IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received August 17, 2020, accepted September 5, 2020, date of publication September 8, 2020,
date of current version September 23, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.3022710

On the Gaussian Cramér-Rao Bound for Blind Single-Input Multiple-Output System Identification: Fast and Asymptotic Computations

MOHAMED NAIT-MEZIANE[1], KARIM ABED-MERAIM[1], (Fellow, IEEE), ZHIPENG ZHAO[2], AND NGUYEN LINH TRUNG[3], (Senior Member, IEEE)
[1]PRISME Laboratory, University of Orléans, 45067 Orléans, France
[2]Mathematical and Algorithmic Sciences Lab, France Research Center, Huawei Technologies Company Ltd., 92100 Boulogne-Billancourt, France
[3]AVITECH, University of Engineering and Technology, Vietnam National University, Hanoi 100000, Vietnam

Corresponding author: Mohamed Nait-Meziane (mohamed.nait-meziane@univ-orleans.fr)

This work was sponsored in part by the Huawei Innovation Research Program (HIRP) under RD Agreement YBN2018035141, and in part by the National Foundation for Science and Technology Development of Vietnam under Grant 01/2019/TN.

ABSTRACT The Cramér-Rao Bound (CRB) is a powerful tool to assess the performance limits of a parameter estimation problem for a given statistical model. In particular, the Gaussian CRB (i.e., the CRB obtained assuming the data are Gaussian) corresponds to the worst case; giving the largest CRB among a large class of data distributions. This makes it very useful in practice since optimizing under the Gaussian data assumption can be interpreted as a min-max optimization (i.e., minimizing the largest CRB). The Gaussian CRB is also the corresponding bound of Second-Order Statistics (SOS)-based estimation methods, which are frequently used in practice. Despite its practicality, computing this bound might be cumbersome in some cases, particularly in the case where the input is assumed deterministic and has a large number of samples. In this paper, we address this computational issue by proposing a fast computation for the deterministic Gaussian CRB of Single-Input Multiple Output (SIMO) blind system identification. More precisely, we exploit circulant matrix properties to reduce the cost from cubic to quadratic with respect to the sample size. Moreover, we derive a closed-form formula for the asymptotic (large sample size) Gaussian CRB and show how it can be computed using the residue theorem.

INDEX TERMS Blind system identification, Cramér-Rao Bound (CRB), Single-Input Multiple-Output (SIMO) systems, fast and asymptotic computations.

I. INTRODUCTION

Parametric estimation plays a central role in many signal processing problems encountered in various fields such as radar, communications, and seismology. When dealing with parametric estimation, the analytic evaluation of the exact performance of an estimator is often intractable. A widely used approach to bypass this issue consists in first finding a lower bound (or limit) on the estimation performance for the problem at hand (independently of the considered estimator), and then comparing the experimental performance of the estimator of interest to this bound. The knowledge of such performance bound is of great importance since: i) it allows to know if, after using a particular estimator, an improvement is possible or not [1]; ii) it allows to know if respecting certain performance requirements is possible in a given context [2]; and iii) it allows a system design with the best achievable accuracy [3], [4]. Among the existing bounds, the CRB (Cramér-Rao Bound) is the most widely used for benchmarking the mean-square-error (MSE) of unbiased estimators [5], [6]. Particularly, the Gaussian CRB is of great practicality as it represents the least favorable (giving the largest CRB) but the most tractable and useful case [7]. Basically, if an estimator performs well under the Gaussian assumption, then we expect to be able to do better under other assumptions. Moreover, the Gaussian CRB is a lower bound

The associate editor coordinating the review of this manuscript and approving it for publication was Qinghua Guo.

on all estimators based on second-order statistics (SOS); a class of estimators frequently considered in practice.

In this work, we consider the problem of identifying the channel coefficients and input signal of a Single-Input Multiple-Output (SIMO) system in the blind context; i.e., the only known quantity is the output. This problem has been studied thoroughly in the literature and several SOS-based solutions have been proposed for it in the past [8]–[10]. Even though the Gaussian CRB formula for this problem is known, its computation for a deterministic input remains, however, a challenging task due to the excessive computational cost it incurs, particularly in cases where a one-time offline computation is not an option; for example, when we need to compute the CRB for different sets of parameters. In this work, we focus on this specific problem and propose an alternative formulation for the elements of the CRB matrix allowing for a significant reduction in its computational cost. This formulation involves expressing the most computationally heavy blocks in terms of a circulant matrix, then use its diagonalized form to enable fast computations. This is possible thanks to the fact that a circulant matrix can be diagonalized using the Discrete Fourier Transform (DFT) matrix, which can be computed fast using the Fast Fourier Transform (FFT) algorithm. Furthermore, we also derive an asymptotic formula (given in an integral-form and assuming (very) large sample sizes) for the stochastic Gaussian CRB and provide an efficient solution for its computation based on the residue theorem [11].

This paper is organized as follows. We start by formulating the Blind System Identification (BSI) problem in Section II, then we present the CRB formula for this problem in Section III. Section IV follows with a presentation of the proposed fast computation of the CRB. In Section V, we present the proposed asymptotic CRB formula and validate it using computer simulations. Finally, we provide some concluding remarks in Section VI.

A. NOTATION

$\mathbb{C}$	Set of complex numbers.
$\mathbf{x}, \mathbf{X}$	Column-vectors are in bold lower-case font and matrices are in bold upper-case font.
$(\mathbf{X})_{ij}$	ij-th element of matrix $\mathbf{X}$.
$\mathrm{Re}(\cdot), \mathrm{Im}(\cdot)$	Real-part and imaginary-part operators.
$(\cdot)^*, (\cdot)^T, (\cdot)^H$	Conjugate, transpose, and conjugate-transpose operators, respectively.
$\mathbf{X}^{-1}, \mathbf{X}^{\frac{1}{2}}$	Matrix inverse and matrix square-root.
$\mathbf{X}^{-H}, \mathbf{X}^{\frac{H}{2}}$	Denote $(\mathbf{X}^H)^{-1} = (\mathbf{X}^{-1})^H$ and $(\mathbf{X}^H)^{\frac{1}{2}} = (\mathbf{X}^{\frac{1}{2}})^H$, respectively. (Similar notation if H is replaced by $*$ or T.)
$\mathrm{rank}(\mathbf{X})$	Rank of matrix $\mathbf{X}$.
$\mathrm{diag}(\mathbf{x})$	Diagonal matrix with components of $\mathbf{x}$ as diagonal elements.
$\mathbf{I}_N, \mathbf{I}$	Identity matrix of size $N \times N$. Subscript N is dropped if understood from context.
$\mathbf{e}_i$	i-th column of $\mathbf{I}$. Size understood from context.
$\mathbf{0}_{N\times N}, \mathbf{0}$	Matrix of size $N \times N$ with all elements 0. Subscript $N\times N$ is dropped if understood from context.
$\otimes$	Kronecker product.
$\delta(\tau)$	Dirac delta function; equals 1 at 0 and equals 0 elsewhere.
$\mathcal{O}(N)$	"Big-Oh" notation indicating that the computational cost of an (complex) operation is of order N, e.g., for $\mathbf{x}, \mathbf{y} \in \mathbb{C}^N$ and $\mathbf{A} \in \mathbb{C}^{M\times N}$, $\mathbf{x}^T\mathbf{y}$ is $\mathcal{O}(N)$ and $\mathbf{Ax}$ is $\mathcal{O}(MN)$.

II. PROBLEM FORMULATION

We consider a SIMO system where each output $x_i[n]$, $i = 1, \ldots, M$, is described using

$$x_i[n] = \sum_{l=0}^{L} h_i[l]s[n-l] + w_i[n], \quad n = 0, \ldots, N-1, \tag{1}$$

where $h_i[l]$ denotes coefficient l of channel i, L is the maximum channel order, $s[n]$ is the input signal (we assume $s[n] = 0$, for $n < 0$) and $w_i[n]$ is a zero-mean additive white (complex circular) Gaussian noise of variance σ^2. All quantities are assumed *complex-valued*. Using matrix-vector notation, (1) can be written

$$\mathbf{x} = \mathcal{H}\mathbf{s} + \mathbf{w} = \mathcal{S}\mathbf{h} + \mathbf{w}, \tag{2}$$

where $\mathbf{x} = [x_1[0], \ldots, x_1[N-1], \ldots, x_M[0], \ldots, x_M[N-1]]^T$, $\mathbf{s} = [s[0], \ldots, s[N-1]]^T$, $\mathbf{w} = [w_1[0], \ldots, w_1[N-1], \ldots, w_M[0], \ldots, w_M[N-1]]^T$, $\mathbf{h} = [h_1[0], \ldots, h_1[L], \ldots, h_M[0], \ldots, h_M[L]]^T$, $\mathcal{H} = [\mathbf{H}_1^T, \ldots, \mathbf{H}_M^T]^T$ where

$$\mathbf{H}_i = \begin{bmatrix} h_i[0] & & & & 0 \\ \vdots & \ddots & & & \\ h_i[L] & & h_i(0) & & \\ & \ddots & & \ddots & \\ 0 & & h_i[L] & \cdots & h_i[0] \end{bmatrix}_{N\times N}, \tag{3}$$

and $\mathcal{S} = \mathbf{I}_M \otimes \mathbf{S}$, where

$$\mathbf{S} = \begin{bmatrix} s[0] & & 0 \\ & \ddots & \\ \vdots & & s[0] \\ & & \vdots \\ s[N-1] & \cdots & s[N-L-1] \end{bmatrix}_{N\times(L+1)}.$$

The problem of blind SIMO system identification consists in "finding" the (unknown) channel coefficients and (unknown) input signal using only the (known) outputs together with certain (statistical) side information. In a deterministic context, we seek $\boldsymbol{\theta} = [\mathbf{h}^T, \mathbf{s}^T]^T$ assuming a known $\mathbf{x}$ and an additive white Gaussian noise with known[1] variance σ^2.

