Versatility of constrained CRB for system analysis and design

Tarek Menni, Jérôme Galy, Eric Chaumette, Pascal Larzabal

Versatility of constrained CRB for system analysis and design*

Tarek Menni, Jerôme Galy, Eric Chaumette, Pascal Larzabal

Abstract

Provided that one keeps in mind the Cramér-Rao bound (CRB) limitations, that is, to become an overly optimistic lower bound when the observation conditions degrades, the CRB is a lower bound of great interest for analysis and design of a system of measurement in the asymptotic region. As a contribution, we introduce an original framework taking into account most (and possibly all) of factors impacting the asymptotic estimation performance of the parameters of interest via equality constraints, leading to direct algebraic computations of constrained CRB. Anyhow, for complex systems, derivation of analytical expression of CRB is either impossible or inefficient. For application to active systems of measurement such as radar, we provide the general form of the Fisher information matrix (FIM) for multiple conditional models which generally precludes the derivation of an analytical expression of the CRB for scenarios including interference and sensors modeling errors. We show that the proposed framework can also be used efficiently to generate new closed-form expressions of CRB, although this is not its main aim. Last, we examplify the proposed framework in the context of a tracking radar facing different target types (in terms of parameters variability).

Key words: Deterministic parameter estimation, Cramér–Rao bound, Conditional model, Active radar, Waveform diversity

Accepted for publication in IEEE Trans. on AES, September 2013.

* Draft' title was « On the versatility of constrained Crame´r-Rao bound for estimation performance analysis and design of a system of measurement »

Tarek Menni is with the Research and Development Center, Algiers, Algeria. E-mail: menni.tarek@gmail.com

J. Galy is with LIRMM (Laboratoire d'Informatique, de Robotique et de Micro-Electronique de Montpellier) 161 rue ada 34392 Montpellier Cedex 5, France. E-mail: galy@lirmm.fr, Phone: +33 4 67 41 86 43.

Eric Chaumette is with the ONERA, the French Aerospace Lab, DEMR/TSI, Chemin de la Hunière 91120 Palaiseau, France. E-mail: eric.chaumette@onera.fr, Phone: +33 1 80 38 62 96

P. Larzabal is with SATIE/CNRS, Ecole Normale Supérieure de Cachan, 61 avenue du Président Wilson - 94235 Cachan Cedex, France. E-mail: pascal.larzabal@satie.ens-cachan.fr, Phone: +33 1 47 40 27 09.

This work has been partly funded by the European Network of excellence NEWCOM#

I. INTRODUCTION

Minimal performance bounds allow for calculation of the best performance that can be achieved in the Mean Square Error (MSE) sense, when estimating parameters of a signal corrupted by noise. There are two main categories of lower bounds [1]. Those that evaluate the "locally best" behaviour of the estimator and those that consider the "globally best" performance. In the first case, the parameters being estimated are considered to be deterministic, whereas the second category considers the parameters as random variables with an *a priori* probability. This paper is concerned with the first category of bounds concerning deterministic parameters (the case in which signal and/or noise models involve some random parameters [2][3] is not taken into account). Historically the first MSE lower bound for deterministic parameters to be derived was the Cramér-Rao Bound (CRB), which was introduced to investigate fundamental limits of a parameter estimation problem or to assess the relative performance of a specific estimator (efficiency) [1]. It has since become the most popular lower bound due to its simplicity of calculation and the fact that in many cases it can be achieved asymptotically (high signal-to-noise ratio (SNR) and/or large number of snapshots) by Maximum Likelihood Estimators (MLE) [1]. However, numerous works (detailed in [4][5]) have shown that in non-linear estimation problems three distinct regions of operation can be observed. In the asymptotic region, the MSE is small and, in many cases, close to the CRB. In the *a priori* performance region where the number of independent snapshots and/or the SNR are very low, the observations provide little information and the MSE is close to that obtained from the prior knowledge about the problem. Between these two extremes, there is the transition region where the MSE of MLEs usually deteriorates rapidly with respect to CRB, and exhibits a threshold behaviour corresponding to a "performance breakdown" [1] revealed by Large-Error bounds [4][5] that can be used to predict the threshold value. Unfortunately, the computational coast of Large-Error bounds is prohibitive in most applications when the number of unknown parameters increases. Therefore, provided that one keeps in mind the CRB limitations, that is, to become an overly optimistic lower bound when the observation conditions degrades (low SNR and/or low number of snapshots), the CRB is still a lower bound of great interest for system analysis and design in the asymptotic region.

As mentioned in the seminal paper [6], the standard form of the CRB is derived under the implicit assumption that the parameter space is an open subset of $\mathbb{R}^n$ (if we consider unknown deterministic real parameters). However, in many applications (see hereinafter), the vector of unknown parameters is constrained to lie in a proper non-open subset of the original parameter space. In [6], authors have shown that for constraint sets defined by a general smooth (differentiable) functional inequality constraint, the constrained CRB is equivalent to the unconstrained CRB at all regular points of the constraint set, e.g., at interior points. However at nonregular point, such as points governed by equality

constraints, the constrained Fisher information matrix (FIM) consists of the orthogonal projection of the unconstrained FIM onto the tangent hyperplanes of the constraint set.

Since then, numerous works have been devoted either to extend the theoretical results introduced in [6] or to apply these results or its extensions. Thus, [7] presents a simple derivation of the constrained CRB with equality constraints and a new necessary condition for an estimator to satisfy the constrained CRB. Then [8] (and later [9]) provide useful extensions of the constrained CRB with equality constraints where the FIM for the unconstrained problem is not of full rank and the estimators may be biased [9], since for several signal processing problems the unconstrained problem is unidentifiable. Regularization of an unidentifiable unconstrained problem via addition of equality constraints is investigated in [10] for convolutive multi-input multi-output (MIMO) systems and in [11] for the blind source separation (BSS) problem with constant modulus constraints on the sources. [12] extends the formulation of the constrained CRB to complex equality constraints including complex parameters vector for application in the context of a semiblind channel estimation problem. [13] unifies the asymptotic constrained maximum likelihood (CML) theory with the constrained CRB theory by showing that the CML estimate (CMLE) is asymptotically efficient with respect to the constrained CRB. Convergence properties and examples verify the usefulness of the proposed approach. [14] characterizes the best achievable MSE in estimating a sparse deterministic parameter from measurements corrupted by Gaussian noise. To this end, an appropriate definition of bias in the sparse setting is developed, and the associated constrained CRB is derived. [15] investigates the behavior of the MSE of low-rank and sparse matrix decomposition, in particular the special case of the robust principal component analysis (RPCA), and its generalization matrix completion and correction (MCC). [15] derives a constrained CRB for any locally unbiased estimator of the low rank matrix and of the sparse matrix, and analyzes the typical behavior of the constrained CRB for MCC where a subset of entries of the underlying matrix are randomly observed, some of which are grossly corrupted. Lately, [16] has introduced a simple derivation of the CRB, whatever the nature (real or complex) of the unknown parameters, that avoids sophisticated matrix manipulations generally used with complex parameters [12][17]. With this derivation based on the minimization of a norm under a set of linear constraints, the study of FIM singularity, constrained CRB (also in [18] for real parameters), regularity conditions become straightforward corollaries of the derivation, condensing to a few lines previous works [6][7][8][9][12][19][20]. [16] also provides technical results useful for system analysis and design in the asymptotic region: the general reparameterization inequality and the equivalence between parameterization change and equality constraints (previously introduced in [11] and [21] for real parameters). Indeed, a way to improve the estimation of a subset of unknown parameters (parameters of interest for example) can be to introduce, by design choices, either a parameterization change or equality constraints among the other parameters (nuisance parameters for

example) [22]. Additionally, for system design it is also worth knowing if the parameters of interest are identifiable, that is, if they can be estimated (with or without bias) whatever their values. And if they are not identifiable, at least on which subset of the parameter space. In [16] it is shown that the FIM is always singular on a subset of the parameter space deriving from a set of equality constraints (m-dimensional manifold in $\mathbb{R}^n$ or $\mathbb{C}^n$ [23]). Therefore a possible regularization of a singular unconstrained FIM can theoretically be obtained by adding or exploiting some equality constraints on the parameter to be estimated [9][10][11][16][24].

Last, the computation of CRB by direct inversion techniques or other forms of matrix decompositions are known for their high complexity in space and time when the number of parameters is large. Moreover, one may be just interested in a portion of the covariance matrix. Pioneering work [25] proved the tremendous savings in memory and computation by presenting several recursive algorithms computing only submatrices of the CRB. [26] shows that the algorithms proposed in [25] are special instances of a more general framework related to solving a quadratic matrix program which provides methods for fast computation of the CRB, including the cases when the parameters are constrained and when the FIM is singular.

The survey of the open literature dedicated to the constrained CRB paradigm as above, shows that most of previous works in the field has been dedicated to study the CRB modified by constraints either required by the model or required to solve identifiability issues. Apart from this main stream, a few works [12][22] have proposed to investigate the use of parameters constraints from a different perspective: the value of side information on estimation performance, in the sense that a side information allows to treat some formerly unknown parameters as known (for instance, the use of a known training sequence in communications).

Our concern is performance analysis and design of a system of measurement, that is to search for or to check the requirements a given system of measurement must be compliant with to ensure some operational or contractual estimation (measure) performance on some parameters of interest. In this context, the first novel contribution of the present paper is tutorial in nature: all the previous contributions can be gathered to introduce an original framework in order to assess and to analyze the asymptotic estimation performance achievable by any system of measurement modelled as a set of L parametric observation models (whatever they are identical or not and/or independent or not), provided an analytical expression of the parameterized density probability function (p.d.f.) associated with the L observations is available (as it is the case for Gaussian observations, whatever the model is conditional or unconditional, circular or non circular). The main advantage of this framework is to take into account most (and possibly all) of factors impacting the estimation performance of the parameters of interest (see section V for examples) via equality constraints (or reparameterizations) leading to direct numerical algebraic computations of constrained CRB from the expression of the

unconstrained FIM. This framework offers a rational heuristic to assess and to analyze the design of a system of measurement, where derivation of dedicated analytical or closed-form expressions of CRB is superfluous. Anyhow, for complex systems of measurement, derivation of analytical expression of CRB is either impossible or inefficient (see Sections IV-B for examples). Indeed, generally the unknown parameters vector contains the parameters of interest to us as well as other unwanted parameters. Then, the usual technique to obtain the portion of the CRB related to the parameters of interests is to partition the FIM and use the formula for the inverse of block matrices to find the analytical expression of the corresponding block. First, this step is normally a tour de force of linear algebra which is seldom successful for complex observation models. Second, even when successful, the complexity of the analytical expression obtained generally prevents from any insight without resorting to oversimplified particular cases (one or two signals, one or two unknown parameters per signal, ...) or to numerical computations.

The second novel contribution is an application to the case of L conditional models (whatever they are identical or not, independent or not) often used to model active systems of measurement such as radar [27][28]. Additionally, we exhibit some particular cases where the proposed framework can be used efficiently to generate new closed-form expressions of CRB (useful to study the impact of waveform diversity on asymptotic performance estimation), although this is not the main aim of this framework. Last, we exemplify the proposed framework in the context of a tracking radar facing different target types (in terms of parameters variability).

An outline of the paper is as follows. Section II introduces the algebraic notations used in the paper. Section III recalls the useful background on constrained CRB for real and complex parameters. Section IV is an introduction of the proposed framework for a system of measurement consisting of multiple band-limited conditional models. This section first establishes the general expression of the FIM for multiple conditional models, then shows how the proposed framework can be used efficiently to generate new closed-forms expressions of CRB in some particular configurations (absence of sensors modeling errors, independent models); expressions of questionable interest in the light of their complexity. Then section (V) further elaborates on the proposed framework based on constrained CRB by enumerating different varieties of constraints to highlight their potential use in estimation performance analysis and design of a system of measurement. Last, Section VI exemplifies the proposed framework in the context of a tracking radar facing different target types in terms of parameters variability.

II. NOTATIONS

The notational convention adopted is as follows: italic indicates a scalar quantity, as in a; lower case boldface indicates a column vector quantity, as in $\mathbf{a}$; upper case boldface indicates a matrix quantity, as

in $\mathbf{A}$. The *n*-th row and *m*-th column element of the matrix $\mathbf{A}$ will be denoted by $a_{n,m}$ or $(\mathbf{A})_{n,m}$. The *n*-th coordinate of the column vector $\mathbf{a}$ will be denoted by a_n or $(\mathbf{a})_n$. $\mathrm{Re}\{A\}$ is the real part of A and $\mathrm{Im}\{A\}$ is the imaginary part of A. The matrix/vector transpose is indicated by a superscript ^{T} as in $\mathbf{A}^T$. The matrix/vector conjugate is indicated by a superscript * as in $\mathbf{A}^*$. The matrix/vector transpose conjugate is indicated by a superscript as in $\mathbf{A}^H$. $|\mathbf{A}|$ is the determinant of the square matrix $\mathbf{A}$. $[\mathbf{A},\mathbf{B}]$ denotes the matrix resulting from the horizontal concatenation of matrices $\mathbf{A}$ and $\mathbf{B}$. $\left(\mathbf{a}^T,\mathbf{b}^T\right)$ denotes the row vector resulting from the horizontal concatenation of row vectors $\mathbf{a}^T$ and $\mathbf{b}^T$. $\mathbf{I}_M$ is the identity matrix of order M. $\mathbf{vec}(\mathbf{A})$ is a column vector obtained from matrix $\mathbf{A}$ by stacking its column vectors one below another. $S = span\{\mathbf{A}\}$ where $\mathbf{A}$ is a matrix denotes the linear span of the set of its column vectors. $S^\perp$ denotes the orthogonal complement of the subspace S. For two matrices $\mathbf{A}$ and $\mathbf{B}$, $\mathbf{A} \geq \mathbf{B}$ means that $\mathbf{A}-\mathbf{B}$ is positive semi-definite. $E[\cdot]$ denotes the expectation operator and $\|.\|$ denotes a norm. $o(.)$ and $O(.)$ denotes respectively the small oh and big Oh notation. If $\boldsymbol{\theta} = (\theta_1,\theta_2,\ldots,\theta_P)^T$, then: $\frac{\partial}{\partial\boldsymbol{\theta}} = \left(\frac{\partial}{\partial\theta_1},\frac{\partial}{\partial\theta_2},\ldots,\frac{\partial}{\partial\theta_P}\right)^T$, $\frac{\partial}{\partial\boldsymbol{\theta}^T} = \left(\frac{\partial}{\partial\theta_1},\frac{\partial}{\partial\theta_2},\ldots,\frac{\partial}{\partial\theta_P}\right)$. $\mathcal{M}_{\mathbb{R}}(N,P)$ denotes the vector space of real matrices with N rows and P columns. $\mathcal{M}_{\mathbb{C}}(N,P)$ denotes the vector space of complex matrices with N rows and P columns. $\odot$ denotes the Hadamard product. $\otimes$ denotes the Kronecker product. $1(\mathbf{x})$ denotes the constant real-valued function with value equal to 1. $\underline{\mathbf{x}}$ denotes the following notation:

$$\underline{\mathbf{x}}: \begin{cases} \underline{\mathbf{x}} = \mathbf{x} & \text{if } \mathbf{x} \in \mathbb{R}^Q \\ \underline{\mathbf{x}} = \left(\mathbf{x}^T,\mathbf{x}^H\right)^T & \text{if } \mathbf{x} \in \mathbb{C}^Q \text{ and } \mathbf{x} \notin \mathbb{R}^Q \\ \underline{\mathbf{x}} = \left(\mathbf{x}_c^T,\mathbf{x}_c^H,\mathbf{x}_r^T\right)^T & \text{if } \mathbf{x} = \left(\mathbf{x}_c^T,\mathbf{x}_r^T\right)^T, \quad \mathbf{x}_c \in \mathbb{C}^Q \text{ and } \mathbf{x}_c \notin \mathbb{R}^Q, \quad \mathbf{x}_r \in \mathbb{R}^{Q'} \end{cases} \tag{1}$$

Additionally, regarding the definition of Hermitian product, we adopt the convention used in [16] coming from books of mathematics including [29][30], where a sesquilinear form is a function in two variables on a complex vector space $\mathbb{U}$ which is linear in the first variable and semi-linear in the second. This convention allows to define the Gram matrix associated to 2 families of vectors of $\mathbb{U}$, $\{\mathbf{u}\}_{[1,Q]} = \{\mathbf{u}_1,\mathbf{u}_2,\ldots,\mathbf{u}_Q\}$ and $\{\mathbf{c}\}_{[1,P]} = \{\mathbf{c}_1,\mathbf{c}_2,\ldots,\mathbf{c}_P\}$ as [30]:

$$\mathbf{G}\left(\{\mathbf{u}\}_{[1,Q]},\{\mathbf{c}\}_{[1,P]}\right) \in \mathcal{M}_{\mathbb{C}}(P,Q) \;/\; \left(\mathbf{G}\left(\{\mathbf{u}\}_{[1,Q]},\{\mathbf{c}\}_{[1,P]}\right)\right)_{p,q} = \langle \mathbf{u}_q \mid \mathbf{c}_p \rangle \tag{2}$$

leading to:

$$\left\langle \sum_{q=1}^{Q} x_q\mathbf{u}_q \mid \sum_{p=1}^{P} y_p\mathbf{c}_p \right\rangle = \mathbf{y}^H\mathbf{G}\left(\mathbf{u}_{[1,Q]},\mathbf{c}_{[1,P]}\right)\mathbf{x}, \quad \mathbf{x} = (x_1,\ldots,x_Q)^T, \mathbf{y} = (y_1,\ldots,y_P)^T \tag{3}$$

For notational convenience: $\mathbf{G}\left(\{\mathbf{u}\}_{[1,Q]}\right) = \mathbf{G}\left(\{\mathbf{u}\}_{[1,Q]},\{\mathbf{u}\}_{[1,Q]}\right)$.

Beware that most reference signal processing books including [1, p1343][31][32] adopt the opposite convention for sesquilinear form, that is to be semi-linear in the first variable and linear in the second. As a consequence, the equivalent form in "signal processing notation" of any equality/inequality derived in the present paper is obtained by transposing equality/inequality terms. Thanks to the adopted

convention, even in the case of complex parameters, the Fisher Information Matrix (see hereinafter) appears to be both Gram matrix and correlation matrix derived from the canonical definition of the MSE, i.e., a norm associated to an Hermitian product depending on the p.d.f. of the observation.

III. Background on constrained CRB

A. (Unconstrained) CRB

Throughout the present paper, unless otherwise stated, $\mathbf{x}$ denotes the random observation vector of dimension N, Ω denotes the observations space and $L^2(\Omega)$ denotes the complex Hilbert space of square integrable functions over Ω. The probability density function (p.d.f.) of $\mathbf{x}$ is denoted $p(\mathbf{x};\boldsymbol{\theta})$ and depends on a vector of P real parameters $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_P) \in \Theta$, where Θ denotes the parameter space. $p(\mathbf{x};\boldsymbol{\theta})$ is "regular" in the following sense: $\forall \boldsymbol{\theta} \in \Theta, \forall \mathbf{x} \in \Omega, p(\mathbf{x};\boldsymbol{\theta}) > 0$ for almost every $\mathbf{x}$ in the observation space Ω; $\forall \boldsymbol{\theta} \in \Theta, \forall \mathbf{x} \in \Omega, p(\mathbf{x};\boldsymbol{\theta})$ is continuous and differentiable with respect to $\boldsymbol{\theta}$; $p(\mathbf{x};\boldsymbol{\theta})$ does not incorporate any probability mass function. Additionally, we assume that the observation vector $\mathbf{x}$ corresponds to a parametric observation model involving $P_r \geq 0$ real unknown parameters (delays, directions of arrival, ...) and $P_c \geq 0$ complex unknown parameters (spatial transfer functions components, complex amplitudes, ...) where $2P_c + P_r = P$, leading to a p.d.f. of the form:

$$p(\mathbf{x};\boldsymbol{\theta}), \quad \boldsymbol{\theta} = (\theta_1, \ldots, \theta_P)^T = \left(\mathrm{Re}\left\{\boldsymbol{\theta}_c^T\right\}, \mathrm{Im}\left\{\boldsymbol{\theta}_c^T\right\}, \boldsymbol{\theta}_r^T\right)^T \in \mathbb{R}^P, \quad \boldsymbol{\theta}_c \in \mathbb{C}^{P_c},\ \boldsymbol{\theta}_r \in \mathbb{R}^{P_r},\ 2P_c + P_r = P$$

Then the p.d.f. of $\mathbf{x}$ can be parameterized in a dual form [16]:

$$p(\mathbf{x};\boldsymbol{\theta}), \quad \boldsymbol{\theta} = \left(\mathrm{Re}\left\{\boldsymbol{\theta}_c^T\right\}, \mathrm{Im}\left\{\boldsymbol{\theta}_c^T\right\}, \boldsymbol{\theta}_r^T\right)^T \in \mathbb{R}^P \tag{4}$$

$$p(\mathbf{x};\underline{\boldsymbol{\theta}}), \quad \underline{\boldsymbol{\theta}} = \left(\boldsymbol{\theta}_c^T, (\boldsymbol{\theta}_c^*)^T, \boldsymbol{\theta}_r^T\right)^T \in \mathbb{C}^{2P_c} \times \mathbb{R}^{P_r}, \quad 2P_c + P_r = P, \quad P_c \geq 0,\ P_r \geq 0 \tag{5}$$

In the following we will only consider the form (5) since it includes (4) when $P_c = 0$.

