\text{Ddiag}(\mathbf{R})$  denotes a diagonal matrix with the same diagonal elements as  $\mathbf{R}$ , such that  $\left|\overline{R}_{ij}\right| \leq 1$  for all  $i, j \in [N]$ ,  $\overline{R}_{ii} = 1$  for all  $i \in [N]$ , and  $\text{sine}^{-1}(z) \triangleq \sin^{-1}(\Re\{z\}) + j \cdot \sin^{-1}(\Im\{z\})$  operates elementwise.

Since the array manifold matrix of a ULA is a Vandermonde matrix (e.g., [17]) and all the signals involved are uncorrelated, the covariance matrix C (4) is a Toeplitz matrix [18], thus we denote  $C_{ij} \triangleq c_{|i-j|+1}$ , so the (i, j)-th element of  $\overline{R}$  can be expressed as

$$\bar{R}_{ij} = \frac{R_{ij}}{\sqrt{D_{ii}D_{jj}}} = \frac{\psi_i\psi_j c_{|i-j|+1}e^{j(\phi_i - \phi_j)} + \delta_{ij}\sigma_w^2}{\sqrt{(\psi_i^2 c_1 + \sigma_w^2)(\psi_j^2 c_1 + \sigma_w^2)}}$$

where  $\delta_{ij}$  denotes the Kronecker delta of  $i, j \in \mathbb{N}$ . Indeed, it is readily verified that  $\overline{R}_{ii} = 1$  for all  $i \in [N]$ . In addition, since both  $\mathbf{R}^s$  and  $\sigma_v^2$  are unknown, and since all  $\{\overline{R}_{ij}\}$  are related to  $\{\psi_i\}$ and  $\{c_{|i-j|+1}\}$  by the products  $\{\psi_i\psi_jc_{|i-j|+1}\}$  only, we can set w.l.o.g.  $c_1 = 1$ . We will shortly see that the (unknown) Signal-to-Interference Ratios SIR<sub>m</sub>  $\triangleq R_{mm}^s/\sigma_v^2$  are immaterial to our blind calibration problem. This is in stark contrast to the Signal-to-Noise Ratios,  $\text{SNR}_n \triangleq \psi_n^2 / \sigma_w^2$ , which, after setting  $c_1 = 1$ , are now expressed via the sensors' gains  $\{\psi_n\}$ , and are also unknown.

To see this, observe that in the limit of high SNR we have

$$\lim_{\sigma_w^2 \to 0} \overline{R}_{ij} = \frac{\psi_i \psi_j c_{|i-j|+1} e^{j(\phi_i - \phi_j)}}{\sqrt{\psi_i^2 \psi_j^2}} = c_{|i-j|+1} e^{j(\phi_i - \phi_j)},$$

namely a complete loss of information regarding the sensors' gains. Note that [12] addresses (only) this particular case in the context of blind calibration for sparse arrays, consequently only calibrating the sensors' phases. Our signal model (1) therefore offers an extended framework (*cf.* [12], Eq. (1)), which allows blind calibration of the sensors' gains as well for any finite (even if large) SNR, provided that the sample size is sufficiently large. Be that as it may, from a practical point of view, at high SNR the noiseless model still leads to "good" solutions for phase calibration (only).

Lastly, recall that even with an ideal  $\infty$ -bit ADC, and therefore in our case as well, at least two phases and one gain references are required for calibration of the array in this blind scenario (e.g., [19]).

## 4. ONE-BIT BLIND CALIBRATION

In light of the discussion above, w.l.o.g., we assume from now on

$$\psi_1 = 1, \ \phi_1 = \phi_2 = 0, \ c_1 = 1.$$
 (8)

With these assumptions, we define the underlying vector of unknown parameters, which fully determines  $\mathbf{R}^y$  with  $\sigma_w^2$  and (8), as

$$\boldsymbol{\theta} \triangleq \begin{bmatrix} \boldsymbol{\psi}_{[2]}^{\mathrm{T}} \ \boldsymbol{\phi}_{[3]}^{\mathrm{T}} \ \boldsymbol{\rho}_{[2]}^{\mathrm{T}} \ \boldsymbol{\iota}_{[2]}^{\mathrm{T}} \end{bmatrix}^{\mathrm{T}} \in \mathbb{R}^{K_{\boldsymbol{ heta}} imes 1}$$

where we use the notation  $\boldsymbol{a}_{[k]} \triangleq [a_k \ a_{k+1} \ \cdots \ a_N]^{\mathrm{T}} \in \mathbb{R}^{(N-k+1)\times 1}$ for any  $\boldsymbol{a} \in \mathbb{R}^{N\times 1}$  and  $k \leq N$ , and where  $\boldsymbol{c} \triangleq \boldsymbol{\rho} + j \cdot \boldsymbol{\iota} \in \mathbb{C}^{N\times 1}$ , such that  $\rho_1 = 1$  and  $\iota_1 = 0$  from (8), and  $K_{\theta} \triangleq 4N - 5$ .