Note that BSI can be achieved only up to an unknown scalar (see [12], [13]). Appropriate constraints need to be considered for a 'full' identification (with the scalar indeterminacy removed). The corresponding CRB is termed constrained CRB.

III. CRB (UNCONSTRAINED AND CONSTRAINED) FORMULAS

Before presenting the proposed fast and asymptotic computations in sections IV and V, we recall in this section the formulas of the unconstrained CRB (referred to as CRB) and its constrained counterpart (referred to as CCRB) for the blind SIMO system identification problem defined in Section II.

A. UNCONSTRAINED CRB

The CRB matrix provides a lower bound on the error covariance matrix of any unbiased estimator. Such lower bound exists only for certain classes of estimators like the considered class of unbiased ones. Indeed the latter is the most used in the literature, mainly because 'good' estimators are asymptotically unbiased in general. Note also that this class of (asymptotically) unbiased estimators is suitable for our context as we consider large sample sizes for our fast or asymptotic CRB derivation.

The unconstrained CRB matrix (denoted **CRB**)[2] is computed as the inverse of the Fisher Information Matrix (FIM) (denoted $\mathbf{J}$ and assumed nonsingular in this subsection). When the parameter to estimate is complex-valued (i.e., $\boldsymbol{\theta} \in \mathbb{C}^P$, $P = M(L+1)+N$), the FIM $\mathbf{J} \in \mathbb{C}^{2P\times 2P}$ is defined as [14]–[16]

$$\mathbf{J} = \begin{bmatrix} \mathbf{J}_{\theta\theta} & \mathbf{J}_{\theta\theta^*} \\ \mathbf{J}_{\theta^*\theta} & \mathbf{J}_{\theta^*\theta^*} \end{bmatrix} = \begin{bmatrix} \mathbf{J}_{\theta\theta} & \mathbf{J}_{\theta\theta^*} \\ \mathbf{J}_{\theta\theta^*}^H & \mathbf{J}_{\theta\theta}^* \end{bmatrix}, \tag{4}$$

where $\mathbf{J}_{\theta\theta} = E\left(\Delta\Delta^H\right)$ and $\mathbf{J}_{\theta\theta^*} = E\left(\Delta\Delta^T\right)$, with $\Delta = \nabla_{\theta^*} \ln p(\mathbf{x};\boldsymbol{\theta})$, where $\nabla_{\theta^*} = \frac{\partial}{\partial\boldsymbol{\theta}^*} = [\frac{\partial}{\partial\theta_1^*}, \ldots, \frac{\partial}{\partial\theta_P^*}]^T$ is the complex gradient operator defined as in [17] and $\ln p(\mathbf{x};\boldsymbol{\theta})$ is the likelihood function. Note that the derivatives $\frac{\partial}{\partial\theta_i^*}$ are Wirtinger derivatives [18],[3] which are defined as $\frac{\partial}{\partial\theta_i^*} = \frac{1}{2}\left(\frac{\partial}{\partial\,\mathrm{Re}(\theta_i)} + j\frac{\partial}{\partial\,\mathrm{Im}(\theta_i)}\right)$, where $\mathrm{Re}(\cdot)$ and $\mathrm{Im}(\cdot)$ denote real-part and imaginary-part operators, respectively, and the derivatives on the right-hand-side are the usual partial derivatives taken with respect to real-valued quantities. Using the notation defined above, we can write

$$\mathbf{J} = E\left(\tilde{\Delta}\tilde{\Delta}^H\right),$$

where $\tilde{\Delta} = [\Delta^T, \Delta^H]^T$. When $\mathbf{x}$ is a noncircular complex Gaussian random vector, the FIM expression is given (elementwise) by [19]

$$(\mathbf{J})_{ij} = \left(\frac{\partial\tilde{\boldsymbol{\mu}}_x}{\partial\theta_i}\right)^H \tilde{\mathbf{R}}_x^{-1}\frac{\partial\tilde{\boldsymbol{\mu}}_x}{\partial\theta_j} + \frac{1}{2}\,\mathrm{Tr}\left[\frac{\partial\tilde{\mathbf{R}}_x}{\partial\theta_i}\tilde{\mathbf{R}}_x^{-1}\frac{\partial\tilde{\mathbf{R}}_x}{\partial\theta_j}\tilde{\mathbf{R}}_x^{-1}\right], \tag{5}$$

where

$$\tilde{\boldsymbol{\mu}}_x = \begin{bmatrix} \boldsymbol{\mu}_x \\ \boldsymbol{\mu}_x^* \end{bmatrix}, \text{ and } \quad \tilde{\mathbf{R}}_x = \begin{bmatrix} \mathbf{R}_x & \overline{\mathbf{R}}_x \\ \overline{\mathbf{R}}_x^* & \mathbf{R}_x^* \end{bmatrix},$$

with $\boldsymbol{\mu}_x = E(\mathbf{x})$ is the expected value of $\mathbf{x}$, $\mathbf{R}_x = E\left((\mathbf{x}-\boldsymbol{\mu}_x)(\mathbf{x}-\boldsymbol{\mu}_x)^H\right)$ is the covariance matrix of $\mathbf{x}$, and $\overline{\mathbf{R}}_x = E\left((\mathbf{x}-\boldsymbol{\mu}_x)(\mathbf{x}-\boldsymbol{\mu}_x)^T\right)$ is the pseudo-covariance matrix of $\mathbf{x}$ [20]. Because the noise is assumed circular Gaussian in (1), we have $\boldsymbol{\mu}_x = \mathcal{H}\mathbf{s} = \mathcal{S}\mathbf{h}$, $\tilde{\mathbf{R}}_x = \sigma^2\mathbf{I}$, and (5) simplifies to

$$(\mathbf{J})_{ij} = \frac{1}{\sigma^2}\left[\left(\frac{\partial\boldsymbol{\mu}_x}{\partial\theta_i}\right)^H\frac{\partial\boldsymbol{\mu}_x}{\partial\theta_j} + \left(\frac{\partial\boldsymbol{\mu}_x^*}{\partial\theta_i}\right)^H\frac{\partial\boldsymbol{\mu}_x^*}{\partial\theta_j}\right]. \tag{6}$$

From (6), we can see that $(\mathbf{J}_{\theta\theta^*})_{i,j} = 0$ and that $(\mathbf{J}_{\theta\theta})_{i,j} = \frac{1}{\sigma^2}\left(\frac{\partial\boldsymbol{\mu}_x}{\partial\theta_i}\right)^H\frac{\partial\boldsymbol{\mu}_x}{\partial\theta_j}$. This indicates that the knowledge of $\mathbf{J} \in \mathbb{C}^{2P\times 2P}$ reduces to the knowledge of $\mathbf{J}_{\theta\theta} \in \mathbb{C}^{P\times P}$. Therefore, the FIM for the unconstrained estimation of $\boldsymbol{\theta}$ for our model (2) is block diagonal given by

$$\mathbf{J} = \begin{bmatrix} \mathbf{J}_{\theta\theta} & \mathbf{0} \\ \mathbf{0} & \mathbf{J}_{\theta\theta}^* \end{bmatrix}, \tag{7}$$

where

$$\mathbf{J}_{\theta\theta} = \frac{1}{\sigma^2}\begin{bmatrix} \mathcal{S}^H\mathcal{S} & \mathcal{S}^H\mathcal{H} \\ \mathcal{H}^H\mathcal{S} & \mathcal{H}^H\mathcal{H} \end{bmatrix} = \frac{1}{\sigma^2}\begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{B}^H & \mathbf{C} \end{bmatrix}. \tag{8}$$

For notational simplicity matrices $\mathbf{A}$, $\mathbf{B}$, $\mathbf{B}^H$, and $\mathbf{C}$ are used instead of $\mathcal{S}^H\mathcal{S}$, $\mathcal{S}^H\mathcal{H}$, $\mathcal{H}^H\mathcal{S}$, and $\mathcal{H}^H\mathcal{H}$, respectively. The unconstrained CRB matrix

$$\mathbf{CRB} = \mathbf{J}^{-1} = \begin{bmatrix} \mathbf{J}_{\theta\theta}^{-1} & \mathbf{0} \\ \mathbf{0} & \left(\mathbf{J}_{\theta\theta}^{-1}\right)^* \end{bmatrix}, \tag{9}$$

is completely determined by $\mathbf{J}_{\theta\theta}^{-1}$ (assumed nonsingular). If $\mathbf{J}$ is singular,[4] we consider a constrained estimation of $\boldsymbol{\theta}$, which leads to the constrained CRB presented in the next section.