Let $\underline{\boldsymbol{\theta}}^0$ be a selected value of the parameter $\underline{\boldsymbol{\theta}}$, and $\widehat{\mathbf{g}\left(\underline{\boldsymbol{\theta}}^0\right)}(\mathbf{x})$ an estimator of $\mathbf{g}\left(\underline{\boldsymbol{\theta}}^0\right)$ where $\mathbf{g}(\underline{\boldsymbol{\theta}}) = \left(g_1(\underline{\boldsymbol{\theta}}), \ldots, g_{Q_c}(\underline{\boldsymbol{\theta}}), g_{Q_c+1}(\underline{\boldsymbol{\theta}}), \ldots, g_{Q_c+Q_r}(\underline{\boldsymbol{\theta}})\right)^T$ is a vector of $Q_c + Q_r$ functions of $\underline{\boldsymbol{\theta}}$, the first $Q_c \geq 0$ ones being complex-valued functions, the last $Q_r \geq 0$ being real-valued functions. For any selected value $\underline{\boldsymbol{\theta}}^0$, $\widehat{\mathbf{g}\left(\underline{\boldsymbol{\theta}}^0\right)}(\mathbf{x})$ stands for a mapping of the observation space Ω into an estimate of $\mathbf{g}\left(\underline{\boldsymbol{\theta}}^0\right)$, designed to have a low MSE at $\underline{\boldsymbol{\theta}}^0$ (possibly the lowest) and some relevant properties for other values of $\underline{\boldsymbol{\theta}}$, as unbiasedness for instance. Let us recall that $\widehat{\mathbf{g}\left(\underline{\boldsymbol{\theta}}^0\right)}(\mathbf{x})$ is an unbiased estimator of $\mathbf{g}\left(\underline{\boldsymbol{\theta}}^0\right)$ at the selected value $\underline{\boldsymbol{\theta}}^0$ if:

$$E_{\underline{\boldsymbol{\theta}}^0}\left[\widehat{\mathbf{g}\left(\underline{\boldsymbol{\theta}}^0\right)}(\mathbf{x})\right] = \mathbf{g}\left(\underline{\boldsymbol{\theta}}^0\right) = \int_{\Omega} \widehat{\mathbf{g}\left(\underline{\boldsymbol{\theta}}^0\right)}(\mathbf{x})\, p\left(\mathbf{x};\underline{\boldsymbol{\theta}}^0\right) d\mathbf{x}, \tag{6}$$

where $E_{\underline{\boldsymbol{\theta}}}[\mathbf{g}(\mathbf{x})]$ is the statistical expectation of the vector of functions $\mathbf{g}(\,)$ with respect to $\mathbf{x}$ parameterized by $\underline{\boldsymbol{\theta}}$. Actually, if the exhaustive characterization - in the sense of statistical performance

- of an estimator $\widehat{\mathbf{g}\left(\boldsymbol{\theta}^0\right)}\left(\mathbf{x}\right)$ containing complex-valued components, is supposed to include the characterization of all its individual components, that is real and imaginary parts, then the characterization of $\widehat{\underline{\mathbf{g}\left(\underline{\boldsymbol{\theta}}^0\right)}}\left(\mathbf{x}\right)$, where $\dim\left\{\widehat{\underline{\mathbf{g}\left(\underline{\boldsymbol{\theta}}^0\right)}}\left(\mathbf{x}\right)\right\} = 2Q_c + Q_r = Q$, is necessary, which can be achieved by the knowledge of $\mathbf{G}_{\underline{\boldsymbol{\theta}}^0}\left(\left\{\widehat{\underline{\mathbf{g}\left(\underline{\boldsymbol{\theta}}^0\right)}}\left(\mathbf{x}\right) - \underline{\mathbf{g}}\left(\underline{\boldsymbol{\theta}}^0\right)\right\}\right)$ where:

$$\left\{\mathbf{h}\left(\mathbf{x}\right)\right\} = \left\{\left(h_1\left(\mathbf{x}\right),\ldots,h_Q\left(\mathbf{x}\right)\right)^T\right\} = \left\{h_1\left(\mathbf{x}\right),\ldots,h_Q\left(\mathbf{x}\right)\right\} \tag{7}$$

denotes a family of vectors whose elements are the vector components, and:

$$\left(\mathbf{G}_{\underline{\boldsymbol{\theta}}^0}\left(\left\{\mathbf{g}\left(\mathbf{x}\right)\right\},\left\{\mathbf{h}\left(\mathbf{x}\right)\right\}\right)\right)_{p,q} = \left\langle g_q\left(\mathbf{x}\right) \mid h_p\left(\mathbf{x}\right)\right\rangle_{\underline{\boldsymbol{\theta}}^0} \tag{8}$$

is a Gram matrix associated to the Hermitian product $\left\langle \mid \right\rangle_{\underline{\boldsymbol{\theta}}^0}$ depending on $p\left(\mathbf{x};\underline{\boldsymbol{\theta}}^0\right)$:

$$\left\langle g\left(\mathbf{x}\right) \mid h\left(\mathbf{x}\right)\right\rangle_{\underline{\boldsymbol{\theta}}^0} = E_{\underline{\boldsymbol{\theta}}^0}\left[g\left(\mathbf{x}\right)h^*\left(\mathbf{x}\right)\right] = \int\limits_{\Omega} g\left(\mathbf{x}\right)h^*\left(\mathbf{x}\right)p\left(\mathbf{x};\underline{\boldsymbol{\theta}}^0\right)d\mathbf{x}. \tag{9}$$

Then, if $\underline{\mathbf{g}}\left(\underline{\boldsymbol{\theta}}\right)$ is differentiable at $\underline{\boldsymbol{\theta}}^0$ (for both real and complex components [16, III.A]) the (unconstrained) CRB is given by:

$$\mathbf{G}_{\underline{\boldsymbol{\theta}}^0}\left(\left\{\widehat{\underline{\mathbf{g}\left(\underline{\boldsymbol{\theta}}^0\right)}}\left(\mathbf{x}\right) - \underline{\mathbf{g}}\left(\underline{\boldsymbol{\theta}}^0\right)\right\}\right) \geq \mathbf{CRB}_{\underline{\mathbf{g}}|\underline{\boldsymbol{\theta}}}\left(\underline{\boldsymbol{\theta}}^0\right) = \frac{\partial\underline{\mathbf{g}}^*\left(\underline{\boldsymbol{\theta}}^0\right)}{\partial\underline{\boldsymbol{\theta}}^H}\mathbf{F}_{\underline{\boldsymbol{\theta}}^0}^{-1}\frac{\partial\underline{\mathbf{g}}^T\left(\underline{\boldsymbol{\theta}}^0\right)}{\partial\underline{\boldsymbol{\theta}}} \tag{10}$$

$$\mathbf{F}_{\underline{\boldsymbol{\theta}}^0} = \mathbf{G}_{\underline{\boldsymbol{\theta}}^0}\left(\left\{\frac{\partial\ln p\left(\mathbf{x};\underline{\boldsymbol{\theta}}\right)}{\partial\underline{\boldsymbol{\theta}}}^*\right\}\right) = E_{\underline{\boldsymbol{\theta}}^0}\left[\frac{\partial\ln p\left(\mathbf{x};\underline{\boldsymbol{\theta}}\right)}{\partial\underline{\boldsymbol{\theta}}}\frac{\partial\ln p\left(\mathbf{x};\underline{\boldsymbol{\theta}}\right)}{\partial\underline{\boldsymbol{\theta}}}^H\right] \tag{11}$$

$$\left(\widehat{\underline{\mathbf{g}\left(\underline{\boldsymbol{\theta}}^0\right)}}\left(\mathbf{x}\right) - \underline{\mathbf{g}}\left(\underline{\boldsymbol{\theta}}^0\right)\right)_{\text{eff}}^T = \frac{\partial\ln p\left(\mathbf{x};\underline{\boldsymbol{\theta}}\right)}{\partial\underline{\boldsymbol{\theta}}^H}\mathbf{F}_{\underline{\boldsymbol{\theta}}^0}^{-1}\frac{\partial\underline{\mathbf{g}}^T\left(\underline{\boldsymbol{\theta}}^0\right)}{\partial\underline{\boldsymbol{\theta}}} \tag{12}$$

provided that the FIM $\mathbf{F}_{\underline{\boldsymbol{\theta}}^0}$ is invertible, that is provided that $\left\{\frac{\partial p(\mathbf{x};\underline{\boldsymbol{\theta}})}{\partial\underline{\boldsymbol{\theta}}}\right\}$ is an independent family of vectors.

B. Constrained CRB

The addition of K non redundant equality constraints on the unknown deterministic parameters :

$$\underline{\mathbf{f}}\left(\underline{\boldsymbol{\theta}}\right) = \mathbf{0} \in \mathbb{C}^{2K_c}\times\mathbb{R}^{K_r},\ 2K_c + K_r = K,\ 1 \leq K < P \tag{13}$$

where the matrix $\frac{\partial\underline{\mathbf{f}}(\underline{\boldsymbol{\theta}})}{\partial\underline{\boldsymbol{\theta}}^T} \in \mathcal{M}_{\mathbb{C}}\left(K,P\right)$ has full row rank (K), leads to the constrained FIM and constrained CRB [6][7][8][9][12][16][18]:

$$\mathbf{G}_{\underline{\boldsymbol{\theta}}^0}\left(\left\{\widehat{\underline{\mathbf{g}\left(\underline{\boldsymbol{\theta}}^0\right)}}\left(\mathbf{x}\right) - \underline{\mathbf{g}}\left(\underline{\boldsymbol{\theta}}^0\right)\right\}\right) \geq \mathbf{CRB}^c_{\underline{\mathbf{g}}|\underline{\boldsymbol{\theta}}}\left(\underline{\boldsymbol{\theta}}^0\right) = \frac{\partial\underline{\mathbf{g}}^*\left(\underline{\boldsymbol{\theta}}^0\right)}{\partial\underline{\boldsymbol{\theta}}^H}\mathbf{U}^*_{\underline{\boldsymbol{\theta}}^0}\left(\mathbf{F}^c_{\underline{\boldsymbol{\theta}}^0}\right)^{-1}\mathbf{U}^T_{\underline{\boldsymbol{\theta}}^0}\frac{\partial\underline{\mathbf{g}}^T\left(\underline{\boldsymbol{\theta}}^0\right)}{\partial\underline{\boldsymbol{\theta}}} \tag{14}$$

$$\mathbf{F}^c_{\underline{\boldsymbol{\theta}}^0} = \mathbf{G}_{\underline{\boldsymbol{\theta}}^0}\left(\left\{\mathbf{U}^H_{\underline{\boldsymbol{\theta}}^0}\frac{\partial\ln p\left(\mathbf{x};\underline{\boldsymbol{\theta}}\right)}{\partial\underline{\boldsymbol{\theta}}}^*\right\}\right) = \mathbf{U}^T_{\underline{\boldsymbol{\theta}}^0}\mathbf{F}_{\underline{\boldsymbol{\theta}}^0}\mathbf{U}^*_{\underline{\boldsymbol{\theta}}^0} \tag{15}$$

$$\left(\widehat{\underline{\mathbf{g}\left(\underline{\boldsymbol{\theta}}^0\right)}}\left(\mathbf{x}\right) - \underline{\mathbf{g}}\left(\underline{\boldsymbol{\theta}}^0\right)\right)_{\text{eff}}^T = \frac{\partial\ln p\left(\mathbf{x};\underline{\boldsymbol{\theta}}\right)}{\partial\underline{\boldsymbol{\theta}}^H}\mathbf{U}^*_{\underline{\boldsymbol{\theta}}^0}\left(\mathbf{F}^c_{\underline{\boldsymbol{\theta}}^0}\right)^{-1}\mathbf{U}^T_{\underline{\boldsymbol{\theta}}^0}\frac{\partial\underline{\mathbf{g}}^T\left(\underline{\boldsymbol{\theta}}^0\right)}{\partial\underline{\boldsymbol{\theta}}}$$

where $\mathbf{U}_{\underline{\boldsymbol{\theta}}^0}$ is a basis of $\ker\left\{\frac{\partial\underline{\mathbf{f}}(\underline{\boldsymbol{\theta}}^0)}{\partial\underline{\boldsymbol{\theta}}^T}\right\}$ and $\mathbf{F}^c_{\underline{\boldsymbol{\theta}}^0}$ denotes the constrained FIM. It is worth noticing that the constrained CRB does not depend on the choice of the basis $\mathbf{U}_{\underline{\boldsymbol{\theta}}^0}$.

Moreover, it has been shown [11][16][21] that, regarding the computation of the CRB, non redundant equality constraints on parameters (13):

$$\underline{\mathbf{f}}\left(\underline{\boldsymbol{\theta}}\right)=\mathbf{0}\in\mathbb{C}^{K},\ \underline{\boldsymbol{\theta}}\in\mathbb{C}^{P},\ 1\leq K<P,$$

restricts $\underline{\boldsymbol{\theta}}$ to a manifold, which amounts to a reparameterization of the unknown parameters $\underline{\boldsymbol{\theta}}=\underline{\boldsymbol{\theta}}\left(\underline{\boldsymbol{\omega}}\right)$, $\dim\left\{\underline{\boldsymbol{\omega}}\right\}=P-K$ [23, Theorem 5-2]. Then, the reparameterization inequality [16]:

$$\mathbf{CRB}_{\underline{\mathbf{g}}|\underline{\boldsymbol{\theta}}}\left(\underline{\boldsymbol{\theta}}\left(\underline{\boldsymbol{\omega}}^{0}\right)\right)\geq\mathbf{CRB}_{\underline{\mathbf{g}}(\underline{\boldsymbol{\theta}})|\underline{\boldsymbol{\omega}}}\left(\underline{\boldsymbol{\omega}}^{0}\right)\tag{16}$$

where:

$$\mathbf{CRB}_{\underline{\mathbf{g}}(\underline{\boldsymbol{\theta}})|\underline{\boldsymbol{\omega}}}\left(\underline{\boldsymbol{\omega}}^{0}\right)=\frac{\partial\underline{\mathbf{g}}^{*}\left(\underline{\boldsymbol{\theta}}\left(\underline{\boldsymbol{\omega}}^{0}\right)\right)}{\partial\underline{\boldsymbol{\theta}}^{H}}\mathbf{CRB}_{\underline{\boldsymbol{\theta}}|\underline{\boldsymbol{\omega}}}\left(\underline{\boldsymbol{\omega}}^{0}\right)\frac{\partial\underline{\mathbf{g}}^{T}\left(\underline{\boldsymbol{\theta}}\left(\underline{\boldsymbol{\omega}}^{0}\right)\right)}{\partial\underline{\boldsymbol{\theta}}},$$

$$\mathbf{CRB}_{\underline{\boldsymbol{\theta}}|\underline{\boldsymbol{\omega}}}\left(\underline{\boldsymbol{\omega}}^{0}\right)=\frac{\partial\underline{\boldsymbol{\theta}}^{*}\left(\underline{\boldsymbol{\omega}}^{0}\right)}{\partial\underline{\boldsymbol{\omega}}^{H}}\left(\frac{\partial\underline{\boldsymbol{\theta}}^{T}\left(\underline{\boldsymbol{\omega}}^{0}\right)}{\partial\underline{\boldsymbol{\omega}}}\mathbf{F}_{\underline{\boldsymbol{\theta}}(\underline{\boldsymbol{\omega}}^{0})}\frac{\partial\underline{\boldsymbol{\theta}}^{*}\left(\underline{\boldsymbol{\omega}}^{0}\right)}{\partial\underline{\boldsymbol{\omega}}^{H}}\right)^{-1}\frac{\partial\underline{\boldsymbol{\theta}}^{T}\left(\underline{\boldsymbol{\omega}}^{0}\right)}{\partial\underline{\boldsymbol{\omega}}},$$

holds provided that $\mathbf{U}_{\underline{\boldsymbol{\theta}}^{0}}\triangleq\frac{\partial\underline{\boldsymbol{\theta}}(\underline{\boldsymbol{\omega}}^{0})}{\partial\underline{\boldsymbol{\omega}}^{T}}$, where $\mathbf{U}_{\underline{\boldsymbol{\theta}}^{0}}\in\mathcal{M}_{\mathbb{C}}\left(P,P-K\right)$ is a basis of $\ker\left\{\frac{\partial\underline{\mathbf{f}}(\underline{\boldsymbol{\theta}}^{0})}{\partial\underline{\boldsymbol{\theta}}^{T}}\right\}$. $\mathbf{U}_{\underline{\boldsymbol{\theta}}^{0}}$ can always be computed - after rearrangement of $\underline{\boldsymbol{\theta}}$ - as [16][21]:

$$\mathbf{U}_{\underline{\boldsymbol{\theta}}^{0}}=\left[\begin{array}{c}\mathbf{I}_{P-K}\\-\left(\frac{\partial\underline{\mathbf{f}}(\underline{\boldsymbol{\theta}}^{0})}{\partial\underline{\boldsymbol{\varepsilon}}^{T}}\right)^{-1}\frac{\partial\underline{\mathbf{f}}(\underline{\boldsymbol{\theta}}^{0})}{\partial\underline{\boldsymbol{\omega}}^{T}}\end{array}\right],\quad\underline{\boldsymbol{\theta}}=\left(\begin{array}{c}\underline{\boldsymbol{\omega}}\\\underline{\boldsymbol{\varepsilon}}\end{array}\right)\tag{17}$$

where $\underline{\boldsymbol{\varepsilon}}$ is a subvector (subset) of K components of $\underline{\boldsymbol{\theta}}$ which K columns of partial derivatives - columns of matrix $\frac{\partial\underline{\mathbf{f}}(\underline{\boldsymbol{\theta}}^{0})}{\partial\underline{\boldsymbol{\theta}}^{T}}$ - are independent.

IV. Estimation performance and design of a system with constrained CRB under the conditional model

Our main interest is the design of system of measurement where the optimization criterion is the system estimation performance for a set of parameters of interest. By way of illustration, we consider the convenient (and quite standard) setting of a system of measurement generating L observations where each observation consists of a signal of interest corrupted by another signal usually called noise signal. In the following, for sake of legibility, the term "signal" will implicitly refer to any signal of interest and the term "noise" will implicitly refer to any noise signal (whatever we consider an "internal" noise source (thermal noise) or an "external" noise source (interference)). Each observation is modelled as a N^{l}-dimensional random vector $\mathbf{x}^{l}$, function of a vector of $P^{l}+Q^{l}\geq1$ deterministic parameters (real or complex) $\underline{\boldsymbol{\theta}}^{l}$:

$$\mathbf{x}^{l}\triangleq\mathbf{x}^{l}\left(\underline{\boldsymbol{\theta}}^{l}\right)=\mathbf{s}^{l}\left(\underline{\boldsymbol{\theta}}_{\mathbf{s}}^{l},\underline{\boldsymbol{\delta}}_{\mathbf{s}}^{l}\right)+\mathbf{n}^{l}\left(\underline{\boldsymbol{\theta}}_{\mathbf{n}}^{l},\underline{\boldsymbol{\delta}}_{\mathbf{n}}^{l}\right),\quad1\leq l\leq L\tag{18}$$

$$\left(\underline{\boldsymbol{\theta}}^{l}\right)^{T}=\left(\left(\underline{\boldsymbol{\theta}}_{\mathbf{s}}^{l}\right)^{T},\left(\underline{\boldsymbol{\delta}}_{\mathbf{s}}^{l}\right)^{T},\left(\underline{\boldsymbol{\theta}}_{\mathbf{n}}^{l}\right)^{T},\left(\underline{\boldsymbol{\delta}}_{\mathbf{n}}^{l}\right)^{T}\right),\quad\left(\underline{\boldsymbol{\theta}}_{\mathbf{s}}^{l}\right)^{T}=\left(\left(\underline{\boldsymbol{\sigma}}_{\mathbf{s}}^{l}\right)^{T},\left(\underline{\boldsymbol{\Xi}}_{\mathbf{s}}^{l}\right)^{T}\right)$$

where :

• $\mathbf{n}^l\left(\underline{\boldsymbol{\theta}}_{\mathbf{n}}^l, \underline{\boldsymbol{\delta}}_{\mathbf{n}}^l\right)$ is a noise vector depending both on a vector $\underline{\boldsymbol{\theta}}_{\mathbf{n}}^l$ of $P_{\mathbf{n}}^l$ parameters associated with sources of noise (thermal noise covariance matrix, interference DOAs, ...) and on a vector $\underline{\boldsymbol{\delta}}_{\mathbf{n}}^l$ of $Q_{\mathbf{n}}^l$ parameters independent from noise (typically parameters of the system physical components contributing to noise observations (thermal noise or interference)),

• $\mathbf{s}^l\left(\underline{\boldsymbol{\theta}}_{\mathbf{s}}^l, \underline{\boldsymbol{\delta}}_{\mathbf{s}}^l\right)$ is signal vector depending both on a vector $\underline{\boldsymbol{\theta}}_{\mathbf{s}}^l$ of $P_{\mathbf{s}}^l$ parameters associated with sources of signal (DOAs, amplitudes, ...) and on a vector $\underline{\boldsymbol{\delta}}_{\mathbf{s}}^l$ of $Q_{\mathbf{s}}^l$ parameters independent from signal (typically parameters of the system physical components contributing to signal observations (waveforms generator, sensors on transmit and on receive, ...)),

• $\underline{\boldsymbol{\Xi}}_{\mathbf{s}}^l$ are the signal parameters of interest and $\underline{\boldsymbol{\sigma}}_{\mathbf{s}}^l$ are the signal nuisance parameters (as they are not of interest but present [1][32][31]),

• $P^l = P_{\mathbf{n}}^l + P_{\mathbf{s}}^l$, $Q^l = Q_{\mathbf{n}}^l + Q_{\mathbf{s}}^l$.

Therefore we look for the requirements on $\left\{\mathbf{s}^l\left(\ ,\underline{\boldsymbol{\delta}}_{\mathbf{s}}^l\right), \mathbf{n}^l\left(\underline{\boldsymbol{\theta}}_{\mathbf{n}}^l, \underline{\boldsymbol{\delta}}_{\mathbf{n}}^l\right)\right\}_{l=1}^L$ a given system of measurement must be compliant with to ensure some operational or contractual estimation performance on (selected values of) the parameters of interest $\left\{\underline{\boldsymbol{\Xi}}_{\mathbf{s}}^l\right\}_{l=1}^L$ for given (selected values of) nuisance parameters $\left\{\underline{\boldsymbol{\sigma}}_{\mathbf{s}}^l\right\}_{l=1}^L$. The metric chosen for estimation performance is the CRB. Therefore in the following, unless otherwise stated, for sake of legibility and conciseness, the wording "estimation performance" alone always stands for the estimation performance provided by the CRB of the parameters of interest $\left\{\underline{\boldsymbol{\Xi}}_{\mathbf{s}}^l\right\}_{l=1}^L$.

A first noteworthy feature of the (Gaussian) conditional model [33] is to be an observation model for which the asymptotic Gaussianity and efficiency of MLEs has been proved [33][44]. A second noteworthy feature is to be a suitable observation model for active systems of measurement such as radar [40][42], which are our primary systems of interest. For a broader perspective, let us consider the general problem of geographical localization (or positioning) which has received ever growing attention by the signal processing community. In the localization/positioning paradigm a distinction can be made between active localization and passive localization. Passive localization refers essentially to military applications where we want to localize a passive transmitter that does not cooperate: the transmitted waveform is unknown from the receiver. On the contrary, active localization usually refers to the case where the transmitter and the receiver cooperate: the transmitted waveform is known from the receiver. In an active localization system such a radar (or sonar or lidar), a known waveform is transmitted and the signals scattered from the targets of interest are used to estimate their parameters. Typically, the received signals are modelled as scaled, delayed, and Doppler-shifted versions of the transmitted signal; see, e.g., [39]. Estimation of the time delay and Doppler shift provides information about the range and radial velocity of the targets. The use of spatial diversity, i.e. antenna arrays, compared with a single sensor, guarantees more accurate range and velocity estimation and allows

estimation of the targets direction. Last, but no least, waveform diversity [43] may be used to improve the estimation of all targets parameters. The same analysis apply for communication systems, except that the wave transmission is designed to be direct and not via scatterers. In an active system of measurement, as the waveform parametric model is known and deterministic (in opposition with a passive system where a probabilistic modelling of the waveform is generally considered), the most accurate statistical prediction for L observations is obtained when the scattered signal amplitudes are modelled as deterministic (since it is well known that the complex Gaussian amplitude modelling provide an average unconditional CRB higher that the corresponding conditional CRB [1][31][32]).

A. General expression of the FIM for L conditional models

Even though there are many FIM formulas on this topic since numerous works have been done in this field (see references in [1][31][32][43]), each FIM formula is specific to some particular systems modelling (narrow band arrays, narrow band transmitted signals, temporally white noise, Doppler effect approximation, single observation model....). Therefore, to the best of our knowledge, what is missing is a neat FIM formula able to take into account all possible diversities, i.e. valid for L observations models independently of underlying approximations. It is the aim of this section to provide such a general FIM formula for the conditional model.

1) FIM for a single band limited conditional model:

In many practical problems of interest (radar, sonar, communication, ...), the complex N-dimensional observation vector $\mathbf{x}$ consists of a bandpass signal with bandwidth B $\left(f \in \left[-\frac{B}{2}, \frac{B}{2}\right]\right)$ observed for a duration T (length of the time interval containing most (as much as desired) of the signal energy), which is the output of an Hilbert filtering leading to an "in-phase" real part associated to a "quadrature" imaginary part [1], *i.e.* a complex circular vector of the form:

$$\mathbf{x}\left(t;\underline{\boldsymbol{\theta}}\right) = \mathbf{s}\left(t;\underline{\boldsymbol{\theta}}_{\mathbf{s}},\underline{\boldsymbol{\delta}}\right) + \mathbf{n}\left(t;\underline{\boldsymbol{\theta}}_{\mathbf{n}},\underline{\boldsymbol{\delta}}\right), \quad \underline{\boldsymbol{\theta}}^T = \left(\underline{\boldsymbol{\theta}}_{\mathbf{s}}^T, \underline{\boldsymbol{\theta}}_{\mathbf{n}}^T, \underline{\boldsymbol{\delta}}^T\right), \quad t_0 \leq t \leq t_0 + T \tag{19}$$

where $\mathbf{s}\left(t;\underline{\boldsymbol{\theta}}_{\mathbf{s}},\underline{\boldsymbol{\delta}}\right)$ is the signal and $\mathbf{n}\left(t;\underline{\boldsymbol{\theta}}_{\mathbf{n}},\underline{\boldsymbol{\delta}}\right)$ is the noise. $\underline{\boldsymbol{\theta}}_{\mathbf{s}}$ denotes all the parameters involved solely in the signal parametric model (signal DOAs, signal waveform parameters, signal amplitudes, ...). $\underline{\boldsymbol{\theta}}_{\mathbf{n}}$ denotes all the parameters involved solely in the noise parametric model (interference DOAs, interference powers, ...). $\underline{\boldsymbol{\delta}}$ denotes all the parameters involved both in the signal and the noise parametric models (sensor locations, sensor modeling errors, ...).