Many of the previously proposed methods for blind calibration are based on the fact that, theoretically,  $\widehat{\mathbf{R}}^{y} \triangleq \frac{1}{T} \sum_{t=1}^{T} \mathbf{y}(t) \mathbf{y}(t)^{\mathrm{H}} \in \mathbb{C}^{N \times N}$  can be made arbitrarily close to  $\mathbf{R}^{y}$  by increasing (appropriately) the sample size T. However, in practice, the available sample size is always limited and is oftentimes fixed. Hence, rather than relying on the coarse approximation  $\widehat{\mathbf{R}}^{y} \approx \mathbf{R}^{y}$ , which typically leads to coarse, sub-optimal estimates of  $\{\psi, \phi\}$ , we propose a refined approach, which obtains more accurate estimates of  $\psi$  and  $\phi$  by accounting for information regarding the estimation errors in  $\widehat{\mathbf{R}}^{y}$ , which are generally correlated [20].

Specifically, for any finite sample size T, we have

$$\widehat{\boldsymbol{R}}^{y} \triangleq \boldsymbol{R}^{y} + \boldsymbol{\mathcal{E}} \implies \widehat{R}_{ij}^{y} = R_{ij}^{y} + \mathcal{E}_{ij}, \ \forall i, j \in [N],$$

where  $\{\mathcal{E}_{ij}\}$  denote the estimation errors in the estimation of  $\{R_{ij}^y\}$ . Using (6), it is readily seen that while sine  $(\frac{\pi}{2}R_{ij}^y) = \overline{R}_{ij}$ ,

sine 
$$\left(\frac{\pi}{2}\widehat{R}_{ij}^{y}\right) = \operatorname{sine}\left(\operatorname{sine}^{-1}(\overline{R}_{ij}) + \frac{\pi}{2}\mathcal{E}_{ij}\right) \triangleq \overline{R}_{ij} + \widetilde{\mathcal{E}}_{ij},$$

where  $\operatorname{sin}(z) \triangleq \operatorname{sin}(\Re\{z\}) + j \cdot \operatorname{sin}(\Im\{z\})$ , and  $\{\widetilde{\mathcal{E}}_{ij}\}$  are the implied "transformed" estimation errors. A direct, explicit SOS-based analytic characterization of  $\widetilde{\mathcal{E}}_{ij}$  is rather involved, so we seek an estimate which would *implicitly* account for correlations among  $\{\mathcal{E}_{ij}\}$ . This leads us to the following solution approach.

#### 4.1. Estimation by KLD Covariance Fitting

We propose to use the KLD of two zero-mean circular CN distributions as a covariance fitting criterion, in order to implicitly extract more information contained is the SOS. Since the zero-mean CN distribution is fully characterized by SOS, note that the KLD of two zero-mean multivariate CN distributions serves as a plausible criterion for *consistent* covariance matrix estimation. Indeed, from Gibbs' inequality, the KLD is always non-negative (e.g., [21]), hence

$$D_{\mathrm{KL}}^{\mathcal{CN}}\left(\widehat{\boldsymbol{\Sigma}}, \boldsymbol{\Sigma}\right) \triangleq \log\left(\frac{\det \boldsymbol{\Sigma}}{\det \widehat{\boldsymbol{\Sigma}}}\right) + \mathrm{Tr}\left(\widehat{\boldsymbol{\Sigma}}\boldsymbol{\Sigma}^{-1}\right) - N \ge 0, \quad (9)$$

$$D_{\mathsf{KL}}^{CN}\left(\widehat{\boldsymbol{\Sigma}}, \boldsymbol{\Sigma}\right) = 0 \iff \widehat{\boldsymbol{\Sigma}} = \boldsymbol{\Sigma} \in \mathbb{C}^{N \times N}.$$
(10)

