REPARAMETERIZATION AND CONSTRAINTS FOR CRB: DUALITY AND A MAJOR INEQUALITY FOR SYSTEM ANALYSIS AND DESIGN IN THE ASYMPTOTIC REGION

(1)*Tarek Menni,* (1)(2)*Eric Chaumette,* (1)*Pascal Larzabal*

(1) SATIE, ENS Cachan, CNRS, UniverSud, 61 av President Wilson, F-94230 Cachan, France
(2) ONERA - DEMR/TSI, The French Aerospace Lab, Chemin de la Hunière, F-91120 Palaiseau, France

ABSTRACT

The CRB is a lower bound of great interest for system analysis and design in the asymptotic region (high SNR and/or large number of snapshots), as it is simple to calculate and it is usually possible to obtain closed form expressions. It is from this perspective that the paper highlights, by means of a classical radar estimation problem, two results useful for system analysis and design: a reparameterization inequality and the equivalence between reparameterization and equality constraints.

Index Terms**—** Parameter estimation, Cramer–Rao bounds

1. 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. 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, the fact that in many cases it can be achieved asymptotically (high SNR and/or large number of snapshots) by Maximum Likelihood Estimators (MLE) , and last but not least, its noticeable property of being the lowest bound on the MSE of unbiased estimators, since it derives from local unbiasedness at the vicinity of any selected value of the parameters. This initial characterization of locally unbiased estimators has been significantly generalized by Barankin work [2][3], who established the general form of the greatest lower bound on MSE (BB) for uniformly unbiased estimators, which is unfortunately incomputable analytically in general. Therefore, since then, numerous works detailed in [2][3] have been devoted to deriving computable approximations of the BB and have shown that the CRB and the BB can be regarded as key representatives of two general classes of bounds, respectively the Small-Error bounds and the Large-Error bounds. These works have also 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 Small-Error bounds. 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 an additional ambiguity region, also called the transition region. In this region, the MSE of MLEs usually deteriorates rapidly with respect to Small-Error bounds and exhibits a threshold behaviour corresponding to a ”performance breakdown”. Small-Error bound such as the CRB are not able to handle the threshold phenomena, whereas it is revealed by Large-Error bounds that can be used to predict the threshold value. Unfortunately, Large-Error bounds closed form expressions hardly ever exist and, even when they exist, each Large-Error bound requests the search of an optimum over a set of test points, leading to a computational cost prohibitive in most applications. Therefore, provided that one keeps in mind the CRB limitations, that is, to become an excessively optimistic lower bound when the observation conditions degrade (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.
It is from this perspective that the paper is devoted to highlight two technical results derived in [4] useful for system analysis and design in the asymptotic region: the general reparameterization inequality and the equivalence between parameterization change and equality constraints.
As an application example, we consider the very well-known problem of delay and velocity estimation with an active radar where a known waveform is transmitted. Typically, the received signals are modelled as scaled, delayed, and Doppler-shifted versions of the transmitted signal [5].

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

2. CRAMÉR-RAO BOUND FOR MIXED (REAL AND COMPLEX) PARAMETERS

The notational convention adopted is as follows: a, $\mathbf{a}$, $\mathbf{A}$ indicates respectively a scalar, a vector and a matrix quantity. The matrix/vector conjugate is indicated by a superscript * and the matrix/vector transpose conjugate is indicated by a superscript ^{H}. $\underline{\mathbf{x}}$ denotes:

$$\begin{cases} \underline{\mathbf{x}} = \mathbf{x} & \text{if } \mathbf{x} \in \mathbb{R}^Q \\ \underline{\mathbf{x}} = \begin{pmatrix} \mathbf{x} \\ \mathbf{x}^* \end{pmatrix} & \text{if } \mathbf{x} \in \mathbb{C}^Q \backslash \mathbb{R}^Q \\ \underline{\begin{pmatrix} \mathbf{x}_c \\ \mathbf{x}_r \end{pmatrix}} = \begin{pmatrix} \mathbf{x}_c \\ \mathbf{x}_c^* \\ \mathbf{x}_r \end{pmatrix} & \text{if } \mathbf{x}_c \in \mathbb{C}^Q \setminus \mathbb{R}^Q,\ \mathbf{x}_r \in \mathbb{R}^{Q'} \end{cases} . \quad (1)$$