Under the assumption of circular complex Gaussian centred noise $\mathbf{n}\left(t;\underline{\boldsymbol{\theta}}_{\mathbf{n}},\underline{\boldsymbol{\delta}}\right)$ and a deterministic signal $\mathbf{s}\left(t;\underline{\boldsymbol{\theta}}_{\mathbf{s}},\underline{\boldsymbol{\delta}}\right)$, (19) belongs to the set of conditional models [1] which p.d.f. at time t is:

$$\begin{aligned} p\left(\mathbf{x};t,\underline{\boldsymbol{\theta}}\right) &= \frac{e^{-\left(\mathbf{x}\left(t;\underline{\boldsymbol{\theta}}\right)-\mathbf{s}\left(t;\underline{\boldsymbol{\theta}}_{\mathbf{s}},\underline{\boldsymbol{\delta}}\right)\right)^H \mathbf{C}_{\mathbf{n}}^{-1}\left(t;\underline{\boldsymbol{\theta}}_{\mathbf{n}},\underline{\boldsymbol{\delta}}\right)\left(\mathbf{x}\left(t;\underline{\boldsymbol{\theta}}\right)-\mathbf{s}\left(t;\underline{\boldsymbol{\theta}}_{\mathbf{s}},\underline{\boldsymbol{\delta}}\right)\right)}}{\pi^N \left|\mathbf{C}_{\mathbf{n}}\left(t;\underline{\boldsymbol{\theta}}_{\mathbf{n}},\underline{\boldsymbol{\delta}}\right)\right|} \\ \mathbf{C}_{\mathbf{n}}\left(t;\underline{\boldsymbol{\theta}}_{\mathbf{n}},\underline{\boldsymbol{\delta}}\right) &= \mathbf{R}_{\mathbf{n}}\left(t;\underline{\boldsymbol{\theta}}_{\mathbf{n}},\underline{\boldsymbol{\delta}}\right) = E_{\underline{\boldsymbol{\theta}}}\left[\mathbf{n}\left(t;\underline{\boldsymbol{\theta}}_{\mathbf{n}},\underline{\boldsymbol{\delta}}\right)\mathbf{n}\left(t;\underline{\boldsymbol{\theta}}_{\mathbf{n}},\underline{\boldsymbol{\delta}}\right)^H\right] \end{aligned}$$

Additionally, if $\mathbf{n}\left(t;\underline{\boldsymbol{\theta}}_{\mathbf{n}},\underline{\boldsymbol{\delta}}\right)$ is a wide sense stationary (WSS) band limited process with spectral density matrix $\mathbf{R}_{\mathbf{n}}\left(f;\underline{\boldsymbol{\theta}}_{\mathbf{n}},\underline{\boldsymbol{\delta}}\right)$ and autocovariance matrix $\mathbf{R}_{\mathbf{n}}\left(t;\underline{\boldsymbol{\theta}}_{\mathbf{n}},\underline{\boldsymbol{\delta}}\right)$, then:

$$\mathbf{R}_{\mathbf{n}}\left(f;\underline{\boldsymbol{\theta}}_{\mathbf{n}},\underline{\boldsymbol{\delta}}\right)=\int_{-\infty}^{+\infty}\mathbf{R}_{\mathbf{n}}\left(t;\underline{\boldsymbol{\theta}}_{\mathbf{n}},\underline{\boldsymbol{\delta}}\right)e^{-j2\pi ft}dt=\sum_{k=-\infty}^{+\infty}\mathbf{R}_{\mathbf{n}}\left(\frac{k}{B};\underline{\boldsymbol{\theta}}_{\mathbf{n}},\underline{\boldsymbol{\delta}}\right)\frac{e^{-j2\pi\frac{f}{B}k}}{B},$$

and using previously released results such as:

- the FIM for temporally white noise [1][27][43],
- the property of the FIM to be invariant to reversible operations on observations [45],
- a theorem due to Whittle [46][47, th. 9],

it can be shown [48] that the FIM associated to (19) observed during the finite duration T is given by (generalization of [42, (13)]):

$$\mathbf{F}_{\underline{\boldsymbol{\theta}}}=\mathbf{F}_{\underline{\boldsymbol{\theta}},\underline{\boldsymbol{\theta}}^*}=\begin{bmatrix}\mathbf{F}_{\underline{\boldsymbol{\theta}}_{\mathbf{s}},\underline{\boldsymbol{\theta}}_{\mathbf{s}}^*} & \mathbf{F}^H_{\underline{\boldsymbol{\theta}}_{\mathbf{n}},\underline{\boldsymbol{\theta}}_{\mathbf{s}}^*} & \mathbf{F}^H_{\underline{\boldsymbol{\delta}},\underline{\boldsymbol{\theta}}_{\mathbf{s}}^*}\\ \mathbf{F}_{\underline{\boldsymbol{\theta}}_{\mathbf{n}},\underline{\boldsymbol{\theta}}_{\mathbf{s}}^*} & \mathbf{F}_{\underline{\boldsymbol{\theta}}_{\mathbf{n}},\underline{\boldsymbol{\theta}}_{\mathbf{n}}^*} & \mathbf{F}^H_{\underline{\boldsymbol{\delta}},\underline{\boldsymbol{\theta}}_{\mathbf{n}}^*}\\ \mathbf{F}_{\underline{\boldsymbol{\delta}},\underline{\boldsymbol{\theta}}_{\mathbf{s}}^*} & \mathbf{F}_{\underline{\boldsymbol{\delta}},\underline{\boldsymbol{\theta}}_{\mathbf{n}}^*} & \mathbf{F}_{\underline{\boldsymbol{\delta}},\underline{\boldsymbol{\delta}}^*}\end{bmatrix}=\begin{bmatrix}\mathbf{F}_{\underline{\boldsymbol{\theta}}_{\mathbf{s}},\underline{\boldsymbol{\theta}}_{\mathbf{s}}^*} & \mathbf{0} & \mathbf{F}^H_{\underline{\boldsymbol{\delta}},\underline{\boldsymbol{\theta}}_{\mathbf{s}}^*}\\ \mathbf{0} & \mathbf{F}_{\underline{\boldsymbol{\theta}}_{\mathbf{n}},\underline{\boldsymbol{\theta}}_{\mathbf{n}}^*} & \mathbf{F}^H_{\underline{\boldsymbol{\delta}},\underline{\boldsymbol{\theta}}_{\mathbf{n}}^*}\\ \mathbf{F}_{\underline{\boldsymbol{\delta}},\underline{\boldsymbol{\theta}}_{\mathbf{s}}^*} & \mathbf{F}_{\underline{\boldsymbol{\delta}},\underline{\boldsymbol{\theta}}_{\mathbf{n}}^*} & \mathbf{F}_{\underline{\boldsymbol{\delta}},\underline{\boldsymbol{\delta}}^*}\end{bmatrix} \tag{20}$$

where:

$$\left(\mathbf{F}_{\underline{\boldsymbol{\theta}}_{\mathbf{s}},\underline{\boldsymbol{\theta}}_{\mathbf{s}}^*}\right)_{k,l}=\int_{-\frac{B}{2}}^{\frac{B}{2}}\frac{\partial\mathbf{s}^H(f;\underline{\boldsymbol{\theta}}_{\mathbf{s}},\underline{\boldsymbol{\delta}})}{\partial(\underline{\boldsymbol{\theta}}_{\mathbf{s}})_k}\mathbf{R}_{\mathbf{n}}^{-1}\left(f;\underline{\boldsymbol{\theta}}_{\mathbf{n}},\underline{\boldsymbol{\delta}}\right)\frac{\partial\mathbf{s}(f;\underline{\boldsymbol{\theta}}_{\mathbf{s}},\underline{\boldsymbol{\delta}})}{\partial(\underline{\boldsymbol{\theta}}_{\mathbf{s}}^*)_l}df+\int_{-\frac{B}{2}}^{\frac{B}{2}}\frac{\partial\mathbf{s}^H(f;\underline{\boldsymbol{\theta}}_{\mathbf{s}},\underline{\boldsymbol{\delta}})}{\partial(\underline{\boldsymbol{\theta}}_{\mathbf{s}}^*)_l}\mathbf{R}_{\mathbf{n}}^{-1}\left(f;\underline{\boldsymbol{\theta}}_{\mathbf{n}},\underline{\boldsymbol{\delta}}\right)\frac{\partial\mathbf{s}(f;\underline{\boldsymbol{\theta}}_{\mathbf{s}},\underline{\boldsymbol{\delta}})}{\partial(\underline{\boldsymbol{\theta}}_{\mathbf{s}})_k}df \tag{21a}$$

$$\left(\mathbf{F}_{\underline{\boldsymbol{\theta}}_{\mathbf{n}},\underline{\boldsymbol{\theta}}_{\mathbf{n}}^*}\right)_{k,l}=T\int_{-\frac{B}{2}}^{\frac{B}{2}}\mathbf{R}_{\mathbf{n}}^{-1}\left(f;\underline{\boldsymbol{\theta}}_{\mathbf{n}},\underline{\boldsymbol{\delta}}\right)\frac{\partial\mathbf{R}_{\mathbf{n}}^{-1}\left(f;\underline{\boldsymbol{\theta}}_{\mathbf{n}},\underline{\boldsymbol{\delta}}\right)}{\partial\left(\underline{\boldsymbol{\theta}}_{\mathbf{n}}\right)_k}\mathbf{R}_{\mathbf{n}}^{-1}\left(f;\underline{\boldsymbol{\theta}}_{\mathbf{n}},\underline{\boldsymbol{\delta}}\right)\frac{\partial\mathbf{R}_{\mathbf{n}}^{-1}\left(f;\underline{\boldsymbol{\theta}}_{\mathbf{n}},\underline{\boldsymbol{\delta}}\right)}{\partial\left(\underline{\boldsymbol{\theta}}_{\mathbf{n}}^*\right)_l}df \tag{21b}$$

$$\left(\mathbf{F}_{\underline{\boldsymbol{\delta}},\underline{\boldsymbol{\delta}}^*}\right)_{k,l}=T\int_{-\frac{B}{2}}^{\frac{B}{2}}\mathbf{R}_{\mathbf{n}}^{-1}\left(f;\underline{\boldsymbol{\theta}}_{\mathbf{n}},\underline{\boldsymbol{\delta}}\right)\frac{\partial\mathbf{R}_{\mathbf{n}}^{-1}(f;\underline{\boldsymbol{\theta}}_{\mathbf{n}},\underline{\boldsymbol{\delta}})}{\partial(\underline{\boldsymbol{\delta}})_k}\mathbf{R}_{\mathbf{n}}^{-1}\left(f;\underline{\boldsymbol{\theta}}_{\mathbf{n}},\underline{\boldsymbol{\delta}}\right)\frac{\partial\mathbf{R}_{\mathbf{n}}^{-1}(f;\underline{\boldsymbol{\theta}}_{\mathbf{n}},\underline{\boldsymbol{\delta}})}{\partial(\underline{\boldsymbol{\delta}}^*)_l}df+\int_{-\frac{B}{2}}^{\frac{B}{2}}\frac{\partial\mathbf{s}^H(f;\underline{\boldsymbol{\theta}}_{\mathbf{s}},\underline{\boldsymbol{\delta}})}{\partial(\underline{\boldsymbol{\delta}})_k}\mathbf{R}_{\mathbf{n}}^{-1}\left(f;\underline{\boldsymbol{\theta}}_{\mathbf{n}},\underline{\boldsymbol{\delta}}\right)\frac{\partial\mathbf{s}(f;\underline{\boldsymbol{\theta}}_{\mathbf{s}},\underline{\boldsymbol{\delta}})}{\partial(\underline{\boldsymbol{\delta}}^*)_l}df+\int_{-\frac{B}{2}}^{\frac{B}{2}}\frac{\partial\mathbf{s}^H(f;\underline{\boldsymbol{\theta}}_{\mathbf{s}},\underline{\boldsymbol{\delta}})}{\partial(\underline{\boldsymbol{\delta}}^*)_l}\mathbf{R}_{\mathbf{n}}^{-1}\left(f;\underline{\boldsymbol{\theta}}_{\mathbf{n}},\underline{\boldsymbol{\delta}}\right)\frac{\partial\mathbf{s}(f;\underline{\boldsymbol{\theta}}_{\mathbf{s}},\underline{\boldsymbol{\delta}})}{\partial(\underline{\boldsymbol{\delta}})_k}df \tag{21c}$$

$$\left(\mathbf{F}_{\underline{\boldsymbol{\delta}},\underline{\boldsymbol{\theta}}_{\mathbf{s}}^*}\right)_{k,l}=\int_{-\frac{B}{2}}^{\frac{B}{2}}\frac{\partial\mathbf{s}^H(f;\underline{\boldsymbol{\theta}}_{\mathbf{s}},\underline{\boldsymbol{\delta}})}{\partial(\underline{\boldsymbol{\delta}})_k}\mathbf{R}_{\mathbf{n}}^{-1}\left(f;\underline{\boldsymbol{\theta}}_{\mathbf{n}},\underline{\boldsymbol{\delta}}\right)\frac{\partial\mathbf{s}(f;\underline{\boldsymbol{\theta}}_{\mathbf{s}},\underline{\boldsymbol{\delta}})}{\partial(\underline{\boldsymbol{\theta}}_{\mathbf{s}}^*)_l}df+\int_{-\frac{B}{2}}^{\frac{B}{2}}\frac{\partial\mathbf{s}^H(f;\underline{\boldsymbol{\theta}}_{\mathbf{s}},\underline{\boldsymbol{\delta}})}{\partial(\underline{\boldsymbol{\theta}}_{\mathbf{s}}^*)_l}\mathbf{R}_{\mathbf{n}}^{-1}\left(f;\underline{\boldsymbol{\theta}}_{\mathbf{n}},\underline{\boldsymbol{\delta}}\right)\frac{\partial\mathbf{s}(f;\underline{\boldsymbol{\theta}}_{\mathbf{s}},\underline{\boldsymbol{\delta}})}{\partial(\underline{\boldsymbol{\delta}})_k}df \tag{21d}$$

$$\left(\mathbf{F}_{\underline{\boldsymbol{\delta}},\underline{\boldsymbol{\theta}}_{\mathbf{n}}^*}\right)_{k,l}=T\int_{-\frac{B}{2}}^{\frac{B}{2}}\mathbf{R}_{\mathbf{n}}^{-1}\left(f;\underline{\boldsymbol{\theta}}_{\mathbf{n}},\underline{\boldsymbol{\delta}}\right)\frac{\partial\mathbf{R}_{\mathbf{n}}^{-1}\left(f;\underline{\boldsymbol{\theta}}_{\mathbf{n}},\underline{\boldsymbol{\delta}}\right)}{\partial\left(\underline{\boldsymbol{\delta}}\right)_k}\mathbf{R}_{\mathbf{n}}^{-1}\left(f;\underline{\boldsymbol{\theta}}_{\mathbf{n}},\underline{\boldsymbol{\delta}}\right)\frac{\partial\mathbf{R}_{\mathbf{n}}^{-1}\left(f;\underline{\boldsymbol{\theta}}_{\mathbf{n}},\underline{\boldsymbol{\delta}}\right)}{\partial\left(\underline{\boldsymbol{\theta}}_{\mathbf{n}}^*\right)_l}df \tag{21e}$$

First of all, to the best of our knowledge, (20) and (21a-21e) have never been released neither in papers [17][22][27][28][49] nor text books [1][39][43]. They encompass all previously released results on this topic, including:

• the standard narrow band case at a single frequency f_0 and a single observation time $(BT = 1)$, where (19) becomes [1][17]:

$$\mathbf{x}\left(f_0, \underline{\boldsymbol{\theta}}\right) = \mathbf{s}\left(f_0; \underline{\boldsymbol{\theta}}_{\mathbf{s}}, \underline{\boldsymbol{\delta}}\right) + \mathbf{n}\left(f_0; \underline{\boldsymbol{\theta}}_{\mathbf{n}}, \underline{\boldsymbol{\delta}}\right), \quad \mathbf{R}_{\mathbf{n}}\left(f; \underline{\boldsymbol{\theta}}_{\mathbf{n}}, \underline{\boldsymbol{\delta}}\right) = \mathbf{R}\left(\underline{\boldsymbol{\theta}}_{\mathbf{n}}, \underline{\boldsymbol{\delta}}\right) \delta\left(f - f_0\right)$$

• the temporally white noise case [17][22][27] where (19) becomes [43]:

$$\mathbf{x}\left(t; \underline{\boldsymbol{\theta}}\right) = \mathbf{s}\left(t; \underline{\boldsymbol{\theta}}_{\mathbf{s}}, \underline{\boldsymbol{\delta}}\right) + \mathbf{n}\left(t; \underline{\boldsymbol{\theta}}_{\mathbf{n}}, \underline{\boldsymbol{\delta}}\right), \quad \mathbf{R}_{\mathbf{n}}\left(f; \underline{\boldsymbol{\theta}}_{\mathbf{n}}, \underline{\boldsymbol{\delta}}\right) = \mathbf{R}_{\mathbf{n}}\left(0; \underline{\boldsymbol{\theta}}_{\mathbf{n}}, \underline{\boldsymbol{\delta}}\right) \Pi_{\left[-\frac{B}{2}, \frac{B}{2}\right]}\left(f\right)$$

2) FIM for multiple band limited conditional models:

Results released in the previous paragraph can be extended to the observation of L mutually Gaussian band limited N^l-dimensional observation models with band B^l :

$$\mathbf{x}^l\left(t; \underline{\boldsymbol{\theta}}^l\right) = \mathbf{s}^l\left(t; \underline{\boldsymbol{\theta}}^l_{\mathbf{s}}, \underline{\boldsymbol{\delta}}^l\right) + \mathbf{n}^l\left(t; \underline{\boldsymbol{\theta}}^l_{\mathbf{n}}, \underline{\boldsymbol{\delta}}^l\right), \left(\underline{\boldsymbol{\theta}}^l\right)^T = \left(\left(\underline{\boldsymbol{\theta}}^l_{\mathbf{s}}\right)^T, \left(\underline{\boldsymbol{\theta}}^l_{\mathbf{n}}\right)^T, \left(\underline{\boldsymbol{\delta}}^l\right)^T\right), \begin{array}{c} t - t_0^l \in \left[0, T^l\right] \\ 1 \leq l \leq L \end{array} \tag{22}$$

Let $\underline{\mathbf{O}}_{\mathbf{s}}^T = \left(\left(\underline{\boldsymbol{\theta}}^1_{\mathbf{s}}\right)^T, \ldots, \left(\underline{\boldsymbol{\theta}}^L_{\mathbf{s}}\right)^T\right)$, $\underline{\mathbf{O}}_{\mathbf{n}}^T = \left(\left(\underline{\boldsymbol{\theta}}^1_{\mathbf{n}}\right)^T, \ldots, \left(\underline{\boldsymbol{\theta}}^L_{\mathbf{n}}\right)^T\right)$, $\underline{\boldsymbol{\Delta}}^T = \left(\left(\underline{\boldsymbol{\delta}}^1\right)^T, \ldots, \left(\underline{\boldsymbol{\delta}}^L\right)^T\right)$ and $\underline{\mathbf{O}}^T = \left(\underline{\mathbf{O}}_{\mathbf{s}}^T, \underline{\mathbf{O}}_{\mathbf{n}}^T, \underline{\boldsymbol{\Delta}}^T\right)$. Then $\mathbf{X}\left(t; \underline{\mathbf{O}}\right)^T = \left(\mathbf{x}^1\left(t; \underline{\boldsymbol{\theta}}^1\right)^T, \ldots, \mathbf{x}^L\left(t; \underline{\boldsymbol{\theta}}^L\right)^T\right)$ is a N-dimensional $\left(N = \sum_{l=1}^L N^l\right)$ band limited conditional model with bandwidth $B = \max\limits_{1 \leq l \leq L}\left\{B^l\right\}$ and observed for a duration T (length of the time interval containing most (as much as desired) of the energy of all $\mathbf{s}^l\left(t; \underline{\boldsymbol{\theta}}^l_{\mathbf{s}}, \underline{\boldsymbol{\delta}}^l\right)$) :

$$\mathbf{X}\left(t; \underline{\mathbf{O}}\right) = \mathbf{S}\left(t; \underline{\mathbf{O}}_{\mathbf{s}}, \underline{\boldsymbol{\Delta}}\right) + \mathbf{N}\left(t; \underline{\mathbf{O}}_{\mathbf{n}}, \underline{\boldsymbol{\Delta}}\right), \quad t - t_0 \in [0, T] \tag{23}$$

where $t_0 = \min\limits_{1 \leq l \leq L}\left\{t_0^l\right\}$, $T \leq \sum_{l=1}^L T^l$, $\mathbf{S}\left(t; \underline{\mathbf{O}}_{\mathbf{s}}, \underline{\boldsymbol{\Delta}}\right)^T = \left(\mathbf{s}^1\left(t; \underline{\boldsymbol{\theta}}^1_{\mathbf{s}}, \underline{\boldsymbol{\delta}}^1\right)^T, \ldots, \mathbf{s}^L\left(t; \underline{\boldsymbol{\theta}}^L_{\mathbf{s}}, \underline{\boldsymbol{\delta}}^L\right)^T\right)$ and $\mathbf{N}\left(t; \underline{\mathbf{O}}_{\mathbf{s}}, \underline{\boldsymbol{\Delta}}\right)^T = \left(\mathbf{n}^1\left(t; \underline{\boldsymbol{\theta}}^1_{\mathbf{n}}, \underline{\boldsymbol{\delta}}^1\right)^T, \ldots, \mathbf{n}^L\left(t; \underline{\boldsymbol{\theta}}^L_{\mathbf{n}}, \underline{\boldsymbol{\delta}}^L\right)^T\right)$. The FIM associated to (23) is an updated form of (20):

$$\mathbf{F}_{\underline{\mathbf{O}}} = \mathbf{F}_{\underline{\mathbf{O}}, \underline{\mathbf{O}}^*} = \begin{bmatrix} \mathbf{F}_{\underline{\mathbf{O}}_{\mathbf{s}}, \underline{\mathbf{O}}_{\mathbf{s}}^*} & \mathbf{0} & \mathbf{F}^H_{\underline{\boldsymbol{\Delta}}, \underline{\mathbf{O}}_{\mathbf{s}}^*} \\ \mathbf{0} & \mathbf{F}_{\underline{\mathbf{O}}_{\mathbf{n}}, \underline{\mathbf{O}}_{\mathbf{n}}^*} & \mathbf{F}^H_{\underline{\boldsymbol{\Delta}}, \underline{\mathbf{O}}_{\mathbf{n}}^*} \\ \mathbf{F}_{\underline{\boldsymbol{\Delta}}, \underline{\mathbf{O}}_{\mathbf{s}}^*} & \mathbf{F}_{\underline{\boldsymbol{\Delta}}, \underline{\mathbf{O}}_{\mathbf{n}}^*} & \mathbf{F}_{\underline{\boldsymbol{\Delta}}, \underline{\boldsymbol{\Delta}}^*} \end{bmatrix} \tag{24}$$

where (21a-21e) are updated by making the following substitutions : $\underline{\boldsymbol{\theta}}_{\mathbf{s}} \rightarrow \underline{\mathbf{O}}_{\mathbf{s}}$, $\underline{\boldsymbol{\theta}}_{\mathbf{n}} \rightarrow \underline{\mathbf{O}}_{\mathbf{n}}$, $\underline{\boldsymbol{\delta}} \rightarrow \underline{\boldsymbol{\Delta}}$, $\mathbf{R}_{\mathbf{n}}\left(f; \underline{\boldsymbol{\theta}}_{\mathbf{n}}, \underline{\boldsymbol{\delta}}\right) \rightarrow \mathbf{R}_{\mathbf{N}}\left(f; \underline{\mathbf{O}}_{\mathbf{n}}, \underline{\boldsymbol{\Delta}}\right)$, $\mathbf{s}\left(f; \underline{\boldsymbol{\theta}}_{\mathbf{s}}, \underline{\boldsymbol{\delta}}\right) \rightarrow \mathbf{S}\left(t; \underline{\mathbf{O}}_{\mathbf{s}}, \underline{\boldsymbol{\Delta}}\right)$.

B. Estimation performance with constrained CRB: ideal conditional model of L independent observations

This section is dedicated to apply the proposed framework in the case of the standard (restricted) conditional model of the open literature (30), for which we exhibit some particular cases where

constraints can be used efficiently to generate extension of existing closed-form expressions of CRB of parameters of interest.

1) FIM for the ideal conditional model of L independent observations:

To start with, (20) clearly reveals the little-known correlation between signal and noise parameters (in the sense that asymptotic estimation performance of $\underline{\boldsymbol{\theta}}_{\mathbf{s}}$ is dependent of the state - known or unknown - of $\underline{\boldsymbol{\theta}}_{\mathbf{n}}$, and vice versa) in the general operational case, i.e. when the observation model attempt to take into account an actual observation where noise occurs in the presence of sensors modeling errors ($\underline{\boldsymbol{\delta}}$), since then $\mathbf{F}_{\underline{\boldsymbol{\delta}},\underline{\boldsymbol{\theta}}_{\mathbf{s}}^*} \neq \mathbf{0}$ a priori.