B. CONSTRAINED CRB

As indicated before, the blind identification of (2) is possible only up to an unknown scalar. This indeterminacy leads to a singular FIM that cannot be directly inverted to obtain the CRB. To remove the indeterminacy, we resort to adding constraints on the parameters. Indeed, the indeterminacy reflects the existence of many solutions and the constraint is used here to target one (unique solution) among them. In the following, we show how to obtain the constrained CRB matrix (denoted **CCRB**) from the unconstrained FIM $\mathbf{J}$. The given derivation

[1] It can be shown that, in this context, knowing the noise variance or not does not affect the CRB of the desired parameters.

[2] We use a boldface font to emphasize the fact that it is a matrix.

[3] All derivatives in the sequel are assumed to be Wirtinger derivatives.

[4] Singularity occurs when full identification of the parameter vector is not possible (in our case due to the inherent scalar indeterminacy of the BSI problem).

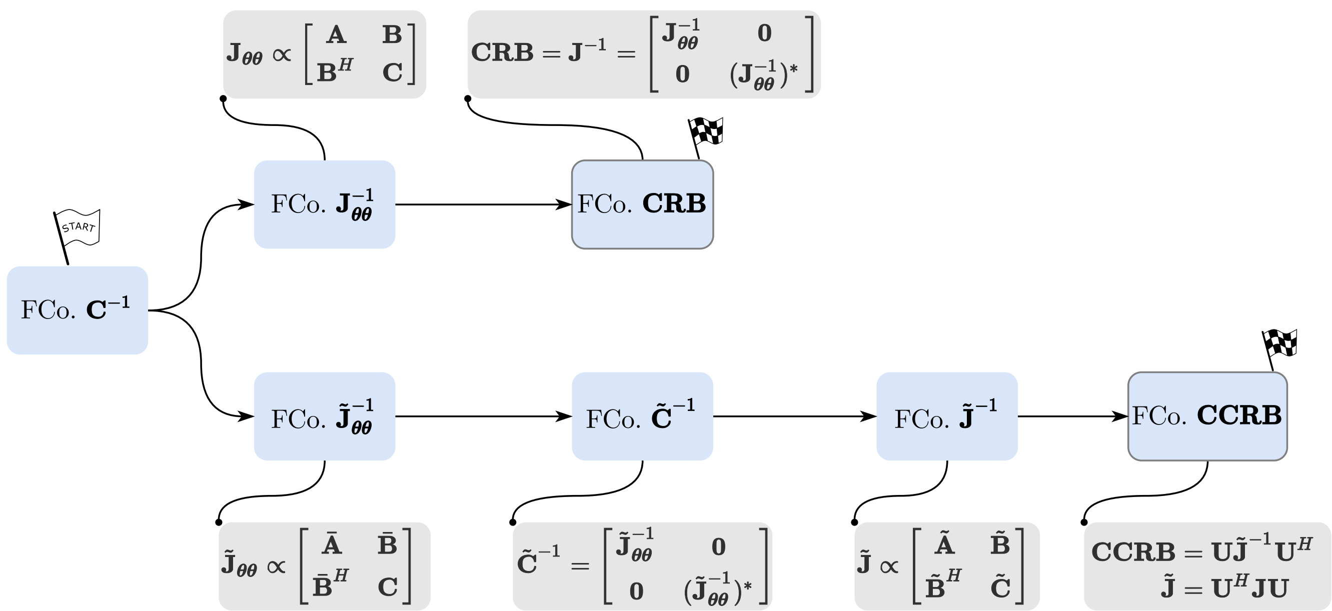


FIGURE 1. Flowchart for the fast computation of CRB and CCRB. The abbreviation "FCo." stands for "Fast computation of". An arrow from block *x* to block *y* should be read as "*x* enables *y*". Matrices with block elements A, B, and C (possibly with a bar or a tilde) hint to the particular structure where C is a square large-size matrix and is the most computationally heavy to invert, A is a square small-size matrix, and B is a rectangular low-rank matrix.

follows the one found in [15] and is given here for completeness. For complex-valued parameters, we need to consider the augmented parameter vector $\tilde{\boldsymbol{\theta}} = [\boldsymbol{\theta}^T, \boldsymbol{\theta}^H]^T$ of size $2P \times 1$.[5] The set of K constraints imposed on the parameter estimation to remove the indeterminacy can be written as

$$\mathbf{g}(\tilde{\boldsymbol{\theta}}) = 0,$$

where $\mathbf{g}(\tilde{\boldsymbol{\theta}})$ is of size $K \times 1$. We then define an augmented constraint set written as

$$\mathbf{f}(\tilde{\boldsymbol{\theta}}) = \begin{bmatrix} \mathbf{g}(\tilde{\boldsymbol{\theta}}) \\ \mathbf{g}(\tilde{\boldsymbol{\theta}})^* \end{bmatrix} = 0,$$

where $\mathbf{f}(\tilde{\boldsymbol{\theta}})$ is of size $2K \times 1$. We also define $\mathbf{F}(\tilde{\boldsymbol{\theta}}) \in \mathbb{C}^{2K \times 2P}$ (assumed to have full row rank) as

$$\mathbf{F}(\tilde{\boldsymbol{\theta}}) = \frac{\partial \mathbf{f}(\tilde{\boldsymbol{\theta}})}{\partial \tilde{\boldsymbol{\theta}}} = \left[\frac{\partial \mathbf{f}(\tilde{\boldsymbol{\theta}})}{\partial \boldsymbol{\theta}}, \frac{\partial \mathbf{f}(\tilde{\boldsymbol{\theta}})}{\partial \boldsymbol{\theta}^*} \right],$$

where

$$\frac{\partial \mathbf{f}(\tilde{\boldsymbol{\theta}})}{\partial \boldsymbol{\theta}} = \begin{bmatrix} \frac{\partial f_1}{\partial \theta_1} & \cdots & \frac{\partial f_1}{\partial \theta_P} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_{2K}}{\partial \theta_1} & \cdots & \frac{\partial f_{2K}}{\partial \theta_P} \end{bmatrix}.$$

If $r = \operatorname{rank}\left(\mathbf{F}(\tilde{\boldsymbol{\theta}})\right)$, then we can find a matrix $\mathbf{U} \in \mathbb{C}^{2P \times (2P-r)}$ with columns that form an orthonormal basis for the nullspace of $\mathbf{F}(\tilde{\boldsymbol{\theta}})$, i.e.,

$$\mathbf{F}(\tilde{\boldsymbol{\theta}})\mathbf{U} = 0.$$

Finally, the constrained CRB matrix is defined as

$$\mathbf{CCRB} = \mathbf{U}(\mathbf{U}^H\mathbf{J}\mathbf{U})^{-1}\mathbf{U}^H. \tag{10}$$

[5] This is similar to the need for an augmented FIM (4) for complex-valued parameters.

IV. FAST CRB AND CCRB COMPUTATIONS

Because we consider $\mathbf{h}$ and $\mathbf{s}$ as deterministic unknowns, the vector $\boldsymbol{\theta} = [\mathbf{h}^T, \mathbf{s}^T]^T$ of desired parameters contains all the values of the channel coefficients and input samples. The excessive computational cost of the CRB matrices comes from the inversion of $\mathbf{J} \in \mathbb{C}^{2P \times 2P}$ for the **CRB** and from the inversion of $\tilde{\mathbf{J}} = \mathbf{U}^H\mathbf{J}\mathbf{U} \in \mathbb{C}^{(2P-r)\times(2P-r)}$ for the **CCRB** (see (9) and (10)), especially when the sample size N is large; recall that $P = M(L+1) + N$. We show hereafter how does the fast computation of $\mathbf{C}^{-1}$ (see (8)) enable the fast computation of both **CRB** and **CCRB**.

The proposed computational cost of the proposed computation is $\mathcal{O}\left(N^2ML\right)$ instead of $\mathcal{O}\left((M(L+1)+N)^3\right)$ resulting from a direct matrix inversion using, for example, Gauss-Jordan elimination.

A. BASIC IDEA

Fig. 1 shows a flowchart for the fast computation of **CRB** and **CCRB**, the details of which are given in the sequel. Both computations depend on the fast computation of matrix $\mathbf{C}^{-1}$. Also, we note the recurrent need for the inversion of matrices of the form

$$\begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{B}^H & \mathbf{C} \end{bmatrix},$$

where block elements $\mathbf{A}$, $\mathbf{B}$, and $\mathbf{C}$ (possibly with a bar or a tilde) are such that $\mathbf{C}$ is a square large-size matrix, $\mathbf{A}$ is a square small-size matrix, and $\mathbf{B}$ is a rectangular low-rank matrix. This structure suggests (through the use of block matrix inversion; see Proposition 1) that the fast computation of the inverse of such matrices is enabled through the fast computation of matrix $\mathbf{C}^{-1}$ (or $\tilde{\mathbf{C}}^{-1}$), which is the most computationally heavy to invert. Moreover, note the similarity between the expression of $\tilde{\mathbf{C}}^{-1}$ and its block elements (details given in the sequel) and the expression of **CRB**. This similarity will allow us to use the same procedure for computing $\tilde{\mathbf{C}}^{-1}$ and **CRB**.

The flowchart together with the derivation steps given in the following subsections represent, for the interested reader, the pseudo-code of our fast CRB computation algorithm.

Before proceeding further, we recall here a result related to block matrix inversion, which we will be using often in the sequel.