Moreover, it can be easily shown that minimization of (9) is equivalent to maximization of the likelihood function for CN-distributed measurements (e.g., [22], Eq. (60)), which, in turn, is asymptotically equivalent to minimization of the optimally weighted nonlinear Least Squares (LS) objective function (see [22], Eq. (62), and Appendix E therein for the full details). Specifically, we propose

$$\widehat{\boldsymbol{\theta}}_{\text{KLD}} \triangleq \underset{\boldsymbol{\theta} \in \mathbb{R}^{K_{\boldsymbol{\theta}} \times 1}}{\operatorname{argmin}} D_{\text{KL}}^{CN} \left( \widehat{\boldsymbol{R}}^{y}, \boldsymbol{R}^{y}(\boldsymbol{\theta}) \right).$$
(11)

Thus, due to all of the above, the proposed estimate (11), from which the KLD-based estimates of  $\psi_{[2]}$  and  $\phi_{[3]}$  are readily extracted, implicitly takes into account cross-correlations between all the estimation errors { $\mathcal{E}_{ij}$ } and utilizes "hidden" SOS information, as desired.

## 4.2. Computation of the Proposed KLD-based Estimate

As mentioned above, since  $\hat{\theta}_{\text{KLD}}$  is equivalent to the maximum likelihood estimate for CN-distributed measurements, we propose to use the respective *Fisher's Scoring Algorithm* (FSA, e.g., [23]), pretending that  $\{y(t)\}_{t=1}^{T}$  are CN-distributed. Of course, in our case the FSA would merely serve as a quasi-Newton algorithm (e.g., [24]), since  $\{y(t)\}_{t=1}^{T}$  are obviously far from being CN-distributed.

The update equation of the FSA for the k-th iteration is given by

$$\widehat{\boldsymbol{\theta}}^{(k)} = \widehat{\boldsymbol{\theta}}^{(k-1)} + \boldsymbol{J}^{-1} \left( \widehat{\boldsymbol{\theta}}^{(k-1)} \right) \left. \nabla_{\boldsymbol{\theta}} \mathcal{L} \right|_{\boldsymbol{\theta} = \widehat{\boldsymbol{\theta}}^{(k-1)}}, \quad (12)$$

where  $\widehat{\boldsymbol{\theta}}^{(k)}$  denotes the estimate of  $\boldsymbol{\theta}$  in the k-th iteration,

$$\mathcal{L}(\boldsymbol{\theta}) \triangleq -T \cdot D_{\mathrm{KL}}^{\mathcal{CN}} \left( \widehat{\boldsymbol{R}}^{y}, \boldsymbol{R}^{y}(\boldsymbol{\theta}) \right) + b, \qquad (13)$$

is the associated log-likelihood function (where  $b \in \mathbb{R}$  is constant w.r.t.  $\boldsymbol{\theta}$ ), and  $\boldsymbol{J}(\boldsymbol{\theta}) \in \mathbb{R}^{K_{\boldsymbol{\theta}} \times K_{\boldsymbol{\theta}}}$  is the associated Fisher Information Matrix (FIM) under the (false) assumption that  $\{\boldsymbol{y}(t)\}_{t=1}^{T}$  are CN, given by [25] (for all  $i, j \in [K_{\boldsymbol{\theta}}]$ )

$$J_{ij}(\boldsymbol{\theta}) = T \cdot \operatorname{Tr}\left( \left( \boldsymbol{R}^{y} \right)^{-1} \left( \frac{\partial}{\partial \theta_{i}} \boldsymbol{R}^{y} \right) \left( \boldsymbol{R}^{y} \right)^{-1} \left( \frac{\partial}{\partial \theta_{j}} \boldsymbol{R}^{y} \right) \right).$$
(14)

Hence, in order to carry out the iterations (12) such that they will successfully converge to  $\hat{\theta}_{\text{KLD}}$ , three ingredients are required: a sufficiently "good" initial solution  $\hat{\theta}^{(0)}$ , and closed-form expressions of the FIM and the score function, i.e.,  $J(\theta)$  and  $\nabla_{\theta} \mathcal{L} \in \mathbb{R}^{K_{\theta} \times 1}$ , resp.