In this case, the form of (24) generally precludes the derivation of a closed-form expression of the CRB for $\underline{\mathbf{O}}_{\mathbf{s}}$. It is the reason why most papers or reference books, following the historical approach of deriving closed-form expressions of CRB, restrict the observation model (19) to the case where $\underline{\boldsymbol{\Delta}} = \varnothing$, since this simplification allows to analyze signal estimation performance from $\mathbf{F}_{\underline{\mathbf{O}}_{\mathbf{s}},\underline{\mathbf{O}}_{\mathbf{s}}^*}$ only. Under this assumption (24) simply reduces to:

$$\mathbf{F}_{\underline{\mathbf{O}}} = \mathbf{F}_{\underline{\mathbf{O}},\underline{\mathbf{O}}^*} = \begin{bmatrix} \mathbf{F}_{\underline{\mathbf{O}}_{\mathbf{s}},\underline{\mathbf{O}}_{\mathbf{s}}^*} & \mathbf{0} \\ \mathbf{0} & \mathbf{F}_{\underline{\mathbf{O}}_{\mathbf{n}},\underline{\mathbf{O}}_{\mathbf{n}}^*} \end{bmatrix}, \quad \underline{\mathbf{O}}^T = \left(\underline{\mathbf{O}}_{\mathbf{s}}^T, \underline{\mathbf{O}}_{\mathbf{n}}^T\right) \tag{25}$$

From a system design point of view, the hypothesis $\underline{\boldsymbol{\Delta}} = \varnothing$ means:

• either that all the parameters of the system components (sensors location and radiation patterns, transmitters and receiver electronics, ...) yielding the L observations are known (perfect calibration of the system components),

• or that matrices $\mathbf{R}_{\mathbf{n}}^l\left(f, \underline{\boldsymbol{\theta}}_{\mathbf{n}}^l, \underline{\boldsymbol{\delta}}^l\right) = \mathbf{R}_{\mathbf{n}}^l\left(f, \underline{\boldsymbol{\theta}}_{\mathbf{n}}^l\right)$, that is are known up to parameters not part of $\mathbf{s}^l\left(t; \underline{\boldsymbol{\theta}}_{\mathbf{s}}^l\right)$ parametric model. For instance this may happen in absence of interference where $\mathbf{R}_{\mathbf{n}}^l\left(f, \underline{\boldsymbol{\theta}}_{\mathbf{n}}^l\right)$ reduces to the thermal noise correlation matrix and the receiver electronics is perfectly known (from sensors output up to Hilbert filters); or if $\mathbf{R}_{\mathbf{n}}^l\left(f, \underline{\boldsymbol{\theta}}_{\mathbf{n}}^l\right)$ has been estimated (measured accurately), for instance from a secondary set of data consisting of noise samples only.

An alternative expression of (25) can be obtained by resorting to constrained CRB with constraints related with the state - known or unknown - of $\underline{\boldsymbol{\Delta}}$. Indeed, the vector of parameters $\underline{\boldsymbol{\Delta}}$ is known if it verifies the following set of constraints:

$$\underline{\boldsymbol{\Delta}} = \underline{\boldsymbol{\Delta}}_0 \quad \Leftrightarrow \quad \underline{\mathbf{f}}\left(\underline{\mathbf{O}}\right) = \underline{\boldsymbol{\Delta}} - \underline{\boldsymbol{\Delta}}_0 = \mathbf{0} \tag{26}$$

Then:

$$\frac{\partial \underline{\mathbf{f}}\left(\underline{\mathbf{O}}\right)}{\partial \underline{\mathbf{O}}^T} = \frac{\partial\left(\underline{\boldsymbol{\Delta}} - \underline{\boldsymbol{\Delta}}_0\right)}{\partial \underline{\mathbf{O}}^T} = \begin{bmatrix} \mathbf{0} & \mathbf{0} & \mathbf{I} \end{bmatrix} \in \mathcal{M}\left(\dim\left(\underline{\boldsymbol{\Delta}}\right), \dim\left(\underline{\mathbf{O}}\right)\right),$$

and:

$$\mathbf{U}_{\underline{\mathbf{O}}} = \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \in \mathcal{M}\left(\dim\left(\underline{\mathbf{O}}\right), \dim\left(\underline{\mathbf{O}}_{\mathbf{s}}\right) + \dim\left(\underline{\mathbf{O}}_{\mathbf{n}}\right)\right) \text{ is a basis of } \ker\left\{\frac{\partial \underline{\mathbf{f}}\left(\underline{\mathbf{O}}\right)}{\partial \underline{\mathbf{O}}^T}\right\}, \tag{27}$$

what leads to the constrained FIM (15):

$$\mathbf{F}^{c}_{\underline{\mathbf{O}}} = \mathbf{U}^{T}_{\underline{\mathbf{O}}} \mathbf{F}_{\underline{\mathbf{O}}} \mathbf{U}^{*}_{\underline{\mathbf{O}}} = \begin{bmatrix} \mathbf{F}_{\underline{\mathbf{O}}_{\mathbf{s}},\underline{\mathbf{O}}^{*}_{\mathbf{s}}} & \mathbf{0} \\ \mathbf{0} & \mathbf{F}_{\underline{\mathbf{O}}_{\mathbf{n}},\underline{\mathbf{O}}^{*}_{\mathbf{n}}} \end{bmatrix} \tag{28}$$

and constrained CRB (14):

$$\mathbf{CRB}^{c}_{\underline{\mathbf{O}}|\underline{\mathbf{O}}} \left(\underline{\mathbf{O}} \right) = \mathbf{U}^{*}_{\underline{\mathbf{O}}} \left(\mathbf{F}^{c}_{\underline{\mathbf{O}}} \right)^{-1} \mathbf{U}^{T}_{\underline{\mathbf{O}}} = \begin{bmatrix} \mathbf{F}^{-1}_{\underline{\mathbf{O}}_{\mathbf{s}},\underline{\mathbf{O}}^{*}_{\mathbf{s}}} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{F}^{-1}_{\underline{\mathbf{O}}_{\mathbf{n}},\underline{\mathbf{O}}^{*}_{\mathbf{n}}} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \end{bmatrix} \tag{29}$$

where $\underline{\mathbf{O}}_{\mathbf{s}}$ and $\underline{\mathbf{O}}_{\mathbf{n}}$ are decorrelated parameters. The all-zero rows of $\mathbf{U}_{\underline{\mathbf{O}}}$ act to remove the corresponding rows and columns from $\mathbf{F}_{\underline{\mathbf{O}}}$ via matrix multiplication. Thus, the contribution of the known parameters to the constrained FIM is removed. Then, the outer transformation acts to restore the CRB matrix to its original size by inserting all-zero rows and columns for all values of that are constrained to be known.

To keep on following the historical approach of deriving closed-form expressions of CRB, we consider the most studied realization of (22) in signal processing, that is the problem of fitting a model composed of a number M of superimposed signals to noisy data when L independent observations are available and (26) is verified (i.e. $\underline{\boldsymbol{\Delta}}$ is known which amounts to drop $\underline{\boldsymbol{\Delta}}$ from the set of unknown parameters for sake of legibility):

$$\mathbf{x}^{l}\left(t;\underline{\boldsymbol{\theta}}^{l}\right) = \mathbf{s}^{l}\left(t;\underline{\boldsymbol{\theta}}^{l}_{\mathbf{s}}\right) + \mathbf{n}^{l}\left(t;\underline{\boldsymbol{\theta}}^{l}_{\mathbf{n}}\right), \quad \mathbf{s}^{l}\left(t;\underline{\boldsymbol{\theta}}^{l}_{\mathbf{s}}\right) = \mathbf{B}^{l}_{\mathbf{s}}\left(t;\boldsymbol{\Xi}^{l}_{\mathbf{s}}\right)\boldsymbol{\sigma}^{l}_{\mathbf{s}} \tag{30}$$

$$\left(\underline{\boldsymbol{\theta}}^{l}\right)^{T} = \left(\left(\underline{\boldsymbol{\theta}}^{l}_{\mathbf{s}}\right)^{T}, \left(\underline{\boldsymbol{\theta}}^{l}_{\mathbf{n}}\right)^{T}\right), \quad \left(\underline{\boldsymbol{\theta}}^{l}_{\mathbf{s}}\right)^{T} = \left(\left(\underline{\boldsymbol{\sigma}}^{l}_{\mathbf{s}}\right)^{T}, \left(\boldsymbol{\Xi}^{l}_{\mathbf{s}}\right)^{T}\right), \quad \left(\boldsymbol{\Xi}^{l}_{\mathbf{s}}\right)^{T} = \left(\left(\boldsymbol{\varepsilon}^{l}_{1}\right)^{T}, \ldots, \left(\boldsymbol{\varepsilon}^{l}_{M}\right)^{T}\right)$$

where:

• $\mathbf{B}^{l}_{\mathbf{s}}\left(t;\boldsymbol{\Xi}^{l}_{\mathbf{s}}\right) = \left[\mathbf{b}^{l}_{\mathbf{s}}\left(t;\boldsymbol{\varepsilon}^{l}_{1}\right), \ldots, \mathbf{b}^{l}_{\mathbf{s}}\left(t;\boldsymbol{\varepsilon}^{l}_{M}\right)\right]$ and $\mathbf{b}^{l}_{\mathbf{s}}\left(t;\boldsymbol{\varepsilon}^{l}\right)$ is a vector of N parametric functions (spatial, temporal, ... transfer function) depending on a vector $\boldsymbol{\varepsilon}^{l}$ of P real parameters of interest associated with a given signal (DOAs, delay, velocity, ...),

• $\boldsymbol{\sigma}^{l}_{\mathbf{s}} = \left(\left(\boldsymbol{\sigma}^{l}_{\mathbf{s}}\right)_{1}, \ldots, \left(\boldsymbol{\sigma}^{l}_{\mathbf{s}}\right)_{M}\right)^{T}$ is the vector of complex amplitudes of the M signals for the l^{th} observation model (nuisance parameters),

• $\mathbf{n}^{l}\left(\underline{\boldsymbol{\theta}}^{l}_{\mathbf{n}}\right)$ is a noise vector depending on a vector $\underline{\boldsymbol{\theta}}^{l}_{\mathbf{n}}$ of parameters (DOAs, correlation matrices, ...), which is clearly an ideal conditional model since all functions $\mathbf{b}^{l}_{\mathbf{s}}\left(t;\boldsymbol{\varepsilon}^{l}\right)$ are assumed to be perfectly known.

Since we consider independent observation models, the associated FIM is given by (25) where:

$$\mathbf{F}_{\underline{\mathbf{O}}_{\mathbf{s}}} = \begin{bmatrix} \mathbf{F}_{\underline{\boldsymbol{\theta}}^{1}_{\mathbf{s}}} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \ddots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{F}_{\underline{\boldsymbol{\theta}}^{L}_{\mathbf{s}}} \end{bmatrix} \triangleq \mathbf{F}_{\underline{\mathbf{O}}_{\mathbf{s}}}\left(\mathbf{O}\right) = \begin{bmatrix} \mathbf{F}_{\underline{\boldsymbol{\theta}}^{1}_{\mathbf{s}}}\left(\underline{\boldsymbol{\theta}}^{1}\right) & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \ddots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{F}_{\underline{\boldsymbol{\theta}}^{L}_{\mathbf{s}}}\left(\underline{\boldsymbol{\theta}}^{L}\right) \end{bmatrix} \triangleq \mathbf{F}_{\underline{\mathbf{O}}_{\mathbf{s}},\underline{\mathbf{O}}^{*}_{\mathbf{s}}}\left(\mathbf{O}\right). \tag{31}$$

Then it is worth considering the following dedicated formalism [48][42] to obtain closed-form expression of $\mathbf{F}_{\underline{\boldsymbol{\theta}}^l_{\mathbf{s}}}\left(\underline{\boldsymbol{\theta}}^l\right)$.

First, the dependency of vectors and matrices on frequency f, e.g. $\mathbf{s}^l\left(f;\underline{\boldsymbol{\theta}}^l_{\mathbf{s}}\right) \rightleftharpoons \mathbf{s}^l\left(t;\underline{\boldsymbol{\theta}}^l_{\mathbf{s}}\right)$, $\mathbf{b}^l_{\mathbf{s}}\left(f;\varepsilon^l\right) \rightleftharpoons \mathbf{b}^l_{\mathbf{s}}\left(t;\varepsilon^l\right)$, $\mathbf{B}^l_{\mathbf{s}}\left(f;\boldsymbol{\Xi}^l_{\mathbf{s}}\right) \rightleftharpoons \mathbf{B}^l_{\mathbf{s}}\left(t;\boldsymbol{\Xi}^l_{\mathbf{s}}\right)$ ..., where $\rightleftharpoons$ denotes the Fourier transform, is omitted wherever this omission is unambiguous. Second, let notation $\{\mathbf{A}\left(f\right)\}$ be the generalization of (7) denoting the family of column vectors of matrix $\mathbf{A}\left(f\right)$:

$$\{\mathbf{A}\left(f\right)\} = \{[\mathbf{a}_1\left(f\right)\ldots\mathbf{a}_Q\left(f\right)]\} = \{\mathbf{a}_1\left(f\right),\ldots,\mathbf{a}_Q\left(f\right)\} \tag{32}$$

Then, if $\mathbf{A}\left(f\right) = [\mathbf{a}_1\left(f\right)\ldots\mathbf{a}_Q\left(f\right)]$ and $\mathbf{C}\left(f\right) = [\mathbf{c}_1\left(f\right)\ldots\mathbf{c}_P\left(f\right)]$, $\mathbf{G}_{\boldsymbol{\theta}_{\mathbf{n}}}\left(\{\mathbf{A}\},\{\mathbf{C}\}\right)$ is the $P\times Q$ complex matrix defined by:

$$\mathbf{G}_{\boldsymbol{\theta}_{\mathbf{n}}}\left(\{\mathbf{A}\},\{\mathbf{C}\}\right)_{p,q} = \langle\mathbf{a}_q \mid \mathbf{c}_p\rangle_{\underline{\boldsymbol{\theta}}_{\mathbf{n}}} = \langle\mathbf{a}_q\left(f\right) \mid \mathbf{c}_p\left(f\right)\rangle_{\underline{\boldsymbol{\theta}}_{\mathbf{n}}} = \int\limits_{-\frac{B}{2}}^{\frac{B}{2}} \mathbf{c}_p^H\left(f\right)\mathbf{R}_{\mathbf{n}}^{-1}\left(f;\underline{\boldsymbol{\theta}}_{\mathbf{n}}\right)\mathbf{a}_q\left(f\right)df \tag{33}$$

Let $\Pi_{\{\mathbf{B}(\boldsymbol{\Xi})\}}$ denote the orthonormal projector on $span\left\{\mathbf{B}\left(f;\boldsymbol{\Xi}\right)\right\}$, i.e. the span of the vector columns of matrix $\mathbf{B}\left(f;\boldsymbol{\Xi}\right)$:

$$\Pi_{\{\mathbf{B}(\boldsymbol{\Xi})\}}\left(\mathbf{a}\right) = \Pi_{\{\mathbf{B}(f;\boldsymbol{\Xi})\}}\left(\mathbf{a}\left(f\right)\right) = \mathbf{B}\left(f;\boldsymbol{\Xi}\right)\mathbf{G}^{-1}_{\boldsymbol{\theta}_{\mathbf{n}}}\left(\{\mathbf{B}\left(\boldsymbol{\Xi}\right)\}\right)\mathbf{G}_{\boldsymbol{\theta}_{\mathbf{n}}}\left(\mathbf{a},\{\mathbf{B}\left(\boldsymbol{\Xi}\right)\}\right)$$

and let $\Pi^{\perp}_{\{\mathbf{B}(\boldsymbol{\Xi})\}}$ denote the orthonormal projector on the orthogonal complement of $span\left\{\mathbf{B}\left(f;\boldsymbol{\Xi}\right)\right\}$:

$$\Pi_{\{\mathbf{B}(\boldsymbol{\Xi})\}}\left(\mathbf{a}\right) + \Pi^{\perp}_{\{\mathbf{B}(\boldsymbol{\Xi})\}}\left(\mathbf{a}\right) = \mathbf{a} \tag{34}$$

Let us define:

$$\Pi_{\{\mathbf{B}(\boldsymbol{\Xi})\}}\left(\{\mathbf{A}\}\right) = \left\{\Pi_{\{\mathbf{B}(\boldsymbol{\Xi})\}}\left(\mathbf{a}_1\right),\ldots,\Pi_{\{\mathbf{B}(\boldsymbol{\Xi})\}}\left(\mathbf{a}_Q\right)\right\} \tag{35}$$

then:

$$\begin{aligned}\mathbf{G}_{\boldsymbol{\theta}_{\mathbf{n}}}\left(\Pi^{\perp}_{\{\mathbf{B}(\boldsymbol{\Xi})\}}\left(\{\mathbf{A}\}\right),\Pi^{\perp}_{\{\mathbf{B}(\boldsymbol{\Xi})\}}\left(\{\mathbf{C}\}\right)\right)_{p,q} &= \mathbf{G}_{\boldsymbol{\theta}_{\mathbf{n}}}\left(\Pi^{\perp}_{\{\mathbf{B}(\boldsymbol{\Xi})\}}\left(\mathbf{a}_q\right),\Pi^{\perp}_{\{\mathbf{B}(\boldsymbol{\Xi})\}}\left(\mathbf{c}_p\right)\right)\\ &= \mathbf{G}_{\boldsymbol{\theta}_{\mathbf{n}}}\left(\mathbf{a}_q,\mathbf{c}_p\right) - \mathbf{G}^H_{\boldsymbol{\theta}_{\mathbf{n}}}\left(\mathbf{c}_p,\{\mathbf{B}\left(\boldsymbol{\Xi}\right)\}\right)\mathbf{G}^{-1}_{\boldsymbol{\theta}_{\mathbf{n}}}\left(\{\mathbf{B}\left(\boldsymbol{\Xi}\right)\}\right)\mathbf{G}_{\boldsymbol{\theta}_{\mathbf{n}}}\left(\mathbf{a}_q,\{\mathbf{B}\left(\boldsymbol{\Xi}\right)\}\right)\end{aligned}$$

and [42][48]:

$$\mathbf{F}_{\underline{\boldsymbol{\theta}}^l_{\mathbf{s}}} = \begin{bmatrix} \mathbf{F}_{\boldsymbol{\sigma}^l_{\mathbf{s}},(\boldsymbol{\sigma}^l_{\mathbf{s}})^*} & \mathbf{0} & \mathbf{F}_{\boldsymbol{\sigma}^l_{\mathbf{s}},\boldsymbol{\Xi}^l_{\mathbf{s}}} \\ \mathbf{0} & \mathbf{F}^*_{\boldsymbol{\sigma}^l_{\mathbf{s}},(\boldsymbol{\sigma}^l_{\mathbf{s}})^*} & \mathbf{F}^*_{\boldsymbol{\sigma}^l_{\mathbf{s}},\boldsymbol{\Xi}^l_{\mathbf{s}}} \\ \mathbf{F}^H_{\boldsymbol{\sigma}^l_{\mathbf{s}},\boldsymbol{\Xi}^l_{\mathbf{s}}} & \mathbf{F}^T_{\boldsymbol{\sigma}^l_{\mathbf{s}},\boldsymbol{\Xi}^l_{\mathbf{s}}} & \mathbf{F}_{\boldsymbol{\Xi}^l_{\mathbf{s}},\boldsymbol{\Xi}^l_{\mathbf{s}}} \end{bmatrix}, \begin{array}{l} \mathbf{F}_{\boldsymbol{\sigma}^l_{\mathbf{s}},(\boldsymbol{\sigma}^l_{\mathbf{s}})^*} = \mathbf{G}_{\underline{\boldsymbol{\theta}}^l_{\mathbf{n}}}\left(\left\{\mathbf{B}^l_{\mathbf{s}}\left(\boldsymbol{\Xi}^l_{\mathbf{s}}\right)\right\},\left\{\mathbf{B}^l_{\mathbf{s}}\left(\boldsymbol{\Xi}^l_{\mathbf{s}}\right)\right\}\right) \\ \mathbf{F}_{\boldsymbol{\sigma}^l_{\mathbf{s}},\boldsymbol{\Xi}^l_{\mathbf{s}}} = \mathbf{G}_{\underline{\boldsymbol{\theta}}^l_{\mathbf{n}}}\left(\left\{\frac{\partial\mathbf{s}^l\left(\underline{\boldsymbol{\theta}}^l_{\mathbf{s}}\right)}{\partial\left(\boldsymbol{\Xi}^l_{\mathbf{s}}\right)^T}\right\},\left\{\mathbf{B}^l_{\mathbf{s}}\left(\boldsymbol{\Xi}^l_{\mathbf{s}}\right)\right\}\right) \\ \mathbf{F}_{\boldsymbol{\Xi}^l_{\mathbf{s}},\boldsymbol{\Xi}^l_{\mathbf{s}}} = 2\,\mathrm{Re}\left\{\mathbf{G}_{\underline{\boldsymbol{\theta}}^l_{\mathbf{n}}}\left(\left\{\frac{\partial\mathbf{s}^l\left(\underline{\boldsymbol{\theta}}^l_{\mathbf{s}}\right)}{\partial\left(\boldsymbol{\Xi}^l_{\mathbf{s}}\right)^T}\right\},\left\{\frac{\partial\mathbf{s}^l\left(\underline{\boldsymbol{\theta}}^l_{\mathbf{s}}\right)}{\partial\left(\boldsymbol{\Xi}^l_{\mathbf{s}}\right)^T}\right\}\right)\right\} \end{array} \tag{36}$$

where:

$$\mathbf{F}_{\boldsymbol{\sigma}_{\mathbf{s}}^l,\boldsymbol{\Xi}_{\mathbf{s}}^l} = \mathbf{D}^l_{\boldsymbol{\Xi}_{\mathbf{s}}}\left(\underline{\boldsymbol{\theta}}^l\right) \odot \left(\left(\boldsymbol{\sigma}_{\mathbf{s}}^l\right)^T \otimes \mathbf{1}_{1\times P}\right), \qquad \mathbf{F}_{\boldsymbol{\Xi}_{\mathbf{s}}^l,\boldsymbol{\Xi}_{\mathbf{s}}^l} = 2\,\mathrm{Re}\left\{\mathbf{C}^l_{\boldsymbol{\Xi}_{\mathbf{s}}^l}\left(\underline{\boldsymbol{\theta}}^l\right) \odot \left(\left(\boldsymbol{\sigma}_{\mathbf{s}}^l\left(\boldsymbol{\sigma}_{\mathbf{s}}^l\right)^H\right)^T \otimes \mathbf{1}_{P\times P}\right)\right\},$$

$$\mathbf{D}^l_{\boldsymbol{\Xi}_{\mathbf{s}}^l}\left(\underline{\boldsymbol{\theta}}^l\right) = \left[\begin{array}{ccc} \mathbf{D}^l_{\boldsymbol{\Xi}_{\mathbf{s}}^l}\left(\underline{\boldsymbol{\theta}}^l\right)_1 & \cdots & \mathbf{D}^l_{\boldsymbol{\Xi}_{\mathbf{s}}^l}\left(\underline{\boldsymbol{\theta}}^l\right)_M \end{array}\right], \qquad \mathbf{D}^l_{\boldsymbol{\Xi}_{\mathbf{s}}^l}\left(\underline{\boldsymbol{\theta}}^l\right)_m = \mathbf{G}_{\underline{\boldsymbol{\theta}}^l_{\mathbf{n}}}\left(\left\{\frac{\partial \mathbf{b}^l_{\mathbf{s}}\left(\boldsymbol{\epsilon}^l_m\right)}{\partial\left(\boldsymbol{\epsilon}^l\right)^T}\right\}, \left\{\mathbf{B}^l_{\mathbf{s}}\left(\boldsymbol{\Xi}^l_{\mathbf{s}}\right)\right\}\right),$$

$$\mathbf{C}^l_{\boldsymbol{\Xi}_{\mathbf{s}}^l}\left(\underline{\boldsymbol{\theta}}^l\right) = \left[\begin{array}{ccc} \mathbf{C}^l_{\boldsymbol{\Xi}_{\mathbf{s}}^l}\left(\underline{\boldsymbol{\theta}}^l\right)_{1,1} & \cdots & \mathbf{C}^l_{\boldsymbol{\Xi}_{\mathbf{s}}^l}\left(\underline{\boldsymbol{\theta}}^l\right)_{1,M} \\ \vdots & \ddots & \vdots \\ \mathbf{C}^l_{\boldsymbol{\Xi}_{\mathbf{s}}^l}\left(\underline{\boldsymbol{\theta}}^l\right)_{M,1} & \cdots & \mathbf{C}^l_{\boldsymbol{\Xi}_{\mathbf{s}}^l}\left(\underline{\boldsymbol{\theta}}^l\right)_{M,M} \end{array}\right], \quad \mathbf{C}^l_{\boldsymbol{\Xi}_{\mathbf{s}}^l}\left(\underline{\boldsymbol{\theta}}^l\right)_{m_1,m_2} = \mathbf{G}_{\underline{\boldsymbol{\theta}}^l_{\mathbf{n}}}\left(\left\{\frac{\partial \mathbf{b}^l_{\mathbf{s}}\left(\boldsymbol{\epsilon}^l_{m_2}\right)}{\partial\left(\boldsymbol{\epsilon}^l\right)^T}\right\}, \left\{\frac{\partial \mathbf{b}^l_{\mathbf{s}}\left(\boldsymbol{\epsilon}^l_{m_1}\right)}{\partial\left(\boldsymbol{\epsilon}^l\right)^T}\right\}\right).$$

2) Derivation of analytical expressions with constrained CRB :

We consider separable constraints: individual constraints are functions of $\underline{\mathbf{O}}_{\mathbf{s}}$ or $\underline{\mathbf{O}}_{\mathbf{n}}$ only. We note that this does not preclude mixing constraints on both $\underline{\mathbf{O}}_{\mathbf{s}}$ and $\underline{\mathbf{O}}_{\mathbf{n}}$: only that each individual constraint is a function of one or the other. A first consequence is that mixing separable subsets of non redundant constraints on $\underline{\mathbf{O}}_{\mathbf{s}}$ and $\underline{\mathbf{O}}_{\mathbf{n}}$ ($\underline{\mathbf{f}}_{\mathbf{s}}\left(\underline{\mathbf{O}}_{\mathbf{s}}\right) = \mathbf{0}$, $\underline{\mathbf{f}}_{\mathbf{n}}\left(\underline{\mathbf{O}}_{\mathbf{n}}\right) = \mathbf{0}$) always yields a global set ($\underline{\mathbf{f}}\left(\underline{\mathbf{O}}\right) = \mathbf{0}$) of non redundant constraints (provided the total number of constraints does not exceed $\dim\left(\underline{\mathbf{O}}\right)$). Additionally the block diagonal structure of $\mathbf{F}_{\underline{\mathbf{O}}}$ (25) combined with the separable constraints assumption leads to (see [21, Subsection II.A][48]):

$$\mathbf{CRB}_{\underline{\mathbf{O}}_{\mathbf{s}}|\underline{\mathbf{O}}} = \mathbf{F}^{-1}_{\underline{\mathbf{O}}_{\mathbf{s}}}, \quad \mathbf{CRB}^{c}_{\underline{\mathbf{O}}_{\mathbf{s}}|\underline{\mathbf{O}}} = \mathbf{U}^*_{\underline{\mathbf{O}}_s}\left(\mathbf{U}^T_{\underline{\mathbf{O}}_s}\mathbf{F}_{\underline{\mathbf{O}}_{\mathbf{s}}}\mathbf{U}^*_{\underline{\mathbf{O}}_s}\right)^{-1}\mathbf{U}^T_{\underline{\mathbf{O}}_s} \tag{37}$$

where $\mathbf{U}_{\underline{\mathbf{O}}_s}$ is a basis of $\ker\left\{\frac{\partial \underline{\mathbf{f}}_{\mathbf{s}}\left(\underline{\mathbf{O}}_{\mathbf{s}}\right)}{\partial \underline{\mathbf{O}}^T}\right\}$.