Proposition 1: Block matrix inversion[6]

Let $\mathbf{A} \in \mathbb{C}^{n\times n}$, $\mathbf{B} \in \mathbb{C}^{n\times m}$, and $\mathbf{C} \in \mathbb{C}^{m\times m}$. If $\mathbf{C}$ and $\mathbf{D} = \mathbf{A} - \mathbf{B}\mathbf{C}^{-1}\mathbf{B}^H$ are nonsingular, then

$$\begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{B}^H & \mathbf{C} \end{bmatrix}^{-1} = \begin{bmatrix} \mathbf{D}^{-1} & -\mathbf{D}^{-1}\mathbf{B}\mathbf{C}^{-1} \\ -\mathbf{C}^{-1}\mathbf{B}^H\mathbf{D}^{-1} & \mathbf{C}^{-1} + \mathbf{C}^{-1}\mathbf{B}^H\mathbf{D}^{-1}\mathbf{B}\mathbf{C}^{-1} \end{bmatrix}. \quad (11)$$

As shown in Fig. 1, the computation of the CRBs depend directly (for **CRB**) or indirectly (for **CCRB** through $\tilde{\mathbf{C}}^{-1}$) on the computation of matrix $\mathbf{C}^{-1}$. On one hand, $\mathbf{C}^{-1}$ has the largest size and is the most computationally heavy when computing **CRB**. On the other hand, $\mathbf{C}^{-1}$ is the most computationally heavy when computing $\tilde{\mathbf{C}}^{-1}$, which in turn is the most computationally heavy in the computation of **CCRB**. Therefore, a fast computation of $\mathbf{C}^{-1}$ should allow a fast computation of both **CRB** and **CCRB**. (The diagonal blocks of these matrices are of particular interest since the trace of the CRB matrix is a lower bound on the MSE of any unbiased estimator of $\boldsymbol{\theta}$.) The basic idea for the proposed fast computation consists in expressing $\mathbf{C}^{-1}$ in a form which involves the use of circulant matrices (Toeplitz matrices with 'wraparound' [23, p. 220]); known for having a particular decomposition that uses the Discrete Fourier Matrix (DFT) for which fast computations are possible using the fast Fourier Transform (FFT) algorithm. The details for the fast computation of $\mathbf{C}^{-1}$ are provided in Section IV-C.

Remark: In our work, we implicitly assumed the source signal to be zero before the initial time instant $n = 0$. If it were not the case, the channel matrix (3) would have a Sylvester structure as shown in [24]. However, the basic idea for the fast CRB computation can still be used for this latter context but with a slight modification of the derivation details.

B. EXPRESSION OF THE CCRB

In this section, we give the expressions of the different matrices shown in Fig. 1 that are involved in the computation of **CCRB**. We chose two constraints to remove the complex scalar indeterminacy defined, following the presentation of Section III-B, by vector $\mathbf{g}(\tilde{\boldsymbol{\theta}}) = [\mathbf{h}^T\mathbf{h}^* - 1, h_{i_0} - h_{i_0}^*]^T = 0$, i.e., a unit-norm constraint and element h_{i_0} constrained to be real-valued. In this case, we have the augmented constraint vector $\mathbf{f}(\tilde{\boldsymbol{\theta}}) = \mathbf{g}(\tilde{\boldsymbol{\theta}})$; since $\mathbf{g}(\tilde{\boldsymbol{\theta}})^*$ gives a redundant set of constraints that are discarded for $\mathbf{F}(\tilde{\boldsymbol{\theta}})$ to have full row-rank. This latter matrix is defined as

$$\mathbf{F}(\tilde{\boldsymbol{\theta}}) = \begin{bmatrix} \mathbf{F}_1 & \mathbf{F}_2 \end{bmatrix},$$

where

$$\mathbf{F}_1 = \begin{bmatrix} \mathbf{h}^H & \mathbf{0} \\ \mathbf{e}_{i_0}^T & \mathbf{0} \end{bmatrix}, \quad \mathbf{F}_2 = \begin{bmatrix} \mathbf{h}^T & \mathbf{0} \\ -\mathbf{e}_{i_0}^T & \mathbf{0} \end{bmatrix}.$$

We define matrix[7] $\mathbf{U}$, for which we have $\mathbf{F}(\tilde{\boldsymbol{\theta}})\mathbf{U} = \mathbf{0}$, as

$$\mathbf{U} = \begin{bmatrix} \mathbf{K}_1 & \mathbf{U}_1 & \mathbf{0} \\ \mathbf{K}_2 & \mathbf{0} & \mathbf{U}_1^* \end{bmatrix} = \mathbf{0},$$

where $\mathbf{U}_1 \in \mathbb{C}^{P\times(P-2)}$ is such that $\mathbf{F}_1\mathbf{U}_1 = \mathbf{0}$ and $\mathbf{F}_2\mathbf{U}_1^* = \mathbf{0}$. Matrices $\mathbf{K}_1 \in \mathbb{C}^{P\times 2}$ and $\mathbf{K}_2 \in \mathbb{C}^{P\times 2}$ complete the number of columns of $\mathbf{U}$ to $2P-2$ columns, where they introduce two columns that are orthogonal to the columns of matrix $\operatorname{diag}\left([\mathbf{U}_1, \mathbf{U}_1^*]\right)$. Taking into account (7) and $\mathbf{U}$ we can write

$$\tilde{\mathbf{J}} = \mathbf{U}^H\mathbf{J}\mathbf{U} = \begin{bmatrix} \tilde{\mathbf{A}} & \tilde{\mathbf{B}} \\ \tilde{\mathbf{B}}^H & \tilde{\mathbf{C}} \end{bmatrix},$$

where $\tilde{\mathbf{A}} \in \mathbb{C}^{2\times 2}$, $\tilde{\mathbf{B}} \in \mathbb{C}^{2\times(2P-4)}$, and $\tilde{\mathbf{C}} \in \mathbb{C}^{(2P-4)\times(2P-4)}$ are defined as $\tilde{\mathbf{A}} = \mathbf{K}_1^H\mathbf{J}_{\boldsymbol{\theta\theta}}\mathbf{K}_1 + \mathbf{K}_2^H\mathbf{J}_{\boldsymbol{\theta\theta}}^*\mathbf{K}_2$, $\tilde{\mathbf{B}} = \left[\mathbf{K}_1^H\mathbf{J}_{\boldsymbol{\theta\theta}}\mathbf{U}_1, \mathbf{K}_2^H\mathbf{J}_{\boldsymbol{\theta\theta}}^*\mathbf{U}_1^*\right]$, and

$$\tilde{\mathbf{C}} = \begin{bmatrix} \tilde{\mathbf{J}}_{\boldsymbol{\theta\theta}} & \mathbf{0} \\ \mathbf{0} & \tilde{\mathbf{J}}_{\boldsymbol{\theta\theta}}^* \end{bmatrix},$$

where $\tilde{\mathbf{J}}_{\boldsymbol{\theta\theta}} = \mathbf{U}_1^H\mathbf{J}_{\boldsymbol{\theta\theta}}\mathbf{U}_1$. Using (11) to find the inverse of $\tilde{\mathbf{J}}$ shows that $\tilde{\mathbf{C}}^{-1}$ is the most computationally heavy. In fact, we can show that the cost of computing the other terms is at most $\mathcal{O}\left(N^2\right)$. Therefore, the cost of computing $\tilde{\mathbf{J}}^{-1}$ is of the same order as the computational cost of $\tilde{\mathbf{C}}^{-1}$ defined as

$$\tilde{\mathbf{C}}^{-1} = \begin{bmatrix} \left(\tilde{\mathbf{J}}_{\boldsymbol{\theta\theta}}\right)^{-1} & \mathbf{0} \\ \mathbf{0} & \left(\tilde{\mathbf{J}}_{\boldsymbol{\theta\theta}}^{-1}\right)^* \end{bmatrix}.$$

Due to the structure of matrix $\mathbf{F}_1$, matrix $\mathbf{U}_1$ defining $\tilde{\mathbf{J}}_{\boldsymbol{\theta\theta}}$ can be written as

$$\mathbf{U}_1 = \begin{bmatrix} \bar{\mathbf{U}}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_N \end{bmatrix},$$

where $\bar{\mathbf{U}}_1$ is the part of $\mathbf{U}_1$ that gets multiplied by the first $M(L+1)$ non-zero columns of $\mathbf{F}_1$. This leads to

$$\tilde{\mathbf{J}}_{\boldsymbol{\theta\theta}} = \frac{1}{\sigma^2}\begin{bmatrix} \bar{\mathbf{A}} & \bar{\mathbf{B}} \\ \bar{\mathbf{B}}^H & \mathbf{C} \end{bmatrix},$$

where $\bar{\mathbf{A}} \in \mathbb{C}^{(M(L+1)-2)\times(M(L+1)-2)}$, $\bar{\mathbf{B}} \in \mathbb{C}^{(M(L+1)-2)\times N}$, and $\mathbf{C} \in \mathbb{C}^{N\times N}$ are defined as $\bar{\mathbf{A}} = \bar{\mathbf{U}}_1^H\mathbf{A}\bar{\mathbf{U}}_1$, $\bar{\mathbf{B}} = \bar{\mathbf{U}}_1^H\mathbf{B}$, and $\mathbf{C} = \mathcal{H}^H\mathcal{H}$ (as defined in (8)). Using (11), we can write

$$\tilde{\mathbf{J}}_{\boldsymbol{\theta\theta}}^{-1} = \sigma^2\begin{bmatrix} \bar{\mathbf{D}}^{-1} & -\bar{\mathbf{D}}^{-1}\bar{\mathbf{B}}\mathbf{C}^{-1} \\ -\mathbf{C}^{-1}\bar{\mathbf{B}}^H\bar{\mathbf{D}}^{-1} & \mathbf{C}^{-1} + \mathbf{C}^{-1}\bar{\mathbf{B}}^H\bar{\mathbf{D}}^{-1}\bar{\mathbf{B}}\mathbf{C}^{-1} \end{bmatrix},$$

[6]This is a particular case of the general one described in [21], [22].