Starting with the elements of  $J(\theta)$ , and denoting  $\overline{R}_{mn} \triangleq \overline{R}_{mn}^{\Re} + j \cdot \overline{R}_{mn}^{\Im}$  for brevity, by the chain rule, we have for every  $\theta_i$ 

$$\frac{\partial R_{mn}^{y}}{\partial \theta_{i}} = \frac{2/\pi}{\sqrt{1 - \left(\bar{R}_{mn}^{\Re}\right)^{2}}} \frac{\partial \bar{R}_{mn}^{\Re}}{\partial \theta_{i}} + \frac{j \cdot 2/\pi}{\sqrt{1 - \left(\bar{R}_{mn}^{\Im}\right)^{2}}} \frac{\partial \bar{R}_{mn}^{\Im}}{\partial \theta_{i}}, \quad (15)$$

where we have used  $\frac{\partial}{\partial x} \sin^{-1}(x) = (1 - x^2)^{-\frac{1}{2}}$ . Given (15), all  $\{\frac{\partial}{\partial \theta_i} \mathbf{R}^y\}$  can now be easily computed using straightforward derivatives, omitted due to lack of space. As a representative example,

$$\frac{\partial \overline{R}_{mn}^{\Re}}{\partial \phi_k} = \left(\rho_{|m-n|+1}\sin(\phi_m - \phi_n) + \iota_{|m-n|+1}\cos(\phi_m - \phi_n)\right) \cdot \frac{\psi_m \psi_n}{\sqrt{D_{mm} D_{nn}}} (\delta_{km} - \delta_{kn}), \ \forall m \neq n \in [N], \forall (k-2) \in [N-2].$$

With  $\{\frac{\partial}{\partial \theta_i} \mathbf{R}^y\}_{i=1}^{K_{\theta_i}}$ , we readily obtain closed-form expressions for

all of (14). As for  $\nabla_{\theta} \mathcal{L}$ , using the chain rule once again, we have

$$\frac{\partial \mathcal{L}(\boldsymbol{\theta})}{\partial \theta_i} = \sum_{m,n=1}^{N} \frac{\partial \mathcal{L}\left(\boldsymbol{\theta}\right)}{\partial R_{mn}^y} \frac{\partial R_{mn}^y}{\partial \theta_i} = \operatorname{Tr}\left(\left(\nabla_{\boldsymbol{R}^y} \mathcal{L}\right) \left(\frac{\partial}{\partial \theta_i} \boldsymbol{R}^y\right)^{\mathrm{T}}\right),$$

for all  $i \in [K_{\theta}]$ , where

$$\nabla_{\mathbf{R}^{y}}\mathcal{L} = -T \cdot \left[ (\mathbf{R}^{y})^{-1} \left( \mathbf{I}_{N} - \widehat{\mathbf{R}}^{y} (\mathbf{R}^{y})^{-1} \right) \right]^{\mathrm{T}} \in \mathbb{C}^{N \times N}.$$
(16)

Thus, with (16) and the already obtained  $\{\frac{\partial}{\partial \theta_i} \mathbf{R}^y\}_{i=1}^{K_{\theta}}$ , we now have a closed-form expression for the score function  $\nabla_{\theta} \mathcal{L}$  as well.

The third ingredient—a "good" initial solution to the iterative procedure (12)—is instrumental from a practical point of view. Indeed, generally, the FSA is likely to converge to (11) when  $\hat{\theta}^{(0)}$  is "close" enough to  $\hat{\theta}_{\text{KLD}}$ , which is the global maximizer of (13).

For this, we propose to use Paulraj and Kailath's (P-K's) classical LS approach [26]. Despite ignoring the estimation errors  $\{\mathcal{E}_{ij}\}$ , P-K's approach yields simple closed-form expressions for consistent estimates of  $\{\psi_{[2]}, \phi_{[3]}\}$ , based on the Toeplitz structure of C, which can still be exploited under one-bit quantization. Indeed,  $\hat{\phi}_{[3],LS}$  can be obtained just as described in [26], using  $\widehat{\overline{R}} \triangleq \operatorname{sine}\left(\frac{\pi}{2}\widehat{R}^{y}\right)$ . Further,  $\widehat{\psi}_{[2],LS}$  can be obtained as well with the following, relatively simple modifications. Since C is Toeplitz, due to (3) and (7), we have (*cf.* [26], Eq. (11))

$$\mu_{ijk\ell} \triangleq \log\left(\frac{|\overline{R}_{ij}|}{|\overline{R}_{k\ell}|}\right) = \beta_i - \beta_j - \beta_k + \beta_\ell,$$

for any four indices satisfying  $i - j = k - \ell \neq 0$ , where

$$\beta_n \triangleq \log\left(\frac{\psi_n}{\sqrt{\psi_n^2 + \sigma_w^2}}\right) \iff \psi_n = \sqrt{\frac{e^{2\beta_n} \sigma_w^2}{1 - e^{2\beta_n}}}, \forall n \in [N].$$
(17)