Whatever the constraints $\underline{\mathbf{f}}_{\mathbf{s}}\left(\underline{\mathbf{O}}_{\mathbf{s}}\right) = \mathbf{0}$ considered, once derived a basis $\mathbf{U}_{\underline{\mathbf{O}}_s}$ of $\ker\left\{\frac{\partial \underline{\mathbf{f}}_{\mathbf{s}}\left(\underline{\mathbf{O}}_{\mathbf{s}}\right)}{\partial \underline{\mathbf{O}}^T}\right\}$, the analytical expression of the associated constrained CRB (37) for the parameters of interest $\left(\left(\boldsymbol{\Xi}^1_{\mathbf{s}}\right)^T, \ldots, \left(\boldsymbol{\Xi}^L_{\mathbf{s}}\right)^T\right)$ of the M signals simply results from combination of three algebraic operations: multiplication of partitioned matrices $\left(\mathbf{U}^T_{\underline{\mathbf{O}}_s}\mathbf{F}_{\underline{\mathbf{O}}_{\mathbf{s}}}\mathbf{U}^*_{\underline{\mathbf{O}}_s}\right)$, inversion of a partitioned matrix $\left(\left(\mathbf{U}^T_{\underline{\mathbf{O}}_s}\mathbf{F}_{\underline{\mathbf{O}}_{\mathbf{s}}}\mathbf{U}^*_{\underline{\mathbf{O}}_s}\right)^{-1}\right)$, multiplication of partitioned matrices $\left(\mathbf{U}^*_{\underline{\mathbf{O}}_s}\left(\mathbf{U}^T_{\underline{\mathbf{O}}_s}\mathbf{F}_{\underline{\mathbf{O}}_{\mathbf{s}}}\mathbf{U}^*_{\underline{\mathbf{O}}_s}\right)^{-1}\mathbf{U}^T_{\underline{\mathbf{O}}_s}\right)$. Therefore we do not provide intermediate results as they are easily reproducible.

Last, derivation of analytical expressions by application of (37) is facilitated if $\underline{\mathbf{O}}_{\mathbf{s}}$ is rearranged as $\underline{\mathbf{O}}^T_{\mathbf{s}} = \left(\left(\underline{\boldsymbol{\sigma}}^1_{\mathbf{s}}\right)^T, \ldots, \left(\underline{\boldsymbol{\sigma}}^L_{\mathbf{s}}\right)^T, \left(\underline{\boldsymbol{\sigma}}^1_{\mathbf{s}}\right)^H, \ldots, \left(\underline{\boldsymbol{\sigma}}^L_{\mathbf{s}}\right)^H, \left(\boldsymbol{\Xi}^1_{\mathbf{s}}\right)^T, \ldots, \left(\boldsymbol{\Xi}^L_{\mathbf{s}}\right)^T\right)$.

The most general approach (that is valid for any type of signal source) consists in assuming that $\boldsymbol{\Xi}^l_{\mathbf{s}}$

and $\boldsymbol{\sigma}_{\mathbf{s}}^{l}$ may vary during the L observations; then estimation performance is given by [42][48]:

$$\begin{aligned}
\mathbf{CRB}_{\boldsymbol{\Xi}_{\mathbf{s}}^{l}|\underline{\mathbf{O}}}\left(\underline{\mathbf{O}}\right) &= \left(2\,\mathrm{Re}\left\{\mathbf{H}_{\boldsymbol{\Xi}_{\mathbf{s}}^{l}}^{l}\left(\underline{\boldsymbol{\theta}}^{l}\right)\odot\left(\left(\boldsymbol{\sigma}_{\mathbf{s}}^{l}\left(\boldsymbol{\sigma}_{\mathbf{s}}^{l}\right)^{H}\right)^{T}\otimes\mathbf{1}_{P\times P}\right)\right\}\right)^{-1} \qquad (38)\\
\mathbf{H}_{\boldsymbol{\Xi}_{\mathbf{s}}^{l}}^{l}\left(\underline{\boldsymbol{\theta}}^{l}\right) &= \begin{bmatrix}\mathbf{H}_{\boldsymbol{\Xi}_{\mathbf{s}}^{l}}^{l}\left(\underline{\boldsymbol{\theta}}^{l}\right)_{1,1} & \cdots & \mathbf{H}_{\boldsymbol{\Xi}_{\mathbf{s}}^{l}}^{l}\left(\underline{\boldsymbol{\theta}}^{l}\right)_{1,M}\\ \vdots & \ddots & \vdots\\ \mathbf{H}_{\boldsymbol{\Xi}_{\mathbf{s}}^{l}}^{l}\left(\underline{\boldsymbol{\theta}}^{l}\right)_{M,1} & \cdots & \mathbf{H}_{\boldsymbol{\Xi}_{\mathbf{s}}^{l}}^{l}\left(\underline{\boldsymbol{\theta}}^{l}\right)_{M,M}\end{bmatrix}\\
\mathbf{H}_{\boldsymbol{\Xi}_{\mathbf{s}}^{l}}^{l}\left(\underline{\boldsymbol{\theta}}^{l}\right)_{m_1,m_2} &= \mathbf{G}_{\underline{\boldsymbol{\theta}}_{\mathbf{n}}^{l}}\left(\Pi_{\left\{\mathbf{B}_{\mathbf{s}}^{l}\left(\boldsymbol{\Xi}_{\mathbf{s}}^{l}\right)\right\}}^{\perp}\left(\left\{\frac{\partial\mathbf{b}_{\mathbf{s}}^{l}\left(\boldsymbol{\varepsilon}_{m_2}^{l}\right)}{\partial\left(\boldsymbol{\varepsilon}^{l}\right)^{T}}\right\}\right),\Pi_{\left\{\mathbf{B}_{\mathbf{s}}^{l}\left(\boldsymbol{\Xi}_{\mathbf{s}}^{l}\right)\right\}}^{\perp}\left(\left\{\frac{\partial\mathbf{b}_{\mathbf{s}}^{l}\left(\boldsymbol{\varepsilon}_{m_1}^{l}\right)}{\partial\left(\boldsymbol{\varepsilon}^{l}\right)^{T}}\right\}\right)\right)\\
\left\langle\mathbf{x}\left(f\right)\mid\mathbf{y}\left(f\right)\right\rangle_{\underline{\boldsymbol{\theta}}_{\mathbf{n}}^{l}} &= \int_{-\frac{B^{l}}{2}}^{\frac{B^{l}}{2}}\mathbf{y}\left(f\right)^{H}\mathbf{R}_{\mathbf{n}}^{l}\left(f,\underline{\boldsymbol{\theta}}_{\mathbf{n}}^{l}\right)^{-1}\mathbf{x}\left(f\right)df.
\end{aligned}$$

The first set of constraints of interest are the $\left(L-1\right)\left(MP\right)$ constraints implicitly considered in the standard signal processing literature relative to the invariability of $\boldsymbol{\Xi}_{\mathbf{s}}^{l}$:

$$\boldsymbol{\Xi}_{\mathbf{s}}^{l}=\boldsymbol{\Xi}_{\mathbf{s}}^{1},\quad 2\leq l\leq L \qquad (39)$$

An easy way to get the associated $\mathbf{U}_{\underline{\mathbf{O}}_{s}}$ is to use (17):

$$\mathbf{U}_{\underline{\mathbf{O}}_{\mathbf{s}}}=\begin{bmatrix}\mathbf{I}_{(LM,LM)} & \mathbf{0} & \mathbf{0}\\ \mathbf{0} & \mathbf{I}_{(LM,LM)} & \mathbf{0}\\ \mathbf{0} & \mathbf{0} & \mathbf{1}_{(L,1)}\otimes\mathbf{I}_{(MP,MP)}\end{bmatrix}\in\mathcal{M}\left(LM\left(P+2\right),M\left(2L+P\right)\right) \qquad (40)$$

what leads to (after substitution of $\boldsymbol{\Xi}_{\mathbf{s}}$ for $\boldsymbol{\Xi}_{\mathbf{s}}^{1}$):

$$\begin{aligned}
\mathbf{CRB}_{\boldsymbol{\Xi}_{\mathbf{s}}|\underline{\mathbf{O}}}^{c}\left(\underline{\mathbf{O}}\right) &= \left(2\sum_{l=1}^{L}\mathrm{Re}\left\{\mathbf{H}_{\boldsymbol{\Xi}_{\mathbf{s}}}^{l}\left(\underline{\boldsymbol{\theta}}^{l}\right)\odot\left(\left(\boldsymbol{\sigma}_{\mathbf{s}}^{l}\left(\boldsymbol{\sigma}_{\mathbf{s}}^{l}\right)^{H}\right)^{T}\otimes\mathbf{1}_{P\times P}\right)\right\}\right)^{-1} \qquad (41)\\
\mathbf{H}_{\boldsymbol{\Xi}_{\mathbf{s}}}^{l}\left(\underline{\boldsymbol{\theta}}^{l}\right) &= \begin{bmatrix}\mathbf{H}_{\boldsymbol{\Xi}_{\mathbf{s}}}^{l}\left(\underline{\boldsymbol{\theta}}^{l}\right)_{1,1} & \cdots & \mathbf{H}_{\boldsymbol{\Xi}_{\mathbf{s}}}^{l}\left(\underline{\boldsymbol{\theta}}^{l}\right)_{1,M}\\ \vdots & \ddots & \vdots\\ \mathbf{H}_{\boldsymbol{\Xi}_{\mathbf{s}}}^{l}\left(\underline{\boldsymbol{\theta}}^{l}\right)_{M,1} & \cdots & \mathbf{H}_{\boldsymbol{\Xi}_{\mathbf{s}}}^{l}\left(\underline{\boldsymbol{\theta}}^{l}\right)_{M,M}\end{bmatrix}\\
\mathbf{H}_{\boldsymbol{\Xi}_{\mathbf{s}}}^{l}\left(\underline{\boldsymbol{\theta}}^{l}\right)_{m_1,m_2} &= \mathbf{G}_{\underline{\boldsymbol{\theta}}_{\mathbf{n}}^{l}}\left(\Pi_{\left\{\mathbf{B}_{\mathbf{s}}^{l}\left(\boldsymbol{\Xi}_{\mathbf{s}}\right)\right\}}^{\perp}\left(\left\{\frac{\partial\mathbf{b}_{\mathbf{s}}^{l}\left(\boldsymbol{\varepsilon}_{m_2}\right)}{\partial\boldsymbol{\varepsilon}^{T}}\right\}\right),\Pi_{\left\{\mathbf{B}_{\mathbf{s}}^{l}\left(\boldsymbol{\Xi}_{\mathbf{s}}\right)\right\}}^{\perp}\left(\left\{\frac{\partial\mathbf{b}_{\mathbf{s}}^{l}\left(\boldsymbol{\varepsilon}_{m_1}\right)}{\partial\boldsymbol{\varepsilon}^{T}}\right\}\right)\right),\\
\left\langle\mathbf{x}\left(f\right)\mid\mathbf{y}\left(f\right)\right\rangle_{\underline{\boldsymbol{\theta}}_{\mathbf{n}}^{l}} &= \int_{-\frac{B^{l}}{2}}^{\frac{B^{l}}{2}}\mathbf{y}\left(f\right)^{H}\mathbf{R}_{\mathbf{n}}^{l}\left(f,\underline{\boldsymbol{\theta}}_{\mathbf{n}}^{l}\right)^{-1}\mathbf{x}\left(f\right)df.
\end{aligned}$$

Note that (41) is a new (to the best of our knowledge) generalization of expressions available in the open literature [1][17][27] where $\mathbf{b}_{\mathbf{s}}^{l}\left(t;\boldsymbol{\varepsilon}\right)=\mathbf{b}_{\mathbf{s}}^{1}\left(t;\boldsymbol{\varepsilon}\right)=\mathbf{b}_{\mathbf{s}}\left(t;\boldsymbol{\varepsilon}\right)$, $\mathbf{R}_{\mathbf{n}}^{l}\left(f,\underline{\boldsymbol{\theta}}_{\mathbf{n}}^{l}\right)=\mathbf{R}_{\mathbf{n}}^{1}\left(f,\underline{\boldsymbol{\theta}}_{\mathbf{n}}^{1}\right)=$

$\mathbf{R}_{\mathbf{n}}\left(f,\underline{\boldsymbol{\theta}}_{\mathbf{n}}\right)$, $\mathbf{H}^{l}_{\underline{\boldsymbol{\Xi}}_{\mathbf{s}}}\left(\underline{\boldsymbol{\theta}}^{l}\right)=\mathbf{H}^{1}_{\underline{\boldsymbol{\Xi}}_{\mathbf{s}}}\left(\underline{\boldsymbol{\theta}}^{1}\right)=\mathbf{H}_{\underline{\boldsymbol{\Xi}}_{\mathbf{s}}}\left(\underline{\boldsymbol{\theta}}\right)$, and (41) reduces to:

$$\begin{aligned}
\mathbf{CRB}^{c}_{\underline{\boldsymbol{\Xi}}_{\mathbf{s}}|\underline{\mathbf{O}}}\left(\underline{\mathbf{O}}\right) &= \left(2L\,\mathrm{Re}\left\{\mathbf{H}_{\underline{\boldsymbol{\Xi}}_{\mathbf{s}}}\left(\underline{\boldsymbol{\theta}}\right)\odot\left(\mathbf{R}^{T}_{\boldsymbol{\sigma}_{\mathbf{s}}}\otimes\mathbf{1}_{P\times P}\right)\right\}\right)^{-1},\quad \mathbf{R}_{\boldsymbol{\sigma}_{\mathbf{s}}}=\frac{1}{L}\sum_{l=1}^{L}\boldsymbol{\sigma}^{l}_{\mathbf{s}}\left(\boldsymbol{\sigma}^{l}_{\mathbf{s}}\right)^{H},\\
\mathbf{H}_{\underline{\boldsymbol{\Xi}}_{\mathbf{s}}}\left(\underline{\boldsymbol{\theta}}\right)_{m_1,m_2} &= \mathbf{G}_{\underline{\boldsymbol{\theta}}_{\mathbf{n}}}\left(\Pi^{\perp}_{\{\mathbf{B}_{\mathbf{s}}(\boldsymbol{\Xi}_{\mathbf{s}})\}}\left(\left\{\frac{\partial\mathbf{b}_{\mathbf{s}}\left(\boldsymbol{\varepsilon}_{m_2}\right)}{\partial\boldsymbol{\varepsilon}^{T}}\right\}\right),\Pi^{\perp}_{\{\mathbf{B}_{\mathbf{s}}(\boldsymbol{\Xi}_{\mathbf{s}})\}}\left(\left\{\frac{\partial\mathbf{b}_{\mathbf{s}}\left(\boldsymbol{\varepsilon}_{m_1}\right)}{\partial\boldsymbol{\varepsilon}^{T}}\right\}\right)\right)\\
\left\langle\mathbf{x}\left(f\right)\mid\mathbf{y}\left(f\right)\right\rangle_{\underline{\boldsymbol{\theta}}_{\mathbf{n}}} &= \int_{-\frac{B}{2}}^{\frac{B}{2}}\mathbf{y}\left(f\right)^{H}\mathbf{R}_{\mathbf{n}}\left(f,\underline{\boldsymbol{\theta}}_{\mathbf{n}}\right)^{-1}\mathbf{x}\left(f\right)df.
\end{aligned}$$

A second set of additional $2LM$ constraints of interest can be relative to the invariability of $\boldsymbol{\sigma}^{l}_{\mathbf{s}}$:

$$\boldsymbol{\Xi}^{l}_{\mathbf{s}}=\boldsymbol{\Xi}^{1}_{\mathbf{s}},\quad \underline{\boldsymbol{\sigma}}^{l}_{\mathbf{s}}=\underline{\boldsymbol{\sigma}}^{1}_{\mathbf{s}},\quad 2\leq l\leq L \tag{42}$$

then $\mathbf{U}_{\underline{\mathbf{O}}_{s}}$ (40) becomes :

$$\mathbf{U}_{\underline{\mathbf{O}}_{\mathbf{s}}}=\begin{bmatrix}\mathbf{1}_{(L,1)}\otimes\mathbf{I}_{(M,M)} & \mathbf{0} & \mathbf{0}\\ \mathbf{0} & \mathbf{1}_{(L,1)}\otimes\mathbf{I}_{(M,M)} & \mathbf{0}\\ \mathbf{0} & \mathbf{0} & \mathbf{1}_{(L,1)}\otimes\mathbf{I}_{(MP,MP)}\end{bmatrix}\in\mathcal{M}\left(LM\left(P+2\right),M\left(2+P\right)\right) \tag{43}$$

what leads to (after substitution of $\boldsymbol{\Xi}_{\mathbf{s}}$ for $\boldsymbol{\Xi}^{1}_{\mathbf{s}}$ and $\boldsymbol{\sigma}_{\mathbf{s}}$ for $\boldsymbol{\sigma}^{1}_{\mathbf{s}}$):

$$\mathbf{CRB}^{c}_{\underline{\boldsymbol{\Xi}}_{\mathbf{s}}|\underline{\mathbf{O}}}\left(\underline{\mathbf{O}}\right)=\left(2\,\mathrm{Re}\left\{\left(\mathbf{C}\left(\underline{\mathbf{O}}\right)-\mathbf{D}^{H}\left(\underline{\mathbf{O}}\right)\mathbf{E}\left(\underline{\mathbf{O}}\right)^{-1}\mathbf{D}\left(\underline{\mathbf{O}}\right)\right)\odot\left(\left(\boldsymbol{\sigma}_{\mathbf{s}}\boldsymbol{\sigma}^{H}_{\mathbf{s}}\right)^{T}\otimes\mathbf{1}_{P\times P}\right)\right\}\right)^{-1} \tag{44}$$

$$\mathbf{C}\left(\underline{\mathbf{O}}\right)=\sum_{l=1}^{L}\mathbf{C}^{l}_{\underline{\boldsymbol{\Xi}}_{\mathbf{s}}}\left(\underline{\boldsymbol{\theta}}^{l}\right),\quad \mathbf{D}\left(\underline{\mathbf{O}}\right)=\sum_{l=1}^{L}\mathbf{D}^{l}_{\underline{\boldsymbol{\Xi}}_{\mathbf{s}}}\left(\underline{\boldsymbol{\theta}}^{l}\right),\quad \mathbf{E}\left(\underline{\mathbf{O}}\right)=\sum_{l=1}^{L}\mathbf{G}_{\underline{\boldsymbol{\theta}}^{l}_{\mathbf{n}}}\left(\left\{\mathbf{B}^{l}_{\mathbf{s}}\left(\boldsymbol{\Xi}_{\mathbf{s}}\right)\right\},\left\{\mathbf{B}^{l}_{\mathbf{s}}\left(\boldsymbol{\Xi}_{\mathbf{s}}\right)\right\}\right)$$

$$\mathbf{C}^{l}_{\underline{\boldsymbol{\Xi}}_{\mathbf{s}}}\left(\underline{\boldsymbol{\theta}}^{l}\right)=\begin{bmatrix}\mathbf{C}^{l}_{\underline{\boldsymbol{\Xi}}_{\mathbf{s}}}\left(\underline{\boldsymbol{\theta}}^{l}\right)_{1,1} & \cdots & \mathbf{C}^{l}_{\underline{\boldsymbol{\Xi}}_{\mathbf{s}}}\left(\underline{\boldsymbol{\theta}}^{l}\right)_{1,M}\\ \vdots & \ddots & \vdots\\ \mathbf{C}^{l}_{\underline{\boldsymbol{\Xi}}_{\mathbf{s}}}\left(\underline{\boldsymbol{\theta}}^{l}\right)_{M,1} & \cdots & \mathbf{C}^{l}_{\underline{\boldsymbol{\Xi}}_{\mathbf{s}}}\left(\underline{\boldsymbol{\theta}}^{l}\right)_{M,M}\end{bmatrix},\quad \mathbf{C}^{l}_{\underline{\boldsymbol{\Xi}}_{\mathbf{s}}}\left(\underline{\boldsymbol{\theta}}^{l}\right)_{m_1,m_2}=\mathbf{G}_{\underline{\boldsymbol{\theta}}^{l}_{\mathbf{n}}}\left(\left\{\frac{\partial\mathbf{b}^{l}_{\mathbf{s}}\left(\boldsymbol{\varepsilon}_{m_2}\right)}{\partial\boldsymbol{\varepsilon}^{T}}\right\},\left\{\frac{\partial\mathbf{b}^{l}_{\mathbf{s}}\left(\boldsymbol{\varepsilon}_{m_1}\right)}{\partial\boldsymbol{\varepsilon}^{T}}\right\}\right)$$

$$\mathbf{D}^{l}_{\underline{\boldsymbol{\Xi}}_{\mathbf{s}}}\left(\underline{\boldsymbol{\theta}}^{l}\right)=\begin{bmatrix}\mathbf{D}^{l}_{\underline{\boldsymbol{\Xi}}_{\mathbf{s}}}\left(\underline{\boldsymbol{\theta}}^{l}\right)_{1} & \cdots & \mathbf{D}^{l}_{\underline{\boldsymbol{\Xi}}_{\mathbf{s}}}\left(\underline{\boldsymbol{\theta}}^{l}\right)_{M}\end{bmatrix},\qquad \mathbf{D}^{l}_{\underline{\boldsymbol{\Xi}}_{\mathbf{s}}}\left(\underline{\boldsymbol{\theta}}^{l}\right)_{m}=\mathbf{G}_{\underline{\boldsymbol{\theta}}^{l}_{\mathbf{n}}}\left(\left\{\frac{\partial\mathbf{b}^{l}_{\mathbf{s}}\left(\boldsymbol{\varepsilon}_{m}\right)}{\partial\boldsymbol{\varepsilon}^{T}}\right\},\left\{\mathbf{B}^{l}_{\mathbf{s}}\left(\boldsymbol{\Xi}_{\mathbf{s}}\right)\right\}\right)$$

$$\left\langle\mathbf{x}\left(f\right)\mid\mathbf{y}\left(f\right)\right\rangle_{\underline{\boldsymbol{\theta}}^{l}_{\mathbf{n}}}=\int_{-\frac{B_l}{2}}^{\frac{B_l}{2}}\mathbf{y}\left(f\right)^{H}\mathbf{R}^{l}_{\mathbf{n}}\left(f,\underline{\boldsymbol{\theta}}^{l}_{\mathbf{n}}\right)^{-1}\mathbf{x}\left(f\right)df.$$