[7]We omit providing the explicit expressions of the matrix blocks of $\mathbf{U}$ as it has no impact on the computational cost.

where $\bar{\mathbf{D}} = \bar{\mathbf{A}} - \bar{\mathbf{B}}\mathbf{C}^{-1}\bar{\mathbf{B}}^H$. Fast computations of $\tilde{\mathbf{J}}_{\theta\theta}^{-1}$ (hence $\tilde{\mathbf{C}}^{-1}$) for **CCRB** and $\mathbf{J}_{\theta\theta}^{-1}$ for **CRB** follow the same procedure that starts by a fast computation of $\mathbf{C}^{-1}$ (presented in Section IV-C) followed by the computation of the remaining blocks of $\tilde{\mathbf{J}}_{\theta\theta}^{-1}$ for **CCRB** or $\mathbf{J}_{\theta\theta}^{-1}$ for **CRB** (presented in Section IV-D).

C. FAST COMPUTATION OF MATRIX $\mathbf{C}^{-1}$

Recall that $\mathbf{C} = \mathcal{H}^H\mathcal{H} = \sum_{i=1}^{M}\mathbf{H}_i^H\mathbf{H}_i$. The different steps involved in the fast computation of $\mathbf{C}^{-1}$ can be summarized as follows.

(a) Express $\mathbf{C}$ in the form $\mathbf{C} = \mathbf{C}_c + \mathbf{C}_r$, i.e., a sum of a circulant matrix $\mathbf{C}_c$ and a matrix $\mathbf{C}_r$ we call the remainder matrix of $\mathbf{C}$.

(b) Express $\mathbf{C}_r$ in the convenient form $\mathbf{C}_r = \mathbf{V}\boldsymbol{\Sigma}\mathbf{V}^H$ to enable the use of the Sherman-Morrison-Woodbury formula [23, p. 65] in order to put $\mathbf{C}^{-1} = \left(\mathbf{C}_c + \mathbf{V}\boldsymbol{\Sigma}\mathbf{V}^H\right)^{-1}$ in the form

$$\mathbf{C}^{-1} = \mathbf{C}_c^{-1} - \mathbf{C}_c^{-1}\mathbf{V}\left(\boldsymbol{\Sigma}^{-1} + \mathbf{V}^H\mathbf{C}_c^{-1}\mathbf{V}\right)^{-1} \times \mathbf{V}^H\mathbf{C}_c^{-1}. \tag{12}$$

(c) Leverage Theorem 1 for a fast computation of $\mathbf{C}_c^{-1}$ to allow the fast computation of the different terms involved in (12), all of which contain $\mathbf{C}_c^{-1}$; except $\boldsymbol{\Sigma}^{-1}$, which is a small-size, easily computed matrix.

Theorem 1 ([23, p. 222]): Let $\mathbf{Z}$ be a circulant matrix with $\mathbf{z} \in \mathbb{C}^N$ representing its first column. Then, $\mathbf{W}_N^{-1}\mathbf{Z}\mathbf{W}_N = \operatorname{diag}(\boldsymbol{\lambda})$, where $\boldsymbol{\lambda} = \mathbf{W}_N^H\mathbf{z}$.

$\mathbf{W}_N$ denotes the DFT matrix of size $N \times N$ defined elementwise as $(\mathbf{W}_N)_{ij} = \omega_N^{(i-1)(j-1)}$, where $\omega_N = \exp(\frac{-j2\pi}{N})$. Matrix $\mathbf{W}_N$ is symmetric, i.e., $\mathbf{W}_N^T = \mathbf{W}_N$, and is a (scaled) unitary matrix, i.e., $\mathbf{W}_N^H\mathbf{W}_N = N\mathbf{I}_N$. Therefore, $\mathbf{W}_N^{-1} = \frac{\mathbf{W}_N^H}{N}$.

Hereafter, we detail the different steps described above.

1) STEP(A)

By adding and subtracting matrix $\mathbf{H}_{ir}$ to $\mathbf{H}_i$ (see (3)), where $\mathbf{H}_{ir}$ is defined as

$$\mathbf{H}_{ir} = \begin{bmatrix} \mathbf{0}_{L\times(N-L)} & \boldsymbol{\Delta}_i \\ \mathbf{0}_{(N-L)\times(N-L)} & \mathbf{0}_{(N-L)\times L} \end{bmatrix}$$

such that

$$\boldsymbol{\Delta}_i = \begin{bmatrix} -h_i(L) & \cdots & -h_i(1) \\ \vdots & \ddots & \vdots \\ 0 & \cdots & -h_i(L) \end{bmatrix}_{L\times L},$$

we can write $\mathbf{H}_i = \mathbf{H}_{ic} + \mathbf{H}_{ir}$, where $\mathbf{H}_{ic} = \mathbf{H}_i - \mathbf{H}_{ir}$ is a circulant matrix. Since $\mathbf{C} = \sum_{i=1}^{M}\mathbf{H}_i^H\mathbf{H}_i$, we can write $\mathbf{C} = \mathbf{C}_c + \mathbf{C}_r$, where

$$\mathbf{C}_c = \sum_{i=1}^{M}\mathbf{H}_{ic}^H\mathbf{H}_{ic} \tag{13}$$

is a circulant matrix (sums and products of circulant matrices give circulant matrices)[8] and

$$\mathbf{C}_r = \sum_{i=1}^{M}\mathbf{H}_{ir}^H\mathbf{H}_{ir} + \mathbf{H}_{ic}^H\mathbf{H}_{ir} + \mathbf{H}_{ir}^H\mathbf{H}_{ic} \tag{14}$$

is the so-called remainder matrix of $\mathbf{C}$.

2) STEP(B)

Let $\mathbf{H}_{ic}(1:L,:)$ denote rows 1 to L of matrix $\mathbf{H}_{ic}$, and let $\mathbf{E}_i^H = \boldsymbol{\Delta}_i^H\mathbf{H}_{ic}(1:L,:)$ and $\mathbf{G}_i = \boldsymbol{\Delta}_i^H\boldsymbol{\Delta}_i$, then we have

$$\mathbf{C}_r = \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{G} \end{bmatrix} + \begin{bmatrix} \mathbf{0} & \mathbf{E} \end{bmatrix} + \begin{bmatrix} \mathbf{0} \\ \mathbf{E}^H \end{bmatrix},$$

where $\mathbf{G} = \sum_{i=1}^{M}\mathbf{G}_i = \mathbf{G}^H$, and $\mathbf{E} = \sum_{i=1}^{M}\mathbf{E}_i$. Now, if we let $\tilde{\mathbf{G}}^H = [\mathbf{0}, \frac{1}{\sqrt{2}}\mathbf{G}^{\frac{H}{2}}]$ and $\tilde{\mathbf{E}}^H = \sqrt{2}\mathbf{G}^{\frac{-H}{2}}\mathbf{E}^H + [\mathbf{0}, \frac{1}{\sqrt{2}}\mathbf{G}^{\frac{H}{2}}]$, where $\mathbf{G}^{\frac{-H}{2}} = (\mathbf{G}^{\frac{-1}{2}})^H$, then we can write (14) as

$$\begin{aligned} \mathbf{C}_r &= \tilde{\mathbf{G}}\tilde{\mathbf{E}}^H + \tilde{\mathbf{E}}\tilde{\mathbf{G}}^H \\ &= \begin{bmatrix} \tilde{\mathbf{G}} & \tilde{\mathbf{E}} \end{bmatrix} \begin{bmatrix} \mathbf{0} & \mathbf{I} \\ \mathbf{I} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \tilde{\mathbf{G}}^H \\ \tilde{\mathbf{E}}^H \end{bmatrix} \\ &= \mathbf{V}\boldsymbol{\Sigma}\mathbf{V}^H, \end{aligned} \tag{15}$$

where $\mathbf{V}$ is $N \times 2L$ and $\boldsymbol{\Sigma}$ is $2L \times 2L$ making $\mathbf{C}_r$ of rank $2L$ (i.e., a low-rank matrix). From (15), we get

$$\mathbf{C} = \mathbf{C}_c + \mathbf{V}\boldsymbol{\Sigma}\mathbf{V}^H. \tag{16}$$

Having $\mathbf{C}$ written in the from (16), enables us to use the Sherman-Morrison-Woodbury formula [23, p. 65], and write the inverse of (16) as

$$\mathbf{C}^{-1} = \mathbf{C}_c^{-1} - \mathbf{C}_c^{-1}\mathbf{V}\left(\boldsymbol{\Sigma}^{-1} + \mathbf{V}^H\mathbf{C}_c^{-1}\mathbf{V}\right)^{-1} \times \mathbf{V}^H\mathbf{C}_c^{-1}, \tag{17}$$

which allows a fast computation of (17) at step (c).