Thus, recalling that  $\sigma_w^2$  and  $\psi_1$  are known, which means that  $\beta_1$  is known, the LS estimate  $\hat{\beta}_{[2],LS}$  based on  $\hat{\overline{R}}$  is obtained via [26], Eq. (13) (excluding equations associated with the diagonal elements). Given  $\hat{\beta}_{[2],LS}$ , the estimates  $\hat{\psi}_{[2],LS}$  are readily computed from (17).

Once  $\widehat{\psi}_{[2],LS}$  and  $\widehat{\phi}_{[3],LS}$  are obtained, we define the diagonal matrix  $\widehat{D}_{LS}$  with diagonal entires  $\widehat{D}_{nn,LS} \triangleq \widehat{\psi}_{n,LS}^2 + \sigma_w^2$ , with which we further define  $\widehat{R}_{LS} \triangleq \widehat{D}_{LS}^{\frac{1}{2}} \widehat{\overline{R}} \widehat{D}_{LS}^{\frac{1}{2}}$  based on (7), and (based on (3))

$$\widehat{\boldsymbol{C}}_{\text{LS}} \triangleq \widehat{\boldsymbol{\Phi}}_{\text{LS}}^* \widehat{\boldsymbol{\Psi}}_{\text{LS}}^{-1} \left( \widehat{\boldsymbol{R}}_{\text{LS}} - \sigma_w^2 \boldsymbol{I}_N \right) \widehat{\boldsymbol{\Psi}}_{\text{LS}}^{-1} \widehat{\boldsymbol{\Phi}}_{\text{LS}},$$
$$\widehat{c}_{|i-j|+1,\text{LS}} \triangleq \widehat{C}_{ij,\text{LS}} \implies \widehat{\boldsymbol{c}}_{\text{LS}} \triangleq \widehat{\boldsymbol{\rho}}_{\text{LS}} + \jmath \cdot \widehat{\boldsymbol{\iota}}_{\text{LS}},$$

where  $\widehat{\Psi}_{LS} \triangleq \text{Diag}(\widehat{\psi}_{LS}), \widehat{\Phi}_{LS} \triangleq \text{Diag}(\widehat{\phi}_{LS}), \widehat{\rho}_{1,LS} = 1$  and  $\widehat{\iota}_{1,LS} = 0$ . Collecting all the respective LS-based estimates into a



Fig. 1: (a) Percentage of convergent FSA (12) trials; (b) Upon convergence, average number of iterations to convergence, both vs. T, for  $\hat{\theta}^{(0)} = \hat{\theta}_{LS}$ .



Fig. 2: MSEs of the gains and phases offsets estimates vs. T. Though both the LS- and KLD-based estimates exhibit a consistency trend, the KLD are superior.



Fig. 3: Empirical probability of error in detecting the number of sources vs. T. Clearly, the KLD-based enhanced calibration significantly improves the accuracy in this post-calibration covariance-based estimation task.

 $K_{\theta}$ -dimensional vector, we now have

$$\widehat{\boldsymbol{\theta}}_{\text{LS}} \triangleq \left[\widehat{\boldsymbol{\psi}}_{[2],\text{LS}}^{\text{T}} \, \widehat{\boldsymbol{\phi}}_{[3],\text{LS}}^{\text{T}} \, \widehat{\boldsymbol{\rho}}_{[2],\text{LS}}^{\text{T}} \, \widehat{\boldsymbol{\iota}}_{[2],\text{LS}}^{\text{T}}\right]^{\text{T}} \triangleq \widehat{\boldsymbol{\theta}}^{(0)} \in \mathbb{R}^{K_{\theta} \times 1}, \quad (18)$$

which is a consistent estimate of  $\theta$  (from the consistency of  $\widehat{R}^{y}$ ), thus serving as a "good" initial solution, presumably "close" to  $\widehat{\theta}_{\text{KLD}}$ . The complete proposed blind calibration scheme is as follows:

**KLD-based One-Bit Blind Calibration:** Given 
$$\{\boldsymbol{y}(t)\}_{t=1}^{T}$$
,(1) Compute  $\widehat{\boldsymbol{R}}^{y} = \frac{1}{T} \sum_{t=1}^{T} \boldsymbol{y}(t) \boldsymbol{y}(t)^{H}$ ;(2) Compute and construct  $\widehat{\boldsymbol{\theta}}_{LS}$  as in (18);(3) Set  $\widehat{\boldsymbol{\theta}}^{(0)} = \widehat{\boldsymbol{\theta}}_{LS}$ , and iterate (12) until convergence (e.g., until  $\left\|\widehat{\boldsymbol{\theta}}^{(k)} - \widehat{\boldsymbol{\theta}}^{(k-1)}\right\|_{2} < \epsilon$ , for some "small"  $\epsilon \in \mathbb{R}_{+}$ );(4) Calibrate the array using  $\widehat{\boldsymbol{\psi}}_{KLD} \triangleq \begin{bmatrix} 1 \ \widehat{\boldsymbol{\psi}}_{[2],KLD}^{T} \end{bmatrix}^{T}$  and  $\widehat{\boldsymbol{\phi}}_{KLD} \triangleq \begin{bmatrix} 0 \ 0 \ \widehat{\boldsymbol{\phi}}_{[3],KLD}^{T} \end{bmatrix}^{T}$ .

We note that even if access to the pre-quantized signals is restricted, so that  $\widehat{\psi}_{\text{KLD}}$  and  $\widehat{\phi}_{\text{KLD}}$  cannot be compensated for *prior* to the one-bit quantization, our method nevertheless provides  $\widehat{C}_{\text{KLD}}$ —the KLD-based estimate of C—which is readily constructed from  $\widehat{\rho}_{\text{KLD}} \triangleq [1 \ \widehat{\rho}_{[2],\text{KLD}}^{\text{T}}]^{\text{T}}$  and  $\widehat{\iota}_{\text{KLD}} \triangleq [0 \ \widehat{\iota}_{[2],\text{KLD}}]^{\text{T}}$ , and can be used for subsequent post-calibration covariance-based estimation tasks.  $\widehat{C}_{\text{KLD}}$  essentially serves as the outcome of implicit KLD-based blind calibration, applied directly to the one-bit quantized measurements.

## 5. SIMULATION RESULTS

We consider model (1) with one-bit measurements (5), a half wavelength inter-element spacing  $d = \lambda/2$  ULA consisting of N = 7 elements, and M = 4 equal power, zero-mean sources arriving from angles  $\alpha = [45^{\circ} 52^{\circ} 9^{\circ} 78^{\circ}]^{\mathrm{T}}$ . The sensors' gains and

phases were set to  $\boldsymbol{\psi} = \begin{bmatrix} 1 & 0.7 & 0.9 & 1.1 & 1.2 & 0.8 & 1.3 \end{bmatrix}^{\mathrm{T}}$  and  $\boldsymbol{\phi} = \begin{bmatrix} 0^{\circ} 0^{\circ} 5^{\circ} 11^{\circ} & -8^{\circ} 4^{\circ} 10^{\circ} \end{bmatrix}^{\mathrm{T}}$ , resp., where w.l.o.g. we assume  $\psi_1, \phi_1$  and  $\phi_2$  are known references. The sources' (equal) power was set such that  $\{\mathrm{SIR}_m = 10\}_{m=1}^4$  and  $c_1 = 1$ , and  $\sigma_w^2 = 1$ . All empirical results were obtained by averaging  $10^4$  independent trials.

Figs. 1a and 1b present the empirical percentage of convergent trials and number of iterations to convergence, resp., of FSA vs. T, when initialized by  $\hat{\theta}_{LS}$ . Here, we say that the iterations (12) converge when the  $\ell^2$ -norm of the step size falls below  $\epsilon = 10^{-7}$  before reaching the maximum number of allowed iteration  $K_{\text{max}} = 100$ . It is clearly demonstrated that when the sample size is sufficiently large, the iterations converge with high probability, which is due to the consistency of  $\hat{\theta}_{LS}$ . We also emphasize that although the nominal values of T considered here may seem typically large, the respective number of *bits per unkonwn* is both realistic and practical owing to the high operational rates of low-complexity 1-bit ADCs.