Note that (44) is another new (to the best of our knowledge) generalization of expressions available in the open literature [1][17][27] where $\mathbf{b}^{l}_{\mathbf{s}}\left(t;\boldsymbol{\varepsilon}\right)=\mathbf{b}^{1}_{\mathbf{s}}\left(t;\boldsymbol{\varepsilon}\right)=\mathbf{b}_{\mathbf{s}}\left(t;\boldsymbol{\varepsilon}\right)$, $\mathbf{R}^{l}_{\mathbf{n}}\left(f,\underline{\boldsymbol{\theta}}^{l}_{\mathbf{n}}\right)=\mathbf{R}^{1}_{\mathbf{n}}\left(f,\underline{\boldsymbol{\theta}}^{1}_{\mathbf{n}}\right)=\mathbf{R}_{\mathbf{n}}\left(f,\underline{\boldsymbol{\theta}}_{\mathbf{n}}\right)$, $\mathbf{C}^{l}_{\underline{\boldsymbol{\Xi}}_{\mathbf{s}}}\left(\underline{\boldsymbol{\theta}}^{l}\right)=\mathbf{C}^{1}_{\underline{\boldsymbol{\Xi}}_{\mathbf{s}}}\left(\underline{\boldsymbol{\theta}}^{1}\right)=\mathbf{C}_{\underline{\boldsymbol{\Xi}}_{\mathbf{s}}}\left(\underline{\mathbf{O}}\right)$, $\mathbf{D}^{l}_{\underline{\boldsymbol{\Xi}}_{\mathbf{s}}}\left(\underline{\boldsymbol{\theta}}^{l}\right)=\mathbf{D}^{1}_{\underline{\boldsymbol{\Xi}}_{\mathbf{s}}}\left(\underline{\boldsymbol{\theta}}^{1}\right)=\mathbf{D}_{\underline{\boldsymbol{\Xi}}_{\mathbf{s}}}\left(\underline{\mathbf{O}}\right)$, and (44) reduces

to:

$$
\begin{aligned}
\mathbf{CRB}^{c}_{\underline{\boldsymbol{\Xi}}_{\mathbf{s}}|\underline{\mathbf{O}}}\left(\underline{\mathbf{O}}\right) &= \left(2L\,\mathrm{Re}\left\{\mathbf{H}_{\underline{\boldsymbol{\Xi}}_{\mathbf{s}}}\left(\underline{\boldsymbol{\theta}}\right)\odot\left(\left(\boldsymbol{\sigma}_{\mathbf{s}}\boldsymbol{\sigma}_{\mathbf{s}}^{H}\right)^{T}\otimes\mathbf{1}_{P\times P}\right)\right\}\right)^{-1} \\
\mathbf{H}_{\underline{\boldsymbol{\Xi}}_{\mathbf{s}}}\left(\underline{\boldsymbol{\theta}}\right) &= \begin{bmatrix} \mathbf{H}_{\underline{\boldsymbol{\Xi}}_{\mathbf{s}}}\left(\underline{\boldsymbol{\theta}}\right)_{1,1} & \cdots & \mathbf{H}_{\underline{\boldsymbol{\Xi}}_{\mathbf{s}}}\left(\underline{\boldsymbol{\theta}}\right)_{1,M} \\ \vdots & \ddots & \vdots \\ \mathbf{H}_{\underline{\boldsymbol{\Xi}}_{\mathbf{s}}}\left(\underline{\boldsymbol{\theta}}\right)_{M,1} & \cdots & \mathbf{H}_{\underline{\boldsymbol{\Xi}}_{\mathbf{s}}}\left(\underline{\boldsymbol{\theta}}\right)_{M,M} \end{bmatrix} \\
\mathbf{H}_{\underline{\boldsymbol{\Xi}}_{\mathbf{s}}}\left(\underline{\boldsymbol{\theta}}\right)_{m_1,m_2} &= \mathbf{G}_{\underline{\boldsymbol{\theta}}_{\mathbf{n}}}\left(\Pi^{\perp}_{\{\mathbf{B}_{\mathbf{s}}(\boldsymbol{\Xi}_{\mathbf{s}})\}}\left(\left\{\frac{\partial\mathbf{b}_{\mathbf{s}}\left(\boldsymbol{\varepsilon}_{m_2}\right)}{\partial\boldsymbol{\varepsilon}^{T}}\right\}\right),\Pi^{\perp}_{\{\mathbf{B}_{\mathbf{s}}(\boldsymbol{\Xi}_{\mathbf{s}})\}}\left(\left\{\frac{\partial\mathbf{b}_{\mathbf{s}}\left(\boldsymbol{\varepsilon}_{m_1}\right)}{\partial\boldsymbol{\varepsilon}^{T}}\right\}\right)\right) \\
\left\langle\mathbf{x}\left(f\right)\mid\mathbf{y}\left(f\right)\right\rangle_{\underline{\boldsymbol{\theta}}_{\mathbf{n}}} &= \int_{-\frac{B}{2}}^{\frac{B}{2}}\mathbf{y}\left(f\right)^{H}\mathbf{R}_{\mathbf{n}}\left(f,\underline{\boldsymbol{\theta}}_{\mathbf{n}}\right)^{-1}\mathbf{x}\left(f\right)df.
\end{aligned}
$$

Many others combinations of constraints on $\boldsymbol{\Xi}_{\mathbf{s}}^{l}$ and $\boldsymbol{\sigma}_{\mathbf{s}}^{l}$ are possible and the use of the computation scheme as above may lead or not to closed-form expressions. Any reader may feel free to use the proposed framework to derive new closed-form expressions, but as far we are concerned, we will not put any additional effort in that perspective, since closed-form expressions are not at all a preliminary requirement to perform a relevant analysis of estimation performance with the proposed framework (see next paragraph).

3) Example of analysis and design of a system with constrained CRB:

For example, let us assume that $\boldsymbol{\Xi}_{\mathbf{s}}^{l}$, $2\leq l\leq L$, reduce to kinematic parameters of interest (delays, velocities, DOAs, ...) in the context of an active tracking radar (or tracking mode of a multifunction active radar). Then (38)(41)(44) are estimation performance according to targets manoeuvrability (fast or slow manoeuvring targets) and targets amplitude fluctuation (which can be caused either by manoeuvres or by radar carrier frequency changes or by both): from an highly manoeuvring (non stationary) set of targets to a steady (stationary) set of targets. From an operational point of view, this parameters variability can be modelled in terms of 3 "canonical" scenarios $\{\mathcal{S}_1,\mathcal{S}_2,\mathcal{S}_3\}$, each scenario consisting of single target and a set of parameters associated with a number of realizations: $\mathcal{S}_1\triangleq\left\{\left\{\boldsymbol{\Xi}_1^l,\sigma_1^l\right\}_{l=1}^{L}\right\}_{i=1}^{I_1}$, $\mathcal{S}_2\triangleq\left\{\boldsymbol{\Xi}_2,\left\{\sigma_2^l\right\}_{l=1}^{L}\right\}_{i=1}^{I_2}$, $\mathcal{S}_3\triangleq\left\{\boldsymbol{\Xi}_3,\sigma_3\right\}_{i=1}^{I_3}$. And we are interested in average estimation performance, that is in $\boldsymbol{\Xi}_1=\frac{1}{L}\sum_{l=1}^{L}\boldsymbol{\Xi}_1^l$, $\boldsymbol{\Xi}_2$ and $\boldsymbol{\Xi}_3$.

Then (38)(41)(44) allows for assessment of estimation performance of a given radar system $\left\{\mathbf{b}_{\mathbf{s}}^{l}\left(t,\boldsymbol{\varepsilon}\right)\right\}_{l=1}^{L}$ in presence of a given noise environment $\left\{\mathbf{R}_{\mathbf{n}}^{l}\left(f,\underline{\boldsymbol{\theta}}_{\mathbf{n}}^{l}\right)\right\}_{l=1}^{L}$. The performance assessment can be based on average performance or on the worst performance over all the realizations of the 3 given scenarios. Contrariwise, (38)(41)(44) can be used to look for the optimal radar system $\left\{\mathbf{b}_{\mathbf{s}}^{l}\left(t,\boldsymbol{\varepsilon}\right)\right\}_{l=1}^{L}$ in presence of a given noise environment $\left\{\mathbf{R}_{\mathbf{n}}^{l}\left(f,\underline{\boldsymbol{\theta}}_{\mathbf{n}}^{l}\right)\right\}_{l=1}^{L}$ and the given 3 scenarios $\{\mathcal{S}_1,\mathcal{S}_2,\mathcal{S}_3\}$. Off course, at least in theory, this design principle can be extended to more stringent scenarios, like several targets on patrols leading to the taking into account of high resolution scenarios.

However, from a practical point of view, the complexity of expressions (38)(41)(44) in the general case prevents from any insight without resorting to oversimplified particular cases (one or two signals, one or two unknown parameters per signal, ...). Additionally, even in the canonical case of a single target per scenario, (38)(41)(44) are rather analytical expressions than closed-form expressions as soon as the transmitted base band signals are band-limited with a non constant amplitude or phase (linear or non-linear frequency modulation, ofdm, ...), since then (38)(41)(44) contain some defined integrals which must be computed by numerical integration. Actually this kind of optimization approach for tracking mode is generally reduced to waveform optimization (design) for a single target per scenario, once the radar hardware design has been set (which may also depends on requirement on detection performance of the surveillance mode for a multifunction radar).

The major drawback of the "analytical" approach leading to the derivation of analytical expressions like (38)(41)(44) is not to allow for assessment of estimation performance of scenarios combination (that is targets type combination) as : $\{\mathcal{S}_1 \cup \mathcal{S}_2\}$, $\{\mathcal{S}_1 \cup \mathcal{S}_3\}$, $\{\mathcal{S}_2 \cup \mathcal{S}_3\}$, $\{\mathcal{S}_1 \cup \mathcal{S}_2 \cup \mathcal{S}_3\}$. Indeed, each scenario combination requires a new FIM/CRB analytical computation. And this issue highlights the major advantage of the proposed framework based on exploitation of constraints, since any scenario combination amounts simply in constraints combination leading to a constrained CRB accessible via a simple algebraic computation (37):

$$\begin{aligned}
&\{\mathcal{S}_1 \cup \mathcal{S}_2\}: && \boldsymbol{\Xi}_2^l = \boldsymbol{\Xi}_2^l, \quad 2 \leq l \leq L \\
&\{\mathcal{S}_1 \cup \mathcal{S}_3\}: && \boldsymbol{\Xi}_3^l = \boldsymbol{\Xi}_3^l, \quad \sigma_3^l = \sigma_3^1, \quad 2 \leq l \leq L \\
&\{\mathcal{S}_2 \cup \mathcal{S}_3\}: && \boldsymbol{\Xi}_2^l = \boldsymbol{\Xi}_2^l, \quad \boldsymbol{\Xi}_3^l = \boldsymbol{\Xi}_3^l, \quad \sigma_3^l = \sigma_3^1, \quad 2 \leq l \leq L \\
&\{\mathcal{S}_1 \cup \mathcal{S}_2 \cup \mathcal{S}_3\}: && \boldsymbol{\Xi}_2^l = \boldsymbol{\Xi}_2^l, \quad \boldsymbol{\Xi}_3^l = \boldsymbol{\Xi}_3^l, \quad \sigma_3^l = \sigma_3^1, \quad 2 \leq l \leq L
\end{aligned}$$

C. Estimation performance and design of a system with constrained CRB

From a more general perspective, closed-form or analytical expressions of $\mathbf{CRB}^c_{\underline{\mathbf{O}}_s|\underline{\mathbf{O}}}(\underline{\mathbf{O}})$ like (38)(41)(44) do not exist (to the best of our knowledge) in the open literature when $\underline{\boldsymbol{\Delta}}$ is known for dependent multiple conditional models (even for the narrow-band formulation) or when $\underline{\boldsymbol{\Delta}}$ is unknown whatever the multiple conditional models are dependent or independent. Nevertheless, it is still possible to compute algebraically $\mathbf{CRB}_{\underline{\mathbf{O}}_s|\underline{\mathbf{O}}}(\underline{\mathbf{O}})$ whatever the scenario considered by:

- identifying the vector of unknown parameters $\underline{\mathbf{O}}$,
- building the associated unconstrained FIM $\mathbf{F}_{\underline{\mathbf{O}}}$ (24) from components (21a-21e),
- inverting the FIM $\mathbf{F}_{\underline{\mathbf{O}}}$.

Even if the above computation principle may seem simple from a theoretical point of view, in practice it is not that simple to build the correct FIM $\mathbf{F}_{\underline{\mathbf{O}}}$ (24) for L observations under different hypothesis of variability for different parameters (as illustrated in the previous paragraph with scenarios $\{\mathcal{S}_1, \mathcal{S}_2, \mathcal{S}_3\}$). Moreover, any change in the variability of any parameter will lead to a new vector of

unknown parameters $\underline{\mathbf{O}}$ and the building a new FIM $\mathbf{F}_{\underline{\mathbf{O}}}$.

The proposed framework is precisely an elegant and versatile solution for this practical issue. Indeed, with the proposed framework it suffices:

• to list once all the parameters $\underline{\boldsymbol{\theta}}^l$ involved in each parametric observation model (22) which may be unknown according to some scenarios,

• to build the unconstrained FIM (24) associated to the vector of all possible unknown parameters for the L observations $\underline{\mathbf{O}}^T = \left(\left(\underline{\boldsymbol{\theta}}^1\right)^T, \ldots, \left(\underline{\boldsymbol{\theta}}^1\right)^T\right)$,

then, for any given parameters scenario (in terms of parameter variability, parameter status known or unknown, ...), it suffices to express the associated set of equality constraints $\underline{\mathbf{f}}\left(\underline{\mathbf{O}}\right) = \mathbf{0}$ and to compute a basis of $\ker\left\{\underline{\mathbf{f}}\left(\underline{\mathbf{O}}\right)\right\}$ (17) $\mathbf{U}_{\underline{\boldsymbol{\theta}}^o}$ in order to get the looked-for constrained FIM $\mathbf{F}^c_{\underline{\mathbf{O}}}$ (15) and constrained CRB $\mathbf{CRB}^c_{\underline{\mathbf{O}}|\underline{\mathbf{O}}}\left(\underline{\mathbf{O}}\right)$ (14).

Note that both the vectors of all unknown parameters $\underline{\mathbf{O}}$ and its associated unconstrained FIM do not change whatever the parameters scenario. Thus it is a very safe and systematic framework to analyze the estimation performance of a system provided computation power (matrix multiplication and inversion) is available.

Therefore, if estimation performance is taken as a metric in order to optimize the design of the system, the proposed framework is a versatile and reliable "brute force" method to design a system compliant with some estimation performance requirements.

An additional illustrative example is easily derived from the previous paragraphs: any subset of constraints (26) may allow to explore the sensitivity of estimation performance to modeling errors configurations: sensitivity to a given modeling error occurring in all observation models $\left(\underline{\delta}^1_q, \ldots, \underline{\delta}^L_q\right)$, sensitivity to all the modeling errors occurring in a given observation model $\underline{\boldsymbol{\delta}}^l$, ... In practice, modeling errors management is handled via sub-assemblies calibration requirements which is a part of the sub-assemblies design specification.

V. On the versatility of constrained CRB for estimation performance analysis and design of a system

As exemplified in the previous section, most of closed-form or analytical expressions of CRB dedicated to a system of L observations are in the general case complicated enough (see (38)(41)(44)) that they do not provide much insight into estimation performance without resorting to oversimplified particular cases or to numerical computations. Therefore the classical approach of deriving CRB in terms of closed-form or analytical expression may be regarded as superfluous from the view point of the relevance of the information these expressions provide. Last, each CRB derivation is valid for a given vector of vector of unknown parameters $\underline{\mathbf{O}}$: any change in the variability of any parameter or

in its status (known or unknown) leads to a new vector of unknown parameters $\underline{\mathbf{O}}$ and the building a new FIM $\mathbf{F}_{\underline{\mathbf{O}}}$ which may not allow for derivation of closed-form or analytical expression of the looked-for CRB.

Contrarily, the propose framework offers a reliable "brute force" heuristic to assess estimation performance of systems in order to analyze or to optimize their design, where no derivation of dedicated closed-form or analytical expressions of CRB is needed. The simplicity, the generality and the versatility of the proposed framework are clearly its major values.

Simplicity: provided that the unconstrained FIM is computable (as it is the case for Gaussian observations involving deterministic parameters, whatever the model is conditional or unconditional, circular or non circular), any factor impacting the estimation performance of system that can be taken into account via equality constraints leads to a constrained CRB accessible via a simple algebraic computation.

Generality: it is the aim of this section to provide an extended (but not exhaustive) set of equality constraints (arbitrarily divided in different varieties) to highlight their potential use in estimation performance analysis and design of a system.

Versatility: any combination of equality constraints from different varieties is possible (indeed suggested), provided that the whole set of equality constraints is not redundant (see section III-B).

By way of illustration of the proposed framework under a more general scope, we still consider the setting of a system generating L observations of a signal corrupted by noise (18) where the signal is composed of M sources:

$$\mathbf{x}^l\left(\underline{\boldsymbol{\theta}}^l\right)=\mathbf{s}^l\left(\underline{\boldsymbol{\theta}}_{\mathbf{s}}^l,\underline{\boldsymbol{\delta}}_{\mathbf{s}}^l\right)+\mathbf{n}^l\left(\underline{\boldsymbol{\theta}}_{\mathbf{n}}^l,\underline{\boldsymbol{\delta}}_{\mathbf{n}}^l\right),\quad \mathbf{s}^l\left(\underline{\boldsymbol{\theta}}_{\mathbf{s}}^l,\underline{\boldsymbol{\delta}}_{\mathbf{s}}^l\right)=\mathbf{B}_{\mathbf{s}}^l\left(\underline{\underline{\boldsymbol{\Xi}}}_{\mathbf{s}}^l,\underline{\boldsymbol{\delta}}_{\mathbf{s}}^l\right)\boldsymbol{\sigma}_{\mathbf{s}}^l \tag{45}$$

$$\left(\underline{\boldsymbol{\theta}}^l\right)^T=\left(\left(\underline{\boldsymbol{\theta}}_{\mathbf{s}}^l\right)^T,\left(\underline{\boldsymbol{\delta}}_{\mathbf{s}}^l\right)^T,\left(\underline{\boldsymbol{\theta}}_{\mathbf{n}}^l\right)^T,\left(\underline{\boldsymbol{\delta}}_{\mathbf{n}}^l\right)^T\right),\quad \left(\underline{\boldsymbol{\theta}}_{\mathbf{s}}^l\right)^T=\left(\left(\underline{\boldsymbol{\sigma}}_{\mathbf{s}}^l\right)^T,\underline{\underline{\boldsymbol{\Xi}}}_{\mathbf{s}}^l\right),\quad \left(\underline{\underline{\boldsymbol{\Xi}}}_{\mathbf{s}}^l\right)^T=\left(\left(\underline{\boldsymbol{\varepsilon}}_1^l\right)^T,\ldots,\left(\underline{\boldsymbol{\varepsilon}}_M^l\right)^T\right)$$

where :

• $\mathbf{n}^l\left(\underline{\boldsymbol{\theta}}_{\mathbf{n}}^l,\underline{\boldsymbol{\delta}}_{\mathbf{n}}^l\right)$ is a noise vector depending both on a vector $\underline{\boldsymbol{\theta}}_{\mathbf{n}}^l$ of $P_{\mathbf{n}}^l$ parameters associated with sources of noise (thermal noise covariance matrix, interference DOAs, ...) and on a vector $\underline{\boldsymbol{\delta}}_{\mathbf{n}}^l$ of $Q_{\mathbf{n}}^l$ parameters independent from noise (typically parameters of the system physical components contributing to noise observations (thermal noise or interference)),

• $\mathbf{B}_{\mathbf{s}}^l\left(\underline{\underline{\boldsymbol{\Xi}}}_{\mathbf{s}}^l,\underline{\boldsymbol{\delta}}_{\mathbf{s}}^l\right)=\left[\mathbf{b}_{\mathbf{s}}^l\left(\underline{\boldsymbol{\varepsilon}}_1^l,\underline{\boldsymbol{\delta}}_{\mathbf{s}}^l\right),\ldots,\mathbf{b}_{\mathbf{s}}^l\left(\underline{\boldsymbol{\varepsilon}}_M^l,\underline{\boldsymbol{\delta}}_{\mathbf{s}}^l\right)\right]$ where $\mathbf{b}_{\mathbf{s}}^l\left(\underline{\boldsymbol{\varepsilon}}^l,\underline{\boldsymbol{\delta}}^l\right)$ is a vector of N^l parametric functions (spatial, temporal, ... transfer function) depending both on a vector $\underline{\boldsymbol{\varepsilon}}^l$ of $P_{\mathbf{s}}^l$ parameters associated with a given signal (DOAs, amplitudes, ...) and on a vector $\underline{\boldsymbol{\delta}}^l$ of $Q_{\mathbf{s}}^l$ parameters independent from signal (typically parameters of the system physical components contributing to signal observations (waveforms generator, sensors on transmit and on receive, ...)),

• $\underline{\underline{\boldsymbol{\Xi}}}_{\mathbf{s}}^l$ are the signal parameters of interest and $\underline{\boldsymbol{\sigma}}_{\mathbf{s}}^l$ are the signal nuisance parameters.

In the standard signal processing framework available in open literature [1][31][32] to model systems,

mostly two different instances of (45) based on Gaussian complex circular observation are considered: the conditional signal model and the unconditional signal model [33]. The discussed signal models are Gaussian and the parameter of interest dependency is given by parameters which are connected with the expectation value in the conditional case (see Section IV) and with the covariance matrix in the unconditional one. Therefore, in the following we will frequently refer to these observations models for illustration purposes since they are supposed to be known references.

A first conceptual partition of the P^l parameters can be defined as follows: the first subset consists of the vector $\underline{\boldsymbol{\theta}}_i^l \left(\left(\underline{\boldsymbol{\theta}}_i^l \right)^T \triangleq \left(\left(\underline{\boldsymbol{\Xi}}_{\mathbf{s}}^1 \right)^T, \ldots, \left(\underline{\boldsymbol{\Xi}}_{\mathbf{s}}^L \right)^T \right) \right)$ of parameters of interest, that is the signal parameters we want to estimate (directions of arrivals, propagation delays, ...) although the second subset consists of the vector $\underline{\boldsymbol{\theta}}_r^l$ of all remaining unknown parameters: nuisance parameters, noise parameters, system components parameters. It is the partition implicitly used in the previous sections.

A second conceptual partition of the P^l parameters can be defined as follows: the first subset consists of the vector $\underline{\boldsymbol{\theta}}_u^l$ of parameters which are always unknown whatever the experimental conditions (directions of arrivals, propagation delays, signal and noise amplitude or power,....) although the second subset consists of the vector $\underline{\boldsymbol{\theta}}_{u/k}^l$ of parameters which can be either known or unknown according to the experimental conditions (typically system components parameters such as sensors parameters (location, complex amplitude gain, ...), receivers parameters (delays, losses, ... introduced by electronics), ..., known or unknown according to the availability of calibration data).

Let $\mathbf{X}^T = \left(\left(\mathbf{x}^1 \right)^T, \ldots, \left(\mathbf{x}^L \right)^T \right)$ be the N-dimensional $\left(N = \sum_{l=1}^L N^l \right)$ random vector resulting from the L available observations and let $\underline{\boldsymbol{\Theta}}^T = \left(\left(\underline{\boldsymbol{\theta}}^1 \right)^T, \ldots, \left(\underline{\boldsymbol{\theta}}^L \right)^T \right)$ be the $P+Q$-dimensional $\left(P+Q = \sum_{l=1}^L \left(P^l + Q^l \right) \right)$ vector containing all the models parameters (whatever they are known or unknown during the experimentation).

We assume that an analytical expression of the p.d.f. $p\left(\mathbf{X}; \underline{\boldsymbol{\Theta}} \right)$ is available such that the (unconstrained) FIM (11):

$$\mathbf{F}_{\underline{\boldsymbol{\Theta}}} = \mathbf{G}_{\underline{\boldsymbol{\Theta}}} \left(\left\{ \frac{\partial \ln p\left(\mathbf{X}; \underline{\boldsymbol{\Theta}} \right)}{\partial \underline{\boldsymbol{\Theta}}}^* \right\} \right) = E_{\underline{\boldsymbol{\Theta}}} \left[\frac{\partial \ln p\left(\mathbf{X}; \underline{\boldsymbol{\Theta}} \right)}{\partial \underline{\boldsymbol{\Theta}}} \frac{\partial \ln p\left(\mathbf{X}; \underline{\boldsymbol{\Theta}} \right)}{\partial \underline{\boldsymbol{\Theta}}}^H \right] \tag{46}$$

can be computed analytically of numerically (see Section IV for the conditional model).