3) STEP(C)

Because $\mathbf{H}_{ic}$ is circulant, we can use Theorem 1 and the properties of the DFT matrix to write (13) as

$$\begin{aligned} \mathbf{C}_c &= \sum_{i=1}^{M}\mathbf{W}_N^{-H}\boldsymbol{\Lambda}_i^H\mathbf{W}_N^H\mathbf{W}_N\boldsymbol{\Lambda}_i\mathbf{W}_N^{-1} \\ &= \mathbf{W}_N\boldsymbol{\Lambda}\mathbf{W}_N^{-1}, \end{aligned}$$

which leads to

$$\mathbf{C}_c^{-1} = \mathbf{W}_N\boldsymbol{\Lambda}^{-1}\mathbf{W}_N^{-1}, \tag{18}$$

where $\boldsymbol{\Lambda} = \sum_{i=1}^{M}\boldsymbol{\Lambda}_i^H\boldsymbol{\Lambda}_i$, and $\boldsymbol{\Lambda}_i = \operatorname{diag}\left(\mathbf{W}_N^H\mathbf{c}_i\right)$, where $\mathbf{c}_i$ is the first column of (3) (same as the first column of $\mathbf{H}_{ic}$), i.e., $\mathbf{c}_i = [h_i(0), \ldots, h_i(L), \mathbf{0}_{1\times(N-L-1)}]^T$. Replacing (18) in (17) gives

$$\begin{aligned} \mathbf{C}^{-1} &= \mathbf{W}_N\boldsymbol{\Lambda}^{-1}\mathbf{W}_N^{-1} - \mathbf{W}_N\boldsymbol{\Lambda}^{-1}\mathbf{W}_N^{-1}\mathbf{U} \\ &\quad \times \left(\boldsymbol{\Sigma} + \mathbf{U}^H\mathbf{W}_N\boldsymbol{\Lambda}^{-1}\mathbf{W}_N^{-1}\mathbf{U}\right)^{-1} \end{aligned}$$

[8]Checking this fact is straightforward for the sum and is easily done for the product using Theorem 1.

TABLE 1. Summary of the computational cost of the different terms involved in (19).

Term identifier	Term	Computational cost
1	$\boldsymbol{\Lambda}^{-1}$	$\mathcal{O}(N\log N)$
2	$\mathbf{W}_N\boldsymbol{\Lambda}^{-1}\mathbf{W}_N^{-1}$	$\mathcal{O}(N\log N)$
3	$\mathbf{W}_N^{-1}\mathbf{U}$	$\mathcal{O}(LN\log N)$
4	$\left(\boldsymbol{\Sigma}+\mathbf{U}^H\mathbf{W}_N\boldsymbol{\Lambda}^{-1}\mathbf{W}_N^{-1}\mathbf{U}\right)^{-1}$	$\mathcal{O}(L^2N)$
5	$\mathbf{W}_N\boldsymbol{\Lambda}^{-1}\mathbf{W}_N^{-1}\mathbf{U}$	$\mathcal{O}(LN\log N)$
6	5 $\times$ 4 $\times$ 5^H	$\mathcal{O}(N^2L)$
7	2 + 6	$\mathcal{O}(N^2)$
—	$\mathbf{C}^{-1}$	$\mathcal{O}(N^2L)$

$$\times\,\mathbf{U}^H\mathbf{W}_N\boldsymbol{\Lambda}^{-1}\mathbf{W}_N^{-1}. \tag{19}$$

Because we are using the DFT matrix, computing $\boldsymbol{\Lambda}_i$ costs $\mathcal{O}(N\log N)$.[9] Computing $\boldsymbol{\Lambda}$ costs $\mathcal{O}(NM)$ (M sums of products of N diagonal elements) and $\boldsymbol{\Lambda}^{-1}$ (inverse of a $N\times N$ diagonal matrix) costs $\mathcal{O}(N)$. Therefore, the total cost for the computation of $\boldsymbol{\Lambda}^{-1}$ is $\mathcal{O}(N\log N)$ (the highest of the intermediary costs).[10] The computation of the remaining terms involved in (19) is conducted in a similar way. Table 1 summarizes the obtained (total) computational cost of the different terms of (19). We find that the overall cost of computing $\mathbf{C}^{-1}$ is $\mathcal{O}\left(N^2L\right)$.

TABLE 2. Summary of the computational cost of the different terms of $\mathbf{J}_{\boldsymbol{\theta\theta}}^{-1}$, or equivalently $\tilde{\mathbf{J}}_{\boldsymbol{\theta\theta}}^{-1}$.

Term identifier	Term	Computational cost
1	$\mathbf{BC}^{-1}\mathbf{B}^H$	$\mathcal{O}(NM^2L^2)$
2	$(\mathbf{A}-$ 1 $)^{-1}$	$\mathcal{O}(M^3L^3)$
3	$\mathbf{BC}^{-1}$	$\mathcal{O}(N^2ML)$
4	2 $\times$ 3	$\mathcal{O}(NM^2L^2)$
5	3^H $\times$ 4	$\mathcal{O}(N^2ML)$
6	$\mathbf{C}^{-1}+$ 5	$\mathcal{O}(N^2)$
—	$\mathbf{J}_{\boldsymbol{\theta\theta}}^{-1}$	$\mathcal{O}(N^2ML)$

D. FAST COMPUTATION OF CRB (OR $\tilde{\mathbf{C}}^{-1}$)

As shown in Fig. 1, the fast computation of **CRB** (resp. $\tilde{\mathbf{C}}^{-1}$) is enabled through the fast computation of $\mathbf{J}_{\boldsymbol{\theta\theta}}^{-1}$ (resp. $\tilde{\mathbf{J}}_{\boldsymbol{\theta\theta}}^{-1}$), which in turn is enabled through the fast computation of $\mathbf{C}^{-1}$ presented in Section IV-C. Based on (11), we give in Table 2 the computational cost of the different terms involved in computing $\mathbf{J}_{\boldsymbol{\theta\theta}}^{-1}$. The computation of $\tilde{\mathbf{J}}_{\boldsymbol{\theta\theta}}^{-1}$ generates the same cost as $\mathbf{J}_{\boldsymbol{\theta\theta}}^{-1}$, since the sizes of the matrix blocks of $\tilde{\mathbf{J}}_{\boldsymbol{\theta\theta}}^{-1}$ are only reduced by two rows or two columns compared to the sizes of the block matrices of $\mathbf{J}_{\boldsymbol{\theta\theta}}^{-1}$. We find in the end that the total computation cost of $\mathbf{J}_{\boldsymbol{\theta\theta}}^{-1}$, hence **CRB**, is $\mathcal{O}\left(N^2ML\right)$.

To illustrate the computational gain we get via the proposed fast computation method, we provide in Table 3 a numerical example highlighting the computational complexity as function of the sample size N.

[9]We use a base-2 logarithm.

[10]For simplicity, we assumed that $M\le\log N$ so that only the dominant cost terms are given in Table 1.

TABLE 3. Comparison between our fast computation ($\mathcal{O}\left(N^2ML\right)$) and a direct computation ($\mathcal{O}\left((M(L+1)+N)^3\right)$) of the CRB for different values of N with $L=4$ and $M=2$.

N	100	500	1000	5000
$\mathcal{O}(N^2ML)$	$\mathcal{O}(8.00\times10^5)$	$\mathcal{O}(2.00\times10^7)$	$\mathcal{O}(8.00\times10^7)$	$\mathcal{O}(2.00\times10^9)$
$\mathcal{O}((M(L+1)+N)^3)$	$\mathcal{O}(1.33\times10^6)$	$\mathcal{O}(1.32\times10^8)$	$\mathcal{O}(1.03\times10^9)$	$\mathcal{O}(1.25\times10^{11})$

V. ASYMPTOTIC CRB COMPUTATION

When the sample size is very large, the inversion of the FIM might become prohibitively expensive even with the fast computation proposed in this paper. In such case, it is useful to consider an asymptotic approximation of the FIM (i.e., an approximation for large sample size N). The goal here is to provide a simple formula that helps getting some insight on the estimation performance in the asymptotic case.