Next, we turn to evaluate the MSEs of the proposed  $\hat{\psi}_{\text{KLD}}$  and  $\hat{\phi}_{\text{KLD}}$ , and their more simple LS-based alternatives  $\hat{\psi}_{\text{LS}}$  and  $\hat{\phi}_{\text{LS}}$ , presented in Fig. 2 vs. *T*. For a pure (and fair) comparison of the MSEs (regardless of convergence issues), the average MSEs of the KLD-based estimates are based only on convergent FSA trials. As seen, the proposed estimates exhibit enhanced performance, reducing the MSEs by up to ~6[dB]. For practical considerations, whenever FSA does not converge, one can resort to calibration by  $\hat{\psi}_{\text{LS}}$  and  $\hat{\phi}_{\text{LS}}$ .

Lastly, we demonstrate the implications of our enhanced solution by a subsequent post-calibration task—blind determination of the number of sources (e.g., [27]). To this end, we use  $\widehat{C}_{\text{KLD}}$  and  $\widehat{C}_{\text{LS}}$ as the post-calibration estimated covariance matrix of the received signal x(t), where the goal now is to detect the number of sources M, recalling  $\sigma_v^2$  is unknown. We use the second order statistic of the eigenvalues (SORTE) algorithm [28], which requires only the eigenvalues of the sample covariance matrix (of a calibrated array). Fig. 3, presenting the empirical probability of error in detecting M vs. T, well demonstrates the considerable performance improvement, a direct implication of the proposed KLD-based enhanced calibration.

## 6. CONCLUSION

We presented a blind calibration scheme for ULAs with coarsely quantized one-bit measurements, using a KLD-based calibration scheme. The KLD covariance fitting criterion—closely related to the optimally weighted nonlinear least squares criterion—implicitly exploit "hidden" SOS information, thus leading to higher accuracy. We also provided an approximate iterative solution algorithm, along with an educated initial estimate (consistent in itself). Significant performance improvements, both in terms of calibration and a postcalibration estimation task, were demonstrated in simulations.

## 7. REFERENCES

- Robert H. Walden, "Analog-to-Digital Converter Survey and Analysis," *IEEE J. Sel. Areas Commun.*, vol. 17, no. 4, pp. 539–550, 1999.
- [2] Bin Le, Thomas W. Rondeau, Jeffrey H. Reed, and Charles W. Bostian, "Analog-to-Digital Converters," *IEEE Signal Process. Mag.*, vol. 22, no. 6, pp. 69–77, 2005.
- [3] Hongjian Sun, Arumugam Nallanathan, Cheng-Xiang Wang, and Yunfei Chen, "Wideband Spectrum Sensing for Cognitive Radio Networks: A Survey," *IEEE Wireless Communications*, vol. 20, no. 2, pp. 74–81, 2013.
- [4] Pietro Stinco, Maria Greco, Fulvio Gini, and Mario La Manna, "Compressed Spectrum Sensing in Cognitive Radar Systems," in *Proc. of ICASSP*, 2014, pp. 81–85.
- [5] Graeme E. Smith, Zach Cammenga, Adam Mitchell, Kristine L. Bell, Joel Johnson, Muralidhar Rangaswamy, and Christopher Baker, "Experiments with Cognitive Radar," *IEEE Aerosp. Electron. Syst. Mag.*, vol. 31, no. 12, pp. 34–46, 2016.
- [6] Junil Choi, Jianhua Mo, and Robert W. Heath, "Near Maximum-Likelihood Detector and Channel Estimator for Uplink Multiuser Massive MIMO Systems with One-Bit ADCs," *IEEE Trans. Commun.*, vol. 64, no. 5, pp. 2005–2018, 2016.
- [7] Yulong Gao, Deshun Hu, Yanping Chen, and Yongkui Ma, "Gridless 1-b DOA Estimation Exploiting SVM Approach," *IEEE Commun. Lett.*, vol. 21, no. 10, pp. 2210–2213, 2017.
- [8] Jiaying Ren and Jian Li, "One-Bit Digital Radar," in Proc. 51st Asilomar Conference on Signals, Systems, and Computers, 2017, pp. 1142–1146.
- [9] Chun-Lin Liu and P.P. Vaidyanathan, "One-Bit Sparse Array DOA Estimation," in *Proc. of ICASSP*, 2017, pp. 3126–3130.
- [10] Aria Ameri, Arindam Bose, Jian Li, and Mojtaba Soltanalian, "One-Bit Radar Processing with Time-Varying Sampling Thresholds," *IEEE Trans. Signal Process.*, vol. 67, no. 20, pp. 5297–5308, 2019.
- [11] Xiaodong Huang and Bin Liao, "One-Bit MUSIC," *IEEE Sig-nal Process. Lett.*, vol. 26, no. 7, pp. 961–965, 2019.
- [12] Krishnaprasad Nambur Ramamohan, Sundeep Prabhakar Chepuri, Daniel Fernández Comesana, and Geert Leus, "Blind Calibration of Sparse Arrays for DOA Estimation with Analog and One-Bit Measurements," in *Proc. of ICASSP*, 2019, pp. 4185–4189.
- [13] Benedikt Loesch and Bin Yang, "Cramér-Rao Bound for Circular and Noncircular Complex Independent Component Analysis," *IEEE Trans. Signal Process.*, vol. 61, no. 2, pp. 365–379, 2013.
- [14] Ofer Bar-Shalom and Anthony J. Weiss, "DOA Estimation Using One-Bit Quantized Measurements," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 38, no. 3, pp. 868–884, 2002.
- [15] Giovanni Jacovitti and Alessandro Neri, "Estimation of the autocorrelation function of complex gaussian stationary processes by amplitude clipped signals," *IEEE Trans. Information theory*, vol. 40, no. 1, pp. 239–245, 1994.
- [16] J. H. Van Vleck and David Middleton, "The Spectrum of Clipped Noise," *Proceedings of the IEEE*, vol. 54, no. 1, pp. 2–19, 1966.