Additionally, let $\underline{\boldsymbol{\Theta}}_{\mathbf{s}}^T = \left(\left(\underline{\boldsymbol{\theta}}_{\mathbf{s}}^1 \right)^T, \ldots, \left(\underline{\boldsymbol{\theta}}_{\mathbf{s}}^L \right)^T \right)$, $\underline{\boldsymbol{\Theta}}_{\mathbf{n}}^T = \left(\left(\underline{\boldsymbol{\theta}}_{\mathbf{n}}^1 \right)^T, \ldots, \left(\underline{\boldsymbol{\theta}}_{\mathbf{n}}^L \right)^T \right)$, $\underline{\boldsymbol{\Delta}}_{\mathbf{s}}^T = \left(\left(\underline{\boldsymbol{\delta}}_{\mathbf{s}}^1 \right)^T, \ldots, \left(\underline{\boldsymbol{\delta}}_{\mathbf{s}}^L \right)^T \right)$, $\underline{\boldsymbol{\Delta}}_{\mathbf{n}}^T = \left(\left(\underline{\boldsymbol{\delta}}_{\mathbf{n}}^1 \right)^T, \ldots, \left(\underline{\boldsymbol{\delta}}_{\mathbf{n}}^L \right)^T \right)$, $\underline{\boldsymbol{\Theta}}_u^T = \left(\left(\underline{\boldsymbol{\theta}}_u^1 \right)^T, \ldots, \left(\underline{\boldsymbol{\theta}}_u^L \right)^T \right)$, $\underline{\boldsymbol{\Theta}}_{u/k}^T = \left(\left(\underline{\boldsymbol{\theta}}_{u/k}^1 \right)^T, \ldots, \left(\underline{\boldsymbol{\theta}}_{u/k}^L \right)^T \right)$, $\underline{\boldsymbol{\Theta}}_i^T = \left(\left(\underline{\boldsymbol{\theta}}_i^1 \right)^T, \ldots, \left(\underline{\boldsymbol{\theta}}_i^L \right)^T \right)$, $\underline{\boldsymbol{\Theta}}_r^T = \left(\left(\underline{\boldsymbol{\theta}}_r^1 \right)^T, \ldots, \left(\underline{\boldsymbol{\theta}}_r^L \right)^T \right)$. Then $\underline{\boldsymbol{\Theta}}^T = \left(\underline{\boldsymbol{\Theta}}_{\mathbf{s}}^T, \underline{\boldsymbol{\Theta}}_{\mathbf{n}}^T, \underline{\boldsymbol{\Delta}}_{\mathbf{s}}^T, \underline{\boldsymbol{\Delta}}_{\mathbf{n}}^T \right) = \left(\underline{\boldsymbol{\Theta}}_u^T, \underline{\boldsymbol{\Theta}}_{u/k}^T \right) = \left(\underline{\boldsymbol{\Theta}}_i^T, \underline{\boldsymbol{\Theta}}_r^T \right)$.

A. Influence of parameters state - known or unknown

In many parametric observation models (systems) some parameters can be either known or unknown according to the experimental conditions. When the value of a parameter is known during an

observation, then its value can be incorporated into any expression involving the parametric model, such as MLEs, lower bounds, etc... Otherwise it must be estimated.

Actually, this issue is inherent to parametric estimation since the relevance of the performance computed or estimated clearly depends on the accuracy of the parameterized observation model. By way of illustration, computer simulations have demonstrated for several decades that in certain cases, advanced estimation techniques as ML, MUSIC, and related algorithms have superior performance compared with conventional processing technique [1]. Nevertheless, in spite of potential advantages of the "high-resolution" techniques, their application to real systems has been rather limited so far. One of the main reasons given for this situation is the relatively high sensitivity of these methods to various system errors. In the deterministic parameters paradigm, these system uncertainties can be modelled conceptually either as pure deterministic values or as random values which p.d.f are parameterized by deterministic parameters. For instance, imprecise knowledge of the sensors location of an antenna system or of their gain and phase characteristics can seriously degrade the antenna system performance [1, 8.11][34][35]. A sensor coordinate x may be, at experimentation time, totally unknown: $x = x^0 \triangleq \theta_q$, or partially unknown: $x = x^0 + dx \triangleq \theta_q + dx\left(\theta_{q+1}, \ldots, \theta_{q+\Delta q}\right)$ where x_0 is an initially calibrated value and $dx\left(\theta_{q+1}, \ldots, \theta_{q+\Delta q}\right)$ is a random discrepancy which p.d.f. depends on parameters $\left(\theta_{q+1}, \ldots, \theta_{q+\Delta q}\right)$ [2]. Experimental systems attempt to eliminate or minimize these errors by careful calibration of the system. Such calibration is sometimes difficult or impossible to carry out in an operational system. Even when initial calibration is possible $\left(x \triangleq x^0\right)$, system parameters may change over time $\left(x \triangleq x^0 + dx\right)$ due to factors such as: gradual changes in the behavior of the sensor itself and of the electronic circuitry between the sensor and the output of the digitizer (due to thermal effects, aging of components, etc.), changes in the location of the sensors (e.g., an antenna array located on the vibrating wing of an aircraft or a hydrophone array towed behind a ship). Due to these changes it may be impossible to maintain array calibration to the required accuracy. A possible solution is self-calibration [36][37][38] where unknown signal parameters and uncertainties are estimated simultaneously. Then, there are generally many more unknowns than equations, and a cost function is typically optimized to solve this problem, which is highly nonlinear making traditional gradient based approaches unsuitable and requesting optimization algorithm prone to rapid and robust convergence. Therefore it is of first importance to be able to quantify the sensitivity of asymptotic performance of the ML and related algorithms to an inadequacy of the reception model [1, 8.11][34][35] in order to assess whether a calibration of the inadequacy (an initial calibration or a self-calibration step) is needed to obtain the requested estimation performance. The sensitivity to an uncertainty parameter θ_q can be assessed by measuring the effect of its state change from unknown to known (or vice versa) on the CRB of parameters of interest, which simply amounts to introduce

an equality constraint:

$$\theta_q = \theta_q^0 \;\Leftrightarrow\; \left(\underline{\boldsymbol{\Theta}}_{u/k}\right)_q = \left(\underline{\boldsymbol{\Theta}}_{u/k}\right)_q^0 \;\Leftrightarrow\; \left(\underline{\boldsymbol{\Theta}}_{u/k}\right)_q - \left(\underline{\boldsymbol{\Theta}}_{u/k}\right)_q^0 = 0 \tag{47}$$

If the model calibration problem is the most general application case of (47) (see Section IV-B), they are many other possible applications of constraint (47) according to the problem under consideration. A first example is the case where a (or several) subset of the components of $\underline{\boldsymbol{\Theta}}$ consists of a (or several) correlation matrix, as in the unconditional model [33]: then (47) can be used to assess the sensitivity of estimation performance to correlation or uncorrelation (all off-diagonal elements are known and nil) hypothesis. A second example is the sensitivity of estimation performance (e.g., bearing, symbol, and channel estimation as in [22]) in communications scenarios to the use of side information, i.e. the use of training data, which amounts to treat some unknown parameters (code symbols) as known deterministic constants.

B. Influence of parameters variability

In some application, a given parameter $\left(\underline{\boldsymbol{\theta}}^l\right)_p$ may be present in the L observation models. According to experimental conditions, its L unknown values $\left(\underline{\boldsymbol{\theta}}^l\right)_p, 1 \leq l \leq L$, may vary independently or remain unchanged. The effect of the parameter variability during the L observations on the CRB of parameters of interest can be assessed by introducing $(L-1)$ equality constraints:

$$\left(\underline{\boldsymbol{\theta}}^l\right)_p - \left(\underline{\boldsymbol{\theta}}^1\right)_p = 0,\ 2 \leq l \leq L \tag{48}$$

leading to the direct algebraic computation of a constrained CRB (14) from the expression of the (unconstrained) FIM (46). A first well known example is the particular case where the a priori p.d.f. of amplitude vectors $\boldsymbol{\sigma}_{\mathbf{s}}^l$ in (45) are unknown: then vectors $\boldsymbol{\sigma}_{\mathbf{s}}^l$ are treated as deterministic but unknown parameter vectors. In some applications of this particular case, the variability or invariability of $\boldsymbol{\sigma}_{\mathbf{s}}^l$ may be a priori stated according to consideration from physics, such as signal kinematics or signal physical properties (as in radar if $\boldsymbol{\sigma}_{\mathbf{s}}^l$ stands for complex backscattering amplitudes, see Section IV-B), or from signal transmission properties (as in telecoms if $\boldsymbol{\sigma}_{\mathbf{s}}^l$ stands for complex transmitted symbols). If we consider a tracking radar, then it is worth knowing the sensitivity of estimation performance of the system to these a priori hypotheses which represent different types of target scenarios $\{\mathcal{S}_1, \mathcal{S}_2, \mathcal{S}3\}$, as illustrated in Sections IV-B2 and IV-B3 when $\left\{\boldsymbol{\Xi}_{\mathbf{s}}^l\right\}_{l=1}^L$ reduce to kinematic parameters (delays, DOAs,...). This analysis can be further refined if we consider that target scenarios have different probability of occurrence: the targets kinematics parameters $\left\{\boldsymbol{\Xi}_{\mathbf{s}}^l\right\}_{l=1}^L$ are expected to be invariable at experimentation time (scenarios $\{\mathcal{S}_2, \mathcal{S}_3\}$) but may be slightly variable occasionally (scenario $\{\mathcal{S}_1\}$). In that case, one would probably want to know, at least from a computational coast point of view, if it worth taking into account the possible but unlikely variability of kinematics parameters (low

probability of occurrence of scenario $\{\mathcal{S}_1\}$). One possible criterion for decision is the loss of performance (see reparameterization inequality (16)) when $\mathbf{\Xi}_{\mathbf{s}}^l$ is actually invariable at experimentation time although it has been modelled as variable. The performance comparison can be done by comparing the (unconstrained) CRB of the mean $\mathbf{g}\left(\underline{\mathbf{\Theta}}\right) = \frac{1}{L}\sum_{l=1}^{L}\mathbf{\Xi}_{\mathbf{s}}^l$ obtained by application of (10) (with the unconstrained FIM given by (46)) with the (constrained) CRB of $\mathbf{\Xi}_{\mathbf{s}} \triangleq \mathbf{\Xi}_{\mathbf{s}}^1$ obtained by application of (14) with constraints $\mathbf{\Xi}_{\mathbf{s}}^l - \mathbf{\Xi}_{\mathbf{s}}^1 = \mathbf{0}$, $2 \leq l \leq L$. If the difference is negligible from an operational point of view and the additional computational coast of the associated ML (or related) estimator is affordable, then it is worth implementing estimators taking into account the parameters variability. In the other hand, mainly if the difference is important from an operational point of view, it is rather preferable to force the parameters invariability hypothesis and to endure a non minimum estimation error occasionally. This type of analysis is typically an example of how performance estimation may influence the design of system processing as well.

C. Influence of parameters constraints/reparameterization

As recalled in section III-B, the assumption that the parameters satisfy functional constraints can be approached from the alternative, yet equivalent, perspective of a global injective reparameterization [16] or at least from a perspective where the parameters locally fit a reduced parametric model [11][21]. Whatever the approach considered (constraints or reparameterization), the parameters are fitted to a lower-dimensional parametric model , i.e., the parameters are assumed to be functions of a distinct reduced parameter vector. Then, the reparamaterization inequality (16) expresses analytically a quite intuitive estimation principle: when the total number of unknown parameters decreases in an observation model, the asymptotic quality of estimation increases (or remain unchanged), in the sense that the CRB decreases (or remain equal), whatever the function of the unknown parameters considered. A consequence of first importance is that the asymptotic quality of estimation increases for all parameters whatever the subset of parameters subject to constraints or reparameterization. Therefore, a first way to improve the estimation of a subset of unknown parameters (parameters of interest for example) can be to introduce, by design choices, either a parameterization change or equality constraints among the other remaining parameters.

As a first illustration, let us consider the case where (45) models multiple observations of an active system, such as a wave transmitter device (radar, sonar, communication device, ...) where the transmitted signal model (base band signal, carrier frequency) is under control, what may allow to introduce some constraints (or a reparameterization) on the amplitudes $\boldsymbol{\sigma}_{\mathbf{s}}^l, 1 \leq l \leq L$. For example in active radar, it is well known that under some conditions on radar target (scatterers) kinematics and waveform design [39], it is possible to introduce a Doppler effect from observation to observation

which can be modelled as:

$$\left(\boldsymbol{\sigma}_{\mathbf{s}}^{l}\right)_{m} = \left(\boldsymbol{\sigma}_{\mathbf{s}}^{1}\right)_{m} e^{j2\pi\omega_{m}(t_{l}-t_{1})}, \quad 1 \leq m \leq M, \quad 2 \leq l \leq L \tag{49}$$

where $t_1, \ldots, t_L$ are the observation times, $\boldsymbol{\omega}^T = (\omega_1, \ldots, \omega_M)$ is the vector of unknown Doppler frequencies and $\boldsymbol{\sigma}_{\mathbf{s}}^1$ is the vector of the unknown targets amplitude (backscattering complex coefficient), constant during the L observations. Therefore, the introduction of a Doppler effect is equivalent to an injective reparameterization $\boldsymbol{\sigma}_{\mathbf{s}}^l = \boldsymbol{\sigma}_{\mathbf{s}}^l\left(\boldsymbol{\sigma}_{\mathbf{s}}^1, \boldsymbol{\omega}\right)$ (49), or to the introduction of the following $(L-1)\times M$ equality constraints:

$$\left(\boldsymbol{\sigma}_{\mathbf{s}}^{l}\right)_{m} - \left(\boldsymbol{\sigma}_{\mathbf{s}}^{1}\right)_{m} e^{j2\pi\omega_{m}(t_{l}-t_{1})} = 0, 1 \leq m \leq M, 2 \leq l \leq L.$$

Then, as detailed in [40], for an active radar system consisting of a 1-element antenna array receiving scaled, timedelayed, and Doppler-shifted echoes of a known complex bandpass signal, the reparameterization inequality leads to the following design principles for delay estimation: if your main requirement is performance estimation, then the Doppler information must always be taken into account when you estimate delay, at the expense of a more complex ML algorithm.

A second illustration is given by the case where a (or several) subset of the components of $\underline{\boldsymbol{\Theta}}$ consists of a (or several) correlation matrix, as in the unconditional model where all the unknown parameters $\left(\underline{\boldsymbol{\Theta}}_{\mathbf{s}}^T, \underline{\boldsymbol{\Theta}}_{\mathbf{n}}^T, \underline{\boldsymbol{\Delta}}_{\mathbf{s}}^T, \underline{\boldsymbol{\Delta}}_{\mathbf{n}}^T\right)$ are connected with the covariance matrix. Then, the array geometry can be used to constrain some of these correlation matrices or some of their components (for instance, the noise correlation matrix function of $\underline{\boldsymbol{\Theta}}_{\mathbf{n}}^T$) to belong to particular matrix subsets such as Toeplitz matrices, persymmetric matrices, centrohermittian matrices .... [1], which amount to introduce constraints on some of the unknown parameters not of primary interest. Although this approach is not mentioned in reference books [1][31][32], it is therefore possible to assess the sensitivity of estimation performance to the type of array geometry with the proposed framework (see also [12] for another constrained matrix estimation example).

A second way to improve the estimation of a subset of unknown parameters can be to introduce, by design choices, either a parameterization change or equality constraints among them.

A first well known example is given by target tracking where the track before detect (TBD) techniques [41, §17] amount to introduce a reparameterization of the targets kinematics parameters $\left(\boldsymbol{\Xi}_{\mathbf{s}}^l, 1 \leq l \leq L\right)$ measurable by a radar system (radial range, radial velocity, radial acceleration, direction of arrival, ...) at each radar observation. In these techniques, the set of parameters $\left(\boldsymbol{\Xi}_{\mathbf{s}}^l, 1 \leq l \leq L\right)$ is reparameterized as a function of a reduced set of parameters $\boldsymbol{\Psi}_{\mathbf{s}}$ $\left(\boldsymbol{\Xi}_{\mathbf{s}}^l = \boldsymbol{\Xi}_{\mathbf{s}}^l\left(\boldsymbol{\Psi}_{\mathbf{s}}\right)\right)$ parameterizing a kinematic model valid during the L observations (for example, an uniformly accelerated motion in the Cartesian coordinate frame where both the radar and the sources are located, for which the unknown parameters consist of 3 vectors (acceleration, initial velocity, initial location) whatever the

number of radar observations L).

A second well known example is given by communications applications based on phase-modulated signals sources with constant modulus (CM) [11][22], which is a constraint on each unknown symbol to be estimated. In [22], it is shown that the CM property combined with training signals improve the accuracy in estimating the channel and the signal waveform.

D. Influence of parameters constraints on the identifiability

Let us recall that an unconstrained singular FIM is always singular on a manifold $\mathcal{C}$ of the parameter space (subset of the parameter space deriving from a set of equality constraints) [16]. The singularity of the FIM over $\mathcal{C}$ is equivalent to the unidentifiability of $\underline{\boldsymbol{\Theta}}$ over $\mathcal{C}$, what leads to the inexistence of the CRB for unbiased estimates over $\mathcal{C}$ resulting in an unbounded CRB at the vicinity of any $\underline{\boldsymbol{\Theta}} \in \mathcal{C}$ (see also [3] for identifiability in the presence of random nuisance parameters). Theoretically [9][16][24], it is possible to regularize an unidentifiable observation model (system), in the sense that it is possible to find the greatest subset - with respect to inclusion - of the parameter space Θ where an unbiased estimate exist for any function of the unknown parameters. It appears that this greatest subset is a manifold as well but which definition depends on the analytical expression for eigenvectors of the FIM, which does not exists in most cases. Thus so far, regularization constraints have rather been investigated under mainly two perspectives:

• the first one consists in noticing in the observation models $\mathbf{x}^l\left(\underline{\boldsymbol{\theta}}^l\right)$ the obvious ambiguity relations between some parameters [22][24] and to introduce some straightforward appropriate constraints that prevent the ambiguity relations from happening,

• the second one consists in the derivation of additional sufficient and/or necessary conditions [10][11] required for local identifiability for particular observation models for which the unconstrained FIM has an analytical expression allowing a rank analysis. For example, results from [10] provide sufficient conditions in MIMO systems for reducing the rank deficiency of the FIM to be as small as possible. Once this is achieved, the remaining degrees of uncertainty may be resolved with the addition of constraints. However, the value of the constrained CRB depends on the choice of the constraint function. Therefore different constraint functions lead to different values of the CRB and there is no rationale yet allowing to predict the efficiency of regularization constraints, in the sense of their ability not only to regularize the unconstrained FIM but also to provide a constrained CRB (14) with variances as low as possible. And this is a problem of practical importance since, if it is possible to know if operational target values for MSEs are lower bounded by the minimum attainable variances given by the Moore-Penrose generalized inverse of the singular unconstrained FIM [9][16][24]), it is not possible to ”guess” which constraints (and the associated constrained MLE [13]) will provide a constrained CRB (14) reaching the target values.

The proposed framework allows to introduce a third heuristic perspective: with the help of the appropriate computation power, an efficient regularization of an unidentifiable observation model may simply result from the computation of all the constrained CRB (14) associated with a library of admissible sets of constraints (compatible with operational conditions), the most efficient set being kept.

VI. Application Example

We consider the tracking mode of a multifunction ground based SIMO (Single Input-Multiple OutPut) radar which operates both surveillance and tracking mode. The radar receiving antenna consists of a uniform planar array (UPA) of N_{R_x} isotropic sensors and the transmitting antenna consists of a single isotropic antenna located at the radar phase-centre, which is also the origin of the coordinate frame. Each observation consists of the reception of a transmitted based band burst $e_T^l(t)$ of N_{pri}^l pulses:

$$e_T^l(t) = \sum_{i=1}^{N_{pri}^l} e_0^l\left(t - t_0^i\right), \quad \int_0^{T_0^l} \left|e_0^l(t)\right|^2 dt = T_0^l,$$

after backscattering by M point scatterers. During each observation, each point scatterer has an uniform radial motion $\mathbf{p}_m^l(t) = \mathbf{u}_m^l\left(r_m^l + v_m^l t\right)$ and a constant backscattering coefficient σ_m^l. Then under the usual assumptions of a single polarization on transmit, a quasi-monostatic radar configuration, a narrow-band receiving array and a spatially and temporally white noise (independent observations), the observations after Hilbert filtering can be modelled as [48]:

$$\mathbf{x}^l\left(t, \underline{\boldsymbol{\theta}}^l\right) = \sum_{m=1}^{M} \mathbf{a}\left(\mathbf{u}_m^l, \lambda_c\right) e^{-j4\pi\left(\frac{r_m^l + v_m^l t}{\lambda_c}\right)} \frac{\sqrt{P_T^l} e_T\left(\alpha_m^l\left(t - \frac{2r_m^l}{c}\right)\right)}{\left(r_m^l + v_m^l t\right)^2} \frac{\lambda_c \sigma_m^l}{(4\pi)^{\frac{3}{2}}} + \mathbf{n}^l(t), \ \alpha_m^l = 1 - \frac{2v_m^l}{c}, \tag{50}$$

$$a_n(\mathbf{u}, \lambda_c) = e^{j\frac{2\pi}{\lambda_c}\mathbf{u}\cdot\mathbf{p}_{R_n}}, \ \mathbf{R}_{\mathbf{n}}^l(\tau) = \left(\sigma_{\mathbf{n}}^2\right)^l \mathbf{I}\delta(\tau), \ \mathbf{u}(\theta,\phi)^T = (\cos(\phi)\sin(\theta), \cos(\phi)\cos(\theta), \sin(\phi)),$$

where P_T^l is the transmitter (peak) power, $\mathbf{p}_{R_n}$ is the location of the n^{th} receiving sensor and λ_c is the radar wavelength. For ground based radar the coordinate frame is implicitly of ENU type (East-North-Up) and the associated angles are azimuth θ and elevation ϕ as defined above. Therefore $\left(\boldsymbol{\Xi}_{\mathbf{s}}^l\right)^T = \left(\left(\boldsymbol{\varepsilon}_1^l\right)^T, \ldots, \left(\boldsymbol{\varepsilon}_M^l\right)^T\right)$ where $\boldsymbol{\varepsilon}_m^l = \left(r_m^l, v_m^l, \theta_m^l, \phi_m^l\right)$.

We are primarily interested by radial range and velocity, since most of multifunction radar are not able to perform an high resolution analysis in the angle domain. Therefore most of multifunction radar simply steers a beam in the direction designated by the surveillance mode, which amounts to consider that the angles of each target is known and equal to the steered beam angles:

$$\left\{\left(\theta_m^l, \phi_m^l\right) = \left(\theta_{steer}^l, \phi_{steer}^l\right)\right\}_{m=1}^{M}, \quad 1 \le l \le L, \tag{51}$$

and (50) becomes:

$$\mathbf{x}^l\left(t,\boldsymbol{\theta}^l\right)=\sum_{m=1}^{M}e^{-j4\pi\left(\frac{r_m^l+v_m^l t}{\lambda_c}\right)}\frac{\sqrt{N_{R_x}P_T^l}e_T\left(\alpha_m^l\left(t-\frac{2r_m^l}{c}\right)\right)}{\left(r_m^l+v_m^l t\right)^2}\frac{\lambda_c\sigma_m^l}{(4\pi)^{\frac{3}{2}}}+\mathbf{n}^l\left(t\right),\quad \underline{\boldsymbol{\varepsilon}}_m^l=\left(r_m^l,v_m^l\right)$$

The scenario we want to investigate from a performance analysis and system design is the following operational one: a patrol of two fighters is incoming when one fighter fires a missile. At that moment, the radar scene consists of 3 types of target $\{\mathcal{S}_1,\mathcal{S}_2,\mathcal{S}_3\}$: the first target (the missile) accelerates and its backscattering amplitude fluctuates, the second target (the fighter firing the missile) keeps an uniform motion and its backscattering amplitude fluctuates, the third target (the second fighter) keeps an uniform motion and its backscattering amplitude do not fluctuate. The behaviour of each target contains an a priori information that can be formulated in terms of parameters constraints:

$$\left(r_2^l,v_2^l\right)=\left(r_2^1,v_2^1\right),\quad \left(r_3^l,v_3^l\right)=\left(r_3^1,v_3^1\right),\quad \underline{\sigma}_3^l=\underline{\sigma}_3^1,\quad 2\leq l\leq L \tag{52}$$

which is clearly an "asymmetric" set of constraints in opposition with "symmetric" sets of constraints such as (39) or (42):

$$\left\{\left(r_m^l,v_m^l\right)=\left(r_m^1,v_m^1\right)\right\}_{m=1}^{M},\quad \underline{\boldsymbol{\sigma}}_{\mathbf{s}}^l=\underline{\boldsymbol{\sigma}}_{\mathbf{s}}^1,\quad 2\leq l\leq L.$$

We want to asses the benefit of taking into account the additional "symmetric" constraints (39)(42) or "asymmetric" constraints (52) when estimating the average value of parameters of interest $r_m=\frac{1}{L}\sum_{l=1}^{L}r_m^l$ and $v_m=\frac{1}{L}\sum_{l=1}^{L}v_m^l$.