For this asymptotic approximation we consider a stochastic perspective, where the input signal is assumed random rather than deterministic. By adapting the multivariate Whittle formula proposed in [25, Eq. (6.3)] to handle complex variables, we get the following expression for the (elementwise) asymptotic FIM

$$[\mathbf{J}_a]_{ij}=\frac{N}{2}\int_{-1/2}^{1/2}\mathrm{Tr}\Bigg[\frac{\partial P_{xx}(f;\boldsymbol{\theta})}{\partial\theta_i^*}P_{xx}(f;\boldsymbol{\theta})^{-1}\times\left(\frac{\partial P_{xx}(f;\boldsymbol{\theta})}{\partial\theta_j^*}\right)^H P_{xx}(f;\boldsymbol{\theta})^{-1}\Bigg]df, \tag{20}$$

where $P_{xx}(f;\boldsymbol{\theta})$ is the power spectral density of observation $\mathbf{x}$. Assuming, for example, a white input signal with auto-correlation $R_{ss}(\tau)=\sigma_s^2\delta(\tau)$ (where $\delta(\tau)$ is the Dirac delta function, i.e., a function that equals 1 at 0 and equals 0 elsewhere), we get

$$P_{xx}(f;\boldsymbol{\theta})=\sigma_s^2\mathbf{h}(f)\mathbf{h}(f)^H+\sigma^2\mathbf{I},$$

where $\mathbf{h}(f)=[h_1(f),\ldots,h_M(f)]^T$ with $h_i(f)=\sum_{n=0}^{L}h_i[n]e^{-j2\pi fn}$. Applying the Sherman-Morrison formula we find that

$$P_{xx}(f;\boldsymbol{\theta})^{-1}=\frac{1}{\sigma^2}\left(\mathbf{I}-\frac{\sigma_s^2\mathbf{h}(f)\mathbf{h}(f)^H}{\sigma^2+\sigma_s^2\mathbf{h}(f)^H\mathbf{h}(f)}\right).$$

We propose to solve (20) in the $\mathcal{Z}$-domain, which gives

$$\begin{aligned}[\mathbf{J}_a]_{ij}&=\frac{N}{4\pi j}\oint_{|z|=1}\mathrm{Tr}\Bigg[\frac{\partial P_{xx}(z;\boldsymbol{\theta})}{\partial\theta_i^*}P_{xx}(z;\boldsymbol{\theta})^{-1}\times\left(\frac{\partial P_{xx}(z;\boldsymbol{\theta})}{\partial\theta_j^*}\right)^H P_{xx}(z;\boldsymbol{\theta})^{-1}\Bigg]\frac{dz}{z}\\&=\frac{N}{4\pi j}\oint_{|z|=1}f(z)dz,\end{aligned} \tag{21}$$

where $f(z)$ is the function within the integral, which simplifies to (assuming $\theta_i=h_i[l]$ and $\theta_j=h_j[m]$)

$$f(z)=\frac{\sigma_s^4\,z^{l-m-1}}{\sigma^2}\left(\delta(i-j)-\frac{h_i(z)h_j(z)^*}{\sigma^2+\sigma_s^2\mathbf{h}(z)^H\mathbf{h}(z)}\right)$$

$$\times \frac{\mathbf{h}(z)^H \mathbf{h}(z)}{\sigma^2 + \sigma_s^2 \mathbf{h}(z)^H \mathbf{h}(z)},$$

and

$$P_{xx}(z; \boldsymbol{\theta}) = \sigma_s^2 \mathbf{h}(z)\mathbf{h}(z^{-*})^H + \sigma^2 \mathbf{I},$$

where $\mathbf{h}(z) = [h_1(z), \ldots, h_M(z)]^T$ with $h_i(z) = \sum_{n=0}^{L} h_i[n]z^{-n}$, and $\mathbf{h}(z^{-*}) = \mathbf{h}(z)$ since $z^{-*} = z$ when evaluated along the unit-circle, and

$$P_{xx}(z; \boldsymbol{\theta})^{-1} = \frac{1}{\sigma^2}\left(\mathbf{I} - \frac{\sigma_s^2 \mathbf{h}(z)\mathbf{h}(z)^H}{\sigma^2 + \sigma_s^2 \mathbf{h}(z)^H \mathbf{h}(z)}\right).$$

The integral (21) can be computed using the residue theorem [11] by writing

$$[\mathbf{J}_a]_{ij} = \frac{N}{2} \sum_k \text{residue}\left(f(z), p_k\right)\Big|_{|p_k|<1},$$

where residue $(f(z), p_k)\,\big|_{|p_k|<1}$ denotes the residue of function $f(z)$ at pole p_k verifying $|p_k| < 1$ (i.e., inside the unit-circle).

Note that $\mathbf{J}_a$ is an asymptotic approximation for the term $\mathbf{D} = \mathbf{A} - \mathbf{B}\mathbf{C}^{-1}\mathbf{B}^H$ we find (inverted) in the upper-left corner of the right-hand-side of (11). Assuming the matrix to invert is (8), this term represents the unconstrained CRB matrix of the channel coefficients. Note also that the diagonal element $[\mathbf{J}_a]_{ii}$ can be considered as an indicator for the estimation performance of element θ_i (a particular channel coefficient) assuming other parameters are known (favorable case). Therefore, a possible indicator for a bad diversity would be

$$I = \min_i [\mathbf{J}_a]_{ii}.$$

If I is 'small', then this indicates a bad diversity condition.

Figs. 2 and 3 illustrate the closeness of the asymptotic FIM expression in (20) to the exact FIM expression for large sample sizes. The plots represent the trace of the corresponding CRB matrices versus the sample size and show the effectiveness of the asymptotic Whittle approximation. The latter allows us to overcome the cumbersome computation of the FIM and its inverse when the data observation period is relatively large. The results were generated using the following parameters: $M = 2$, $\sigma_s^2 = 1$, ten Monte-Carlo runs for each value of the sample size N, and (complex Gaussian) randomly generated channels and source signal. In addition, Fig. 2 gives a comparison for two values of the channel order: $L = 4$ and $L = 6$ at a signal-to-noise ratio (SNR) of 10 dB, whereas Fig. 3 gives a comparison for two values of the SNR: 0 dB and 10 dB, for a fixed channel order $L = 4$. We notice that attaining the exact CRB is faster for low channel orders (zoom in Fig. 2) and high SNRs (zoom in Fig. 3) which illustrates the fact that the rate of convergence depends on the parameters of the system under consideration. Finally, in all simulated scenarios, we observe that the asymptotic CRB seems to be more optimistic (i.e. lower) than the exact one.

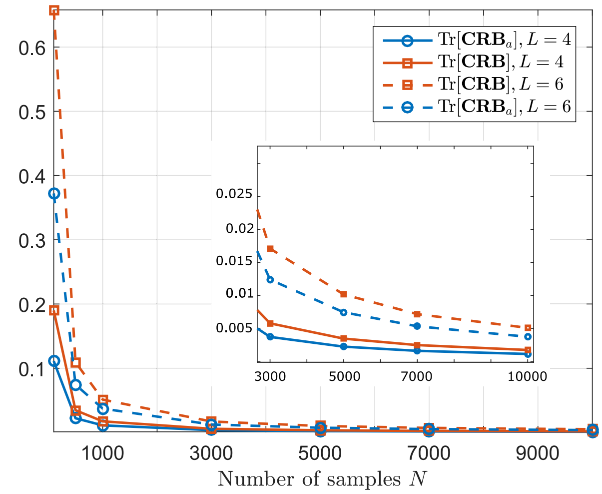


FIGURE 2. Traces of the asymptotic (blue curves) and exact (red curves) CCRB matrices versus the sample size N for channel orders $L = 4$ (solid curves) and $L = 6$ (dashed curves) at SNR = 10 dB.

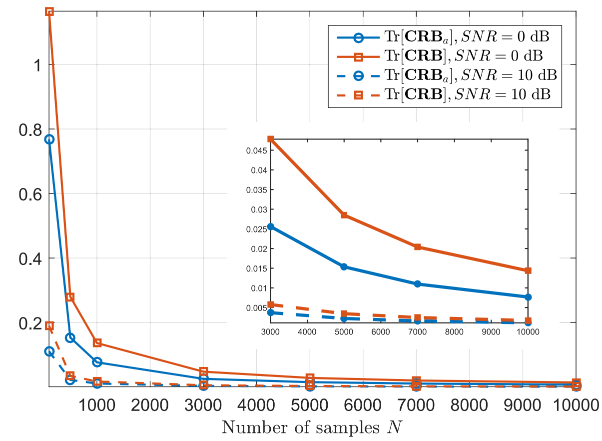


FIGURE 3. Traces of the asymptotic (blue curves) and exact (red curves) CCRB matrices versus the sample size N for SNR values 0 dB (solid curves) and 10 dB (dashed curves), and channel order $L = 4$.

VI. CONCLUSION

The Gaussian CRB is of great interest in signal processing for many purposes and application fields. Its computation in the deterministic case, however, can be heavy or even impractical as it involves the inversion of a matrix which dimension grows linearly with the sample size. We propose in this paper a fast implementation of the deterministic Gaussian CRB in the context of multi-channel blind system identification. This implementation preserves the exact CRB formula while reducing the numerical computation from a cubic to a quadratic cost with respect to the parameter vector dimension. In addition, we propose an asymptotic (approximate) expression of the stochastic Gaussian CRB, which has the advantage of compactness, easy derivation, and applicability for large sample sizes. An illustrative numerical example is provided to show that the asymptotic regime is reachable when the sample size is sufficiently large. Finally, we show how the asymptotic FIM expression can be used as an indicator on how difficult the channel identification problem at hand is.

As perspective, a generalization using the same basic ideas of this work, can be considered for the Multiple Input Multiple Output (MIMO) or semi-blind cases.

REFERENCES

[1] E. Grosicki, K. Abed-Meraim, and Y. Hua, ''A weighted linear prediction method for near-field source localization,'' *IEEE Trans. Signal Process.*, vol. 53, no. 10, pp. 3651–3660, Oct. 2005.

[2] Y. Begriche, M. Thameri, and K. Abed-Meraim, ''Exact cramer rao bound for near field source localization,'' in *Proc. 11th Int. Conf. Inf. Sci., Signal Process. Appl. (ISSPA)*, Jul. 2012, pp. 718–721.