- [17] Kaare Brandt Petersen, Michael Syskind Pedersen, et al., "The matrix cookbook," *Technical University of Denmark*, vol. 7, no. 15, pp. 510, 2008.
- [18] Robert M Gray, "Toeplitz and circulant matrices: A review," *Foundations and Trends*® in *Communications and Information Theory*, vol. 2, no. 3, pp. 155–239, 2006.
- [19] Astély, David and Swindlehurst, A. Lee and Ottersten, Bjorn, "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.
- [20] Amir Weiss and Arie Yeredor, "Asymptotically Optimal Blind Calibration of Uniform Linear Sensor Arrays for Narrowband Gaussian Signals," *IEEE Trans. Signal Process.*, vol. 68, pp. 5322–5333, 2020.
- [21] MacKay, David J.C., Information Theory, Inference and Learning Algorithms, Cambridge University Press, 2003.
- [22] Amir Weiss, "Blind Direction-of-Arrival Estimation in Acoustic Vector-Sensor Arrays via Tensor Decomposition and Kullback-Leibler Divergence Covariance Fitting," submitted to *IEEE Trans. Signal Process.*; also: ArXiv e-prints, May 2020, arXiv:2005.08318 [eess.SP], https://arxiv.org/abs/2005.08318.
- [23] Robert I. Jennrich and P.F. Sampson, "Newton-Raphson and Related Algorithms for Maximum Likelihood Variance Component Estimation," *Technometrics*, vol. 18, no. 1, pp. 11–17, 1976.
- [24] Joseph-Frédéric Bonnans, Jean Charles Gilbert, Claude Lemaréchal, and Claudia A. Sagastizábal, Numerical Optimization: Theoretical and Practical Aspects, Springer Science & Business Media, 2006.
- [25] Sandra L. Collier, "Fisher Information for a Complex Gaussian Random Variable: Beamforming Applications for Wave Propagation in a Random Medium," *IEEE Trans. Signal Process.*, vol. 53, no. 11, pp. 4236–4248, 2005.
- [26] Arogyaswami Paulraj and Thomas Kailath, "Direction of Arrival Estimation by Eigenstructure Methods with Unknown Sensor Gain and Phase," in *Proc. of ICASSP*, 1985, vol. 10, pp. 640–643.
- [27] Amir Weiss and Arie Yeredor, "Blind Determination of the Number of Sources Using Distance Correlation," *IEEE Signal Process. Lett.*, vol. 26, no. 6, pp. 828–832, 2019.
- [28] Zhaoshui He, Andrzej Cichocki, Shengli Xie, and Kyuwan Choi, "Detecting the Number of Clusters in n-Way Probabilistic Clustering," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 32, no. 11, pp. 2006–2021, 2010.