A. Illustration of the versatility of constrained CRB

In this section we show how the constrained approach allows easily to take into account both design constraint (51) and target features (39)(42)(52). For sake of legibility, we assume that $L=3$ and $M=3$ (three bursts sent and three targets). Let consider the vector of all the targets parameters:

$$\begin{aligned}\underline{\mathbf{O}}_{\mathbf{s}}^T=\;&\left(\sigma_1^1,\sigma_1^2,\sigma_1^3,\ldots,\sigma_3^1,\sigma_3^2,\sigma_3^3,\left(\sigma_1^1\right)^*,\left(\sigma_1^2\right)^*,\left(\sigma_1^3\right)^*\ldots,\left(\sigma_3^1\right)^*,\left(\sigma_3^2\right)^*,\left(\sigma_3^3\right)^*,\right.\\&\left.r_1^1,v_1^1,\theta_1^1,\phi_1^1,\ldots,r_1^3,v_1^3,\theta_1^3,\phi_1^3,\ldots,r_3^1,v_3^1,\theta_3^1,\phi_3^1,\ldots,r_3^3,v_3^3,\theta_3^3,\phi_3^3\right)\end{aligned} \tag{53}$$

and the associated FIM $\mathbf{F}_{\underline{\mathbf{O}}_{\mathbf{s}}}$ for L independent conditional models (31)(36). As mentioned above, a beam steering architecture of the radar can be taken into account with constraints (51). The matrix $\mathbf{U}_{\underline{\mathbf{O}}_{\mathbf{s}}}^0$ (17) associated to constraint (51) can be easily computed when parameters are gathered in two sets: known and unknown parameters, which amount to a permutation $\mathbf{P}_0$:

$$\begin{aligned}\left(\mathbf{P}_0\underline{\mathbf{O}}_{\mathbf{s}}\right)^T=\;&\left(\sigma_1^1,\sigma_1^2,\sigma_1^3,\ldots,\sigma_3^1,\sigma_3^2,\sigma_3^3,\left(\sigma_1^1\right)^*,\left(\sigma_1^2\right)^*,\left(\sigma_1^3\right)^*\ldots,\left(\sigma_3^1\right)^*,\left(\sigma_3^2\right)^*,\left(\sigma_3^3\right)^*,\right.\\&\left.r_1^1,v_1^1,\ldots,r_1^3,v_1^3,\ldots,r_3^1,v_3^1,\ldots,r_3^3,v_3^3,\theta_1^1,\phi_1^1,\ldots,\theta_1^3,\phi_1^3,\theta_3^1,\phi_3^1,\ldots\theta_3^3,\phi_3^3\right)\end{aligned}$$

Then [16, Section V.A]:

$$\mathbf{U}_{\mathbf{P}_0\underline{\mathbf{O}}_{\mathbf{s}}}^0=\begin{bmatrix}\mathbf{I}_{36\times 36}\\ \mathbf{0}_{18\times 36}\end{bmatrix}\in\mathcal{M}_{\mathbb{R}}\left(54,36\right)$$

As:

$$\mathbf{F}^{c}_{\underline{\mathbf{O}}_{\mathbf{s}}} = \mathbf{U}^{T}_{\underline{\mathbf{O}}_{\mathbf{s}}} \mathbf{F}_{\underline{\mathbf{O}}_{\mathbf{s}}} \mathbf{U}^{*}_{\underline{\mathbf{O}}_{\mathbf{s}}} = \left(\mathbf{P}_0 \mathbf{U}_{\underline{\mathbf{O}}_{\mathbf{s}}}\right)^{T} \left(\mathbf{P}_0 \mathbf{F}_{\underline{\mathbf{O}}_{\mathbf{s}}} \mathbf{P}_0^{T}\right) \left(\mathbf{P}_0 \mathbf{U}_{\underline{\mathbf{O}}_{\mathbf{s}}}\right)^{*} = \left(\mathbf{P}_0^{T} \mathbf{U}^{0}_{\mathbf{P}_0 \underline{\mathbf{O}}_{\mathbf{s}}}\right)^{T} \mathbf{F}_{\underline{\mathbf{O}}_{\mathbf{s}}} \left(\mathbf{P}_0^{T} \mathbf{U}^{0}_{\mathbf{P}_0 \underline{\mathbf{O}}_{\mathbf{s}}}\right)^{*}$$

therefore:

$$\mathbf{U}^{0}_{\underline{\mathbf{O}}_{\mathbf{s}}} = \mathbf{P}_0^{T} \mathbf{U}^{0}_{\mathbf{P}_0 \underline{\mathbf{O}}_{\mathbf{s}}} \tag{54}$$

When (51) is satisfied (beam steering mode), any additional constraint is expressed more concisely when expressed from the constrained FIM $\left(\mathbf{U}^{0}_{\underline{\mathbf{O}}_{\mathbf{s}}}\right)^{T} \mathbf{F}_{\underline{\mathbf{O}}_{\mathbf{s}}} \left(\mathbf{U}^{0}_{\underline{\mathbf{O}}_{\mathbf{s}}}\right)^{*}$. Indeed the constrained CRB taking into account (51) and the additional "symmetric" constraints (39) is:

$$\left\{\boldsymbol{\Xi}^{l}_{\mathbf{s}} = \boldsymbol{\Xi}^{1}_{\mathbf{s}}\right\}_{l=2}^{L} \rightarrow \mathbf{CRB}^{c}_{1}\left(\underline{\mathbf{O}}_{\mathbf{s}}\right) = \left(\mathbf{U}^{1}_{\underline{\mathbf{O}}_{\mathbf{s}}}\right)^{*} \left(\left(\mathbf{U}^{1}_{\underline{\mathbf{O}}_{\mathbf{s}}}\right)^{T} \left(\left(\mathbf{U}^{0}_{\underline{\mathbf{O}}_{\mathbf{s}}}\right)^{T} \mathbf{F}_{\underline{\mathbf{O}}_{\mathbf{s}}} \left(\mathbf{U}^{0}_{\underline{\mathbf{O}}_{\mathbf{s}}}\right)^{*}\right) \left(\mathbf{U}^{1}_{\underline{\mathbf{O}}_{\mathbf{s}}}\right)^{*}\right)^{-1} \left(\mathbf{U}^{1}_{\underline{\mathbf{O}}_{\mathbf{s}}}\right)^{T} \tag{55}$$

where:

$$\mathbf{U}^{1}_{\underline{\mathbf{O}}_{\mathbf{s}}} = \begin{bmatrix} \mathbf{I}_{9\times 9} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_{9\times 9} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{1}_{3\times 1} \otimes \mathbf{I}_{(2\times 2)} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{1}_{3\times 1} \otimes \mathbf{I}_{(2\times 2)} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{1}_{3\times 1} \otimes \mathbf{I}_{(2\times 2)} \end{bmatrix} \tag{56}$$

The constrained CRB taking into account (51) and the additional "symmetric" constraints (42) is:

$$\left.\begin{matrix} \left\{\boldsymbol{\Xi}^{l}_{\mathbf{s}} = \boldsymbol{\Xi}^{1}_{\mathbf{s}}\right\}_{l=2}^{L} \\ \left\{\boldsymbol{\sigma}^{l}_{\mathbf{s}} = \boldsymbol{\sigma}^{1}_{\mathbf{s}}\right\}_{l=2}^{L} \end{matrix}\right| \rightarrow \mathbf{CRB}^{c}_{2}\left(\underline{\mathbf{O}}_{\mathbf{s}}\right) = \left(\mathbf{U}^{2}_{\underline{\mathbf{O}}_{\mathbf{s}}}\right)^{*} \left(\left(\mathbf{U}^{2}_{\underline{\mathbf{O}}_{\mathbf{s}}}\right)^{T} \left(\left(\mathbf{U}^{0}_{\underline{\mathbf{O}}_{\mathbf{s}}}\right)^{T} \mathbf{F}_{\underline{\mathbf{O}}_{\mathbf{s}}} \left(\mathbf{U}^{0}_{\underline{\mathbf{O}}_{\mathbf{s}}}\right)^{*}\right) \left(\mathbf{U}^{2}_{\underline{\mathbf{O}}_{\mathbf{s}}}\right)^{*}\right)^{-1} \left(\mathbf{U}^{2}_{\underline{\mathbf{O}}_{\mathbf{s}}}\right)^{T}, \tag{57}$$

where:

$$\mathbf{U}^{2}_{\underline{\mathbf{O}}_{\mathbf{s}}} = \begin{bmatrix} \mathbf{1}_{3\times 1} \otimes \mathbf{I}_{(3\times 3)} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{1}_{3\times 1} \otimes \mathbf{I}_{(3\times 3)} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{1}_{3\times 1} \otimes \mathbf{I}_{(2\times 2)} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{1}_{3\times 1} \otimes \mathbf{I}_{(2\times 2)} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{1}_{3\times 1} \otimes \mathbf{I}_{(2\times 2)} \end{bmatrix} \tag{58}$$

Last, the constrained CRB taking into account (51) and the additional "asymmetric" constraints (52) is:

$$\left.\begin{matrix} \left\{\varepsilon^{l}_{2} = \varepsilon^{1}_{2}\right\}_{l=2}^{L} \\ \left\{\varepsilon^{l}_{3} = \varepsilon^{1}_{3}\right\}_{l=2}^{L} \\ \left\{\sigma^{l}_{3} = \sigma^{1}_{3}\right\}_{l=2}^{L} \end{matrix}\right| \rightarrow \mathbf{CRB}^{c}_{3}\left(\underline{\mathbf{O}}_{\mathbf{s}}\right) = \left(\mathbf{U}^{3}_{\underline{\mathbf{O}}_{\mathbf{s}}}\right)^{*} \left(\left(\mathbf{U}^{3}_{\underline{\mathbf{O}}_{\mathbf{s}}}\right)^{T} \left(\left(\mathbf{U}^{0}_{\underline{\mathbf{O}}_{\mathbf{s}}}\right)^{T} \mathbf{F}_{\underline{\mathbf{O}}_{\mathbf{s}}} \left(\mathbf{U}^{0}_{\underline{\mathbf{O}}_{\mathbf{s}}}\right)^{*}\right) \left(\mathbf{U}^{3}_{\underline{\mathbf{O}}_{\mathbf{s}}}\right)^{*}\right)^{-1} \left(\mathbf{U}^{3}_{\underline{\mathbf{O}}_{\mathbf{s}}}\right)^{T}, \tag{59}$$

where:

$$\mathbf{U}^3_{\underline{\mathbf{O}}_\mathbf{s}} = \begin{bmatrix} \mathbf{I}_{6\times 6} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{1}_{3\times 1} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{I}_{6\times 6} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{1}_{3\times 1} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{I}_{(6\times 6)} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{1}_{3\times 1} \otimes \mathbf{I}_{(2\times 2)} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{1}_{3\times 1} \otimes \mathbf{I}_{(2\times 2)} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \end{bmatrix} \tag{60}$$

Actually, $\mathbf{U}^1_{\underline{\mathbf{O}}_\mathbf{s}}$ (56), $\mathbf{U}^2_{\underline{\mathbf{O}}_\mathbf{s}}$ (58) and $\mathbf{U}^3_{\underline{\mathbf{O}}_\mathbf{s}}$ (60) are easy to program in any computing language whatever the constraint (51) is satisfied or not, whatever the number of targets M or the number of observations L. Therefore there is no additional value in expressing (56)(58) and (60) when (51) is not satisfied whatever M or L.

B. numerical results

For sake of simplicity, the 3 bursts waveforms are identical. The 3 bursts are transmitted at time $(0.01, 0.11, 0.21)$(s). The first target has a radial acceleration of $-50ms^{-2}$. The two cases of correlated and uncorrelated backscattering amplitudes are considered. The SNR at output of coherent matched filter is $30dB$ for all targets and all bursts. The main features of targets and bursts are gathered below:

Burst features		**Targets features**		
Modulation Type	LFM	$\left(r_1^1, r_1^2, r_1^3\right)$	$10^4 - (2, 22, 43)$	(m)
Range Ambiguity	20 Km	$\left(r_2^1, r_2^2, r_2^3\right)$	$10^4 + (0, 0, 0)$	(m)
Range Resolution	15 m	$\left(r_3^1, r_3^2, r_3^3\right)$	$10^4 + (8, 8, 8)$	(m)
Blind Range	1155 m	$\left(v_1^1, v_1^2, v_1^3\right)$	$-200 - (0.5, 5, 10)$	(m/s)
Velocity Ambiguity	600 m/s	$\left(v_2^1, v_2^2, v_2^3\right)$	$-200 + (0, 0, 0)$	(m/s)
Velocity Resolution	20 m/s	$\left(v_3^1, v_3^2, v_3^3\right)$	$-200 + (10, 10, 10)$	(m/s)
Beam Width	5°	$\left(\theta_m^1, \theta_m^2, \theta_m^3\right)$	$(0, 0, 0)$	(deg)
Wave Length	0.16 m	$\left(\phi_m^1, \phi_m^2, \phi_m^3\right)$	$(0, 0, 0)$	(deg)
		$\left(\sigma_1^1, \sigma_1^2, \sigma_1^3\right)_{Cor}$	$(1, 1, 1)$	(m)
		$\left(\sigma_2^1, \sigma_2^2, \sigma_2^3\right)_{Cor}$	$(1, 1, 1)$	(m)
		$\left(\sigma_3^1, \sigma_3^2, \sigma_3^3\right)_{Cor}$	$(1, 1, 1)$	(m)
		$\left(\sigma_1^1, \sigma_1^2, \sigma_1^3\right)_{UnCor}$	$\left(1, e^{j\frac{4\pi}{3}}, e^{j\frac{8\pi}{3}}\right)$	(m)
		$\left(\sigma_2^1, \sigma_2^2, \sigma_2^3\right)_{UnCor}$	$\left(1, e^{j\frac{2\pi}{3}}, e^{j\frac{4\pi}{3}}\right)$	(m)
		$\left(\sigma_3^1, \sigma_3^2, \sigma_3^3\right)_{UnCor}$	$(1, 1, 1)$	(m)

(61)

The computation of the unconstrained $\mathbf{F}_{\underline{\mathbf{O}}_\mathbf{s}}$ (31) for L independent conditional models (36) has been obtained with a software developed in [48] and cross-checked with results released in [27] for a single target.

As the influence of "symmetric" constraints (39)(42) on estimation performance is well known in the open literature, we only focus on the influence of "asymmetric" constraints (52).

First we address the beam steering mode. The CRB values are expressed in dB relatively to the waveform resolution:

Scenario	**Target 1**	**Target 2**	**Target 3**
$\mathbf{\Xi_s}$	$CRB_r = -38.9$	$CRB_r = -32.9$	$CRB_r = -37.1$
correlated amplitudes	$CRB_v = -40.2$	$CRB_v = -33.5$	$CRB_v = -37.2$
$\mathbf{\Xi_s}$ under (52)	$CRB_r = -42.6$	$CRB_r = -42.5$	$CRB_r = -42.1$
correlated amplitudes	$CRB_v = -43.2$	$CRB_v = -42.9$	$CRB_v = -42.6$
$\mathbf{\Xi_s}$	$CRB_r = -38.9$	$CRB_r = -33.1$	$CRB_r = -37.6$
uncorrelated amplitudes	$CRB_v = -40.3$	$CRB_v = -33.7$	$CRB_v = -37.6$
$\mathbf{\Xi_s}$ under (52)	$CRB_r = -42.7$	$CRB_r = -41.6$	$CRB_r = -41.5$
uncorrelated amplitudes	$CRB_v = -43.2$	$CRB_v = -41.5$	$CRB_v = -41.6$

(62)

Second we have checked numerically that the above results do not change when $\left(\theta_m^l, \phi_m^l\right)$ status change from known to unknown, which would be the configuration for a radar able to implement angular high resolution technique. This decorrelation between (range,velocity) and angles is not surprising since it is present in all $\mathbf{F}_{\underline{\boldsymbol{\theta}}_\mathbf{s}^l}$ (36) [27].

The main surprising results (to be confirmed with more scenarios) are the insensitivity of the gain on performance estimation to amplitude correlation and the averaging on performance estimation (same performance for all targets) when (52) is taken into account.

VII. Conclusion

The present paper introduces an original framework in order to assess and to analyze the estimation performance and the design of a system of measurement modelled as a set of L parametric observation models. Its main advantage is to take into account most (and possibly all) of factors impacting the estimation performance of the parameters of interest via equality constraints leading to direct numerical algebraic computations of constrained CRB from the expression of the unconstrained FIM. This framework offers a rational heuristic to assess and to analyze performance estimation where derivation of dedicated analytical expressions of CRB is superfluous. Anyhow, for complex systems, derivation of analytical expression of CRB is either impossible or inefficient. For application, we have provided the general form of the FIM for conditional models (often used to model active system of measurement such radar or sonar systems) which generally precludes the derivation of an analytical

expression of the CRB for the parameters of interest for a realistic scenario where interference occurs in the presence of sensors modeling errors. Last, we have shown that the proposed framework can also be used efficiently to generate new closed-form expressions of CRB, although this is not the main aim of this framework.

REFERENCES

[1] H.L. Van Trees, *Optimum Array Processing*, New-York, Wiley-Interscience, 2002

[2] S. Fortunati, A. Farina, F. Gini, M. S. Greco, A. Graziano, S. Giompapa, "Least Squares Algorithm and Cramér-Rao Type Lower Bounds for Relative sensor Registration Process", IEEE Trans. on SP, 59(3): 1075-1087, 2011

[3] S. Fortunati, F. Gini, M. S. Greco, A. Farina, A. Graziano, S. Giompapa, "On the Identifiability Problem in the presence of Random Nuisance Parameters", Signal Processing 92: 2545-2551, 2012

[4] E. Chaumette, J. Galy, A. Quinlan, P. Larzabal, "A New Barankin Bound Approximation for the Prediction of the Threshold Region Performance of Maximum-Likelihood Estimators", IEEE Trans. on SP, 56(11): 5319-5333, 2008

[5] K. Todros and J. Tabrikian, "General Classes of Performance Lower Bounds for Parameter Estimation-Part I: Non-Bayesian Bounds for Unbiased Estimators", IEEE Trans. on IT, 56(10): 5045-5063, 2010

[6] J. D. Gorman and A. O. Hero, "Lower bounds for parametric estimation with constraints", IEEE Trans. on IT, 36(6): 1285-1301, 1990

[7] T.L. Marzetta, "A simple derivation of the constrained multiple parameter Cramér-Rao bound", IEEE Trans. on SP, 41(6): 2247-2249, 1993

[8] P. Stoica and B. C. Ng, "On the Cramér-Rao bound under parametric constraints", IEEE SP Letters, 5(7): 177-179, 1998

[9] Z. Ben-Haim, Y. C. Eldar, "On the Constrained Cramér-Rao Bound With a Singular Fisher Information Matrix", IEEE SP Letters, 16(6): 453-456, 2009

[10] T. J. Moore and B. M. Sadler, "Sufficient conditions for regularity and strict identifiability in MIMO systems", IEEE Trans. on SP, 52(9): 2650-2655, 2004

[11] Y. Yao, G. B. Giannakis, "On Regularity and Identifiability of Blind Source Separation Under Constant-Modulus Constraints", IEEE Trans. on SP, 53(4):1272-1281, 2005

[12] A.K. Jagannatham, B.D. Rao, "Cramér-Rao Lower Bound for Constrained Complex Parameters", IEEE SP letters, 11(11): 875-878, 2004

[13] T.J. Moore, B.M. Sadler, R.J. Kozick, "Maximum-Likelihood Estimation, the CramÉr–Rao Bound, and the Method of Scoring With Parameter Constraints", IEEE Trans. on SP, 56(3): 895-908, 2008

[14] Z. Ben-Haim, Y. C. Eldar, "The Cramér-Rao Bound for Estimating a Sparse Parameter Vector", IEEE Trans. on SP, 58(6): 3384-3389, 2010

[15] G. Tang and A. Nehorai, "Constrained Cramér–Rao Bound on Robust Principal Component Analysis", IEEE Trans. on SP, 59(10): 5070-5076, 2011

[16] T. Menni, E. Chaumette, P. Larzabal and J. P. Barbot, "New results on Deterministic Cramér-Rao bounds for real and complex parameters", IEEE Trans. on SP, 60(3): 1032-1049, 2012

[17] S. F. Yau and Y. Bresler, "A Compact Cramér-Rao Bound Expression for Parametric Estimation of Superimposed Signals", IEEE Trans. on SP, 40(5): 1226-1230, 1992

[18] S. Zhiguang, Z. Jianxiong, H. Lei, L. Jicheng, "A New Derivation of Constrained Cramér-Rao Bound via Norm Minimization", IEEE Trans. on SP, 59(4): 1879-1882, 2011

[19] A. Van Den Bos, "A Cramér-Rao Lower Bound for Complex Parameters", IEEE Trans. on SP, 42(10): 2859, 1994

[20] P. Stoica and T. L. Marzetta, “Parameter estimation problems with singular information matrices”, IEEE Trans. on SP, 49(1): 87-90, 2001

[21] T.J. Moore, R.J. Kozick, B.M. Sadler, ”The Constrained Cramér–Rao Bound From the Perspective of Fitting a Model”, IEEE SP Letters, 14(8): 564-567, 2007

[22] B. M. Sadler, R. J. Kozick, and T. Moore, “Bounds on bearing and symbol estimation with side information” IEEE Trans. on SP, 49(4): 822-834, 2001

[23] M. Spivak, *Calculus on Manifolds*. Cambridge, MA: Perseus Books, 1965

[24] Y.H. Li and P.C. Yeh, ”An Interpretation of the Moore-Penrose Generalized Inverse of a Singular Fisher Information Matrix”, IEEE Trans. on SP, 60(10): 5532-5536, 2012

[25] A. O. Hero and J. Fessler, ”A recursive algorithm for computing Cramér-Rao-type bounds on estimator covariance”, IEEE Trans. on IT, 40(4): 1205-1210, 1994.

[26] P. Tune, ”Computing Constrained Cramer-Rao Bounds”, IEEE Trans. on SP, 60(10): 5543-5548, 2012

[27] A. Dogandzic and A. Nehorai, ”Cramér-Rao Bounds for Estimating Range, Velocity, and Direction with an Active Array”, IEEE Trans. on SP, 49(6): 1122-1137, 2001

[28] M. Greco; P. Stinco; F. Gini; A. Farina, ”Cramér-Rao Bounds and Selection of Bistatic Channels for Multistatic Radar Systems”, IEEE Trans. on AES, 47(4): 2934-2948, 2011

[29] M. Hazewinkel, *Encyclopaedia of Mathematics*, Springer-Verlag Berlin Heidelberg New York, 2002

[30] R.A. Horn, C.R. Johnson, *Matrix Analysis*. Cambridge University Press, 1999

[31] S.M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*. Englewood Cliffs, NJ: Prentice-Hall, 1993.

[32] L. Scharf, *Statistical Signal Processing*, Addison-Wiley, 1991

[33] P. Stoica and A. Nehorai, ”Performances study of conditional and unconditional direction of arrival estimation”, IEEE Trans. on SP, 38(10): 1783-1795, 1990.

[34] A. Ferréol, P. Larzabal, and M. Viberg, ”Performance Prediction of Maximum Likelihood Direction-of-Arrival Estimation in the Presence of Modeling Errors”, IEEE Trans. SP, 56(10): 4785 4793, 2008

[35] B. Friedlander, “Sensitivity analysis of the Maximum Likelihood direction finding algorithm”, IEEE Trans. on AES, 26(6): 953-968, 1990

[36] A. Weiss and B. Friedlander, “Array shape calibration using sources in unknown locations-a maximum likelihood approach”, IEEE Trans. on ASSP, 37(12): 1958–1966, 1989

[37] P. Chung and S. Wang, “Array self-calibration using SAGE algorithm”, IEEE SAM Workshop, 2008

[38] Z. Li, S. Chen , H. Leung, E. Bossé, ”Joint Data Association, Registration, and Fusion using EM-KF”, IEEE Trans. on AES, 46(2): 496-507, 2010

[39] N Levanon, E. Mozeson, *Radar Signals*, Wiley-Interscience 2004

[40] T. Menni, E. Chaumette and P. Larzabal, ”Reparameterization and Constraints for CRB: duality and a major inequality for system analysis and design in the asymptotic region”, ICASSP 2012, Kyoto Japan

[41] S. Blackman and R. Popoli, *Design and Analysis of Modern Tracking Systems,* Artech House, 1999

[42] T. Menni, E. Chaumette, P. Larzabal and J. P. Barbot, ”Crb for Active Radar”, EUSIPCO 2011, Barcelona Spain

[43] M.C. Wicks, E.L. Mokole, S.D. Blunt, R.S. Schneible, V.J. Amuso , *Principles of Waveform Diversity and Design*, SciTech Publishing 2010

[44] A. Renaux, P. Forster, E. Chaumette, P. Larzabal, ”On the High RSB CML Estimator Full Statistical Characterization”, IEEE Trans. on SP, vol 54(12): 4840-4843, 2006

[45] M. Wax, “The joint estimation of differential delay, Doppler, and phase”, IEEE Trans. on IT, 28(5): 817-820, 1982.

[46] B. Friedlander, ”On the Cramer-Rao Bound for Time Delay and Doppler Estimation”, IEEE Trans. on IT, 30(3): 575-580, 1984.

[47] P. Whittle, “The analysis of multiple stationary time series”, J. Royal Statist. Sot., 15: 125-139, 1953

[48] T. Menni, ”Deterministic Cramér-Rao Bounds for the analysis of the asymptotic estimation performance of an active radar”, PhD Thesis, Ecole Normale Supérieure de Cachan, France, 2012

[49] A. Dogandzic and A. Nehorai, ”Space-time fading channel estimation and symbol detection in unknown spatially correlated noise, IEEE Trans. on SP, 50(3): 457-474, 2002