[3] F. Ahmed Sid, K. Abed-Meraim, R. Harba, and F. Oulebsir-Boumghar, ''Analytical performance bounds for multi-tensor diffusion-MRI,'' *Magn. Reson. Imag.*, vol. 36, pp. 146–158, Feb. 2017.

[4] H. Gazzah and K. Abed-Meraim, ''Optimum ambiguity-free directional and omnidirectional planar antenna arrays for DOA estimation,'' *IEEE Trans. Signal Process.*, vol. 57, no. 10, pp. 3942–3953, Oct. 2009.

[5] S. M. Kay, *Fundamentals of Statistical Signal Processing*. Upper Saddle River, NJ, USA: Prentice-Hall, 1993.

[6] P. Stoica and B. Ng, ''Performance bounds for blind channel estimation,'' in *Signal Processing Advances in Wireless & Mobile Communications*, vol. 1, G. B. Giannakis, Y. Hua, P. Stoica, and L. Tong, Eds. Upper Saddle River, NJ, USA: Prentice-Hall, 2001, pp. 41–62.

[7] S. Park, E. Serpedin, and K. Qaraqe, ''Gaussian assumption: The least favorable but the most useful [Lecture Notes],'' *IEEE Signal Process. Mag.*, vol. 30, no. 3, pp. 183–186, May 2013.

[8] E. Moulines, P. Duhamel, J.-F. Cardoso, and S. Mayrargue, ''Subspace methods for the blind identification of multichannel FIR filters,'' *IEEE Trans. Signal Process.*, vol. 43, no. 2, pp. 516–525, Feb. 1995.

[9] K. Abed-Meraim, E. Moulines, and P. Loubaton, ''Prediction error method for second-order blind identification,'' *IEEE Trans. Signal Process.*, vol. 45, no. 3, pp. 694–705, Mar. 1997.

[10] H. Liu, G. Xu, and L. Tong, ''A deterministic approach to blind identification of multi-channel FIR systems,'' in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, vol. 4, Apr. 1994, pp. 581–584.

[11] L. V. Ahlfors, *Complex Analysis: An Introduction to the Theory of Analytic Functions of One Complex Variable*. New York, NY, USA: McGraw-Hill, 1979.

[12] K. Abed-Meraim, W. Qiu, and Y. Hua, ''Blind system identification,'' *Proc. IEEE*, vol. 85, no. 8, pp. 1310–1322, Aug. 1997.

[13] Y. Hua, ''Blind methods of system identification,'' *Circuits, Syst., Signal Process.*, vol. 21, no. 1, pp. 91–108, Jan. 2002.

[14] E. de Carvalho, J. Cioffi, and D. Slock, ''Cramer-rao bounds for blind multichannel estimation,'' in *Proc. Globecom IEEE. Global Telecommun. Conf. Rec.*, vol. 2, Nov./Dec. 2000, pp. 1036–1040.

[15] A. K. Jagannatham and B. D. Rao, ''Cramer-rao lower bound for constrained complex parameters,'' *IEEE Signal Process. Lett.*, vol. 11, no. 11, pp. 875–878, Nov. 2004.

[16] E. Ollila, V. Koivunen, and J. Eriksson, ''On the Cramér-rao bound for the constrained and unconstrained complex parameters,'' in *Proc. 5th IEEE Sensor Array Multichannel Signal Process. Workshop*, Jul. 2008, pp. 414–418.

[17] D. H. Brandwood, ''A complex gradient operator and its application in adaptive array theory,'' *IEE Proc. H-Microw., Opt. Antennas*, vol. 130, no. 1, pp. 11–16, Feb. 1983.

[18] W. Wirtinger, ''Zur formalen theorie der funktionen von mehr komplexen veränderlichen,'' *Math. Annalen*, vol. 97, no. 1, pp. 357–375, 1927.

[19] J.-P. Delmas and H. Abeida, ''Stochastic cramér-rao bound for noncircular signals with application to doa estimation,'' *IEEE Trans. Signal Process.*, vol. 52, no. 11, pp. 3192–3199, Nov. 2004.

[20] F. D. Neeser and J. L. Massey, ''Proper complex random processes with applications to information theory,'' *IEEE Trans. Inf. Theory*, vol. 39, no. 4, pp. 1293–1302, Jul. 1993.

[21] W. J. Duncan, ''LXXVIII. Some devices for the solution of large sets of simultaneous linear equations: With an appendix on the reciprocation of partitioned matrices,'' *London, Edinburgh, Dublin Philos. Mag. J. Sci.*, vol. 35, no. 249, pp. 660–670, 1944.

[22] W. W. Hager, ''Updating the inverse of a matrix,'' *SIAM Rev.*, vol. 31, no. 2, pp. 221–239, Jun. 1989.

[23] G. H. Golub and C. F. Van Loan, *Matrix Computations*, vol. 3. Baltimore, MD, USA: Johns Hopkins Univ. Press, 2012.

[24] K. Abed-Meraim, J.-F. Cardoso, A. Y. Gorokhov, P. Loubaton, and E. Moulines, ''On subspace methods for blind identification of single-input multiple-output FIR systems,'' *IEEE Trans. Signal Process.*, vol. 45, no. 1, pp. 42–55, Jan. 1997.

[25] P. Whittle, ''The analysis of multiple stationary time series,'' *J. Roy. Stat. Soc., B (Methodol.)*, vol. 15, no. 1, pp. 125–139, Jan. 1953.

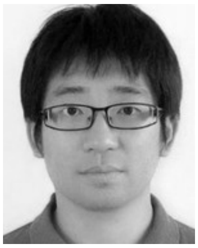
MOHAMED NAIT-MEZIANE was born in Algiers, Algeria, in 1987. He received the B.Eng. degree in electronic engineering from the École Nationale Polytechnique (ENP), Algiers, in 2010, the Magister degree in signal processing from the École Militaire Polytechnique (EMP), Algiers, in 2012, the M.S. degree in signal processing from the Institut Polytechnique de Grenoble (Grenoble INP), Grenoble, France, in 2013, and the Ph.D. degree in signal processing from the Université d'Orléans, Orléans, France, in 2016.

From 2016 to 2020, he held different postdoctoral positions with the Université d'Orléans, where he worked on problems related to blind system/channel identification, blind source separation, and sparse representations of signals. His research interests include statistical signal processing, modeling, estimation, and classification of signals, blind system identification, blind source separation, statistical performance analysis, and sparse representations.

KARIM ABED-MERAIM (Fellow, IEEE) was born in 1967. He received the State Engineering Degree from the École Polytechnique, Palaiseau, France, in 1990, the State Engineering Degree from the École Nationale Supérieure des Télécommunications (ENST), Paris, France, in 1992, the M.Sc. degree from Paris XI University, Orsay, France, in 1992, and the Ph.D. degree in the field of signal processing and communications from ENST, in 1995.

From 1995 to 1998, he was a Research Staff with the Electrical Engineering Department, The University of Melbourne, where he worked on several research project related to ''Blind System Identification for Wireless Communications,'' "Blind Source Separation," and "Array Processing for Communications." From 1998 to 2012, he has been an Assistant Professor and then as an Associate Professor with the Signal and Image Processing Department, Télécom ParisTech. His research interests include signal processing for communications, adaptive filtering and tracking, array processing, and statistical performance analysis. In September 2012, he joined the University d'Orléans (PRISME Laboratory), as a Full Professor. He is the author of over 450 scientific publications, including book chapters, international journal and conference papers, and patents.

Dr. Abed-Meraim is a member of the IEEE SAM-TC and a past Associate Editor of the IEEE Transactions on Signal Processing.



ZHIPENG ZHAO received the M.Eng. and Ph.D. degrees in digital communications from Télécom ParisTech, Paris, France, in 2005 and 2009, respectively.

From 2009 to 2014, he worked as an Algorithm and Architecture Engineer at Comsis, Boussy-Saint-Antoine, France, a telecommunication-based company. Since 2014, he has been a Senior Research Engineer/Project Manager with Huawei Technologies, Paris. His research interests include digital communications, 3G mobile communication, and radio over fiber for cellular networks.

NGUYEN LINH TRUNG (Senior Member, IEEE) received the B.Eng. and Ph.D. degrees in electrical engineering from the Queensland University of Technology, Brisbane, Australia, in 1998 and 2005, respectively.

He was a Postdoctoral Research Fellow with the French National Space Agency (CNES), and held visiting positions with Télécom ParisTech, Vanderbilt University, CentraleSupélec, Université Paris-Sud, Université Paris 13, and the University of Illinois at Urbana-Champaign. Since 2006, he has been on the Faculty of the VNU University of Engineering and Technology (VNU-UET), a member of Vietnam National University, Hanoi (VNU), where he is currently an Associate Professor in electronic engineering with the Faculty of Electronics and Telecommunications, and the Director of the Advanced Institute of Engineering and Technology of VNU-UET. Since 2016, he has also been an Adjunct Professor with the School of Electrical and Data Engineering, University of Technology Sydney. His research interest includes signal processing methods, including time-frequency signal analysis, blind source separation, compressive sampling, tensor-based signal analysis, graph signal processing, and apply them to wireless communication and networking and biomedical engineering, with a current focus on large-scale processing.