$\mathcal{M}_{\mathbb{C}}(N, P)$ denotes the vector space of complex matrices with N rows and P columns. If $\boldsymbol{\theta} = (\theta_1, \theta_2, \ldots, \theta_P)^T$, then: $\frac{\partial}{\partial \boldsymbol{\theta}} = \left(\frac{\partial}{\partial \theta_1}, \frac{\partial}{\partial \theta_2}, \ldots, \frac{\partial}{\partial \theta_P}\right)^T$. $\odot$ denotes the Hadamard product. $\otimes$ denotes the Kronecker product. $1(\mathbf{x})$ denotes the constant real-valued function equal to 1.

Regarding the definition of Hermitian product $\langle \mid \rangle$, we adopt the convention used in books of mathematics [6] 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 $\mathbf{G}\left(\{\mathbf{u}\}_{[1,Q]}, \{\mathbf{c}\}_{[1,P]}\right)$ ($P \times Q$ complex matrix) associated to 2 families of vectors of $\mathbb{U}$, $\{\mathbf{u}\}_{[1,Q]} = \{\mathbf{u}_1, \ldots, \mathbf{u}_Q\}$ and $\{\mathbf{c}\}_{[1,P]} = \{\mathbf{c}_1, \ldots, \mathbf{c}_P\}$ as the one verifying [6]:

$$\left\langle \sum_{q=1}^{Q} x_q \mathbf{u}_q \mid \sum_{p=1}^{P} y_p \mathbf{c}_p \right\rangle = \mathbf{y}^H \mathbf{G}\left(\mathbf{u}_{[1,Q]}, \mathbf{c}_{[1,P]}\right) \mathbf{x} \tag{2}$$

where $\mathbf{x} = (x_1, \ldots, x_Q)^T, \mathbf{y} = (y_1, \ldots, y_P)^T$. 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 [1] 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 inequality introduced in the present paper is obtained by transposing inequality terms (matrices).

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. 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 dual form:

$$p(\mathbf{x}; \boldsymbol{\theta}), \ \boldsymbol{\theta} = \left(\mathrm{Re}\left\{\boldsymbol{\theta}_c^T\right\}, \mathrm{Im}\left\{\boldsymbol{\theta}_c^T\right\}, \boldsymbol{\theta}_r^T\right)^T \tag{3}$$

$$p(\mathbf{x}; \underline{\boldsymbol{\theta}}), \ \underline{\boldsymbol{\theta}} = \left(\boldsymbol{\theta}_c^T, (\boldsymbol{\theta}_c^*)^T, \boldsymbol{\theta}_r^T\right)^T \tag{4}$$

In the following we will only consider the form (4) since it includes (3) when $P_c = 0$. Let $\underline{\boldsymbol{\theta}}^0$ be a selected value of the parameter $\underline{\boldsymbol{\theta}}$, and $\widehat{\mathbf{g}(\underline{\boldsymbol{\theta}}^0)}(\mathbf{x})$ an estimator of $\mathbf{g}(\underline{\boldsymbol{\theta}}^0)$ where $\mathbf{g}(\underline{\boldsymbol{\theta}}) = (g_1(\underline{\boldsymbol{\theta}}), \ldots, g_{Q_c}(\underline{\boldsymbol{\theta}}), g_{Q_c+1}(\underline{\boldsymbol{\theta}}), \ldots, g_Q(\underline{\boldsymbol{\theta}}))^T$ is a vector of Q functions of $\underline{\boldsymbol{\theta}}$, the first Q_c ones being complex-valued functions, the last $Q_r = Q - Q_c$ being real-valued functions, where $Q_c \in [0, Q]$. Then, the statistical performance of any estimator of $\mathbf{g}(\underline{\boldsymbol{\theta}}^0)$ is fully characterized - including characterization of real and imaginary parts [4] - in the MSE sense, by the computation of:

$$MSE_{\underline{\boldsymbol{\theta}}^0}\left[\boldsymbol{\delta}^T \underline{\widehat{\mathbf{g}(\underline{\boldsymbol{\theta}}^0)}}(\mathbf{x})\right] = \int_{\Omega} \left|\boldsymbol{\delta}^T \left(\underline{\widehat{\mathbf{g}(\underline{\boldsymbol{\theta}}^0)}}(\mathbf{x}) - \underline{\mathbf{g}}(\underline{\boldsymbol{\theta}}^0)\right)\right|^2 p(\mathbf{x}; \underline{\boldsymbol{\theta}}^0)\, d\mathbf{x},$$

which is a norm deriving from an Hermitian product $\langle \mid \rangle_{\underline{\boldsymbol{\theta}}^0}$:

$$MSE_{\underline{\boldsymbol{\theta}}^0}\left[\boldsymbol{\delta}^T \underline{\widehat{\mathbf{g}(\underline{\boldsymbol{\theta}}^0)}}(\mathbf{x})\right] = \boldsymbol{\delta}^H \mathbf{G}_{\underline{\boldsymbol{\theta}}^0}\left(\left\{\underline{\widehat{\mathbf{g}(\underline{\boldsymbol{\theta}}^0)}}(\mathbf{x}) - \underline{\mathbf{g}}(\underline{\boldsymbol{\theta}}^0)\right\}\right) \boldsymbol{\delta} \tag{5}$$

$$\langle g(\mathbf{x}) \mid h(\mathbf{x}) \rangle_{\underline{\boldsymbol{\theta}}} = E_{\underline{\boldsymbol{\theta}}}\left[g(\mathbf{x})\, h^*(\mathbf{x})\right]$$

where:

$$\{\mathbf{h}(\mathbf{x})\} = \{h_1(\mathbf{x}), \ldots, h_Q(\mathbf{x})\} \tag{6}$$

denotes a family of vectors which elements are the vector components. The problem of finding a lower bound of $\mathbf{G}_{\underline{\boldsymbol{\theta}}^0}\left(\{\mathbf{u}\}_{[1,Q]}\right)$, $\{\mathbf{u}\}_{[1,Q]} = \left\{\underline{\widehat{\mathbf{g}(\underline{\boldsymbol{\theta}}^0)}}(\mathbf{x}) - \underline{\mathbf{g}}(\underline{\boldsymbol{\theta}}^0)\right\}$, in (5) for locally unbiased estimators, amounts to the minimization of $\mathbf{G}_{\underline{\boldsymbol{\theta}}^0}\left(\{\mathbf{u}\}_{[1,Q]}\right)$ (with respect to the Löwner ordering [6, §7.7]) under a set of linear constraints, which solution is a standard algebra result [4]:

$$\min\left\{\mathbf{G}_{\underline{\boldsymbol{\theta}}^0}\left(\{\mathbf{u}\}_{[1,Q]}\right)\right\} = \left(\frac{\partial \underline{\mathbf{g}}(\underline{\boldsymbol{\theta}}^0)^T}{\partial \underline{\boldsymbol{\theta}}}\right)^H \mathbf{F}_{\underline{\boldsymbol{\theta}}^0}^{-1} \frac{\partial \underline{\mathbf{g}}(\underline{\boldsymbol{\theta}}^0)^T}{\partial \underline{\boldsymbol{\theta}}} \tag{7}$$

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

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

where $\mathbf{CRB}_{\underline{\boldsymbol{\theta}}^0} = \mathbf{F}_{\underline{\boldsymbol{\theta}}^0}^{-1}$ and $\mathbf{F}_{\underline{\boldsymbol{\theta}}^0}$ is the Fisher Information Matrix (FIM).

2.1. Reparameterization and constraints: duality and a major inequality

Let us define the following notation convention:

$$\begin{aligned} \mathbf{CRB}_{\underline{\mathbf{g}}|\underline{\boldsymbol{\theta}}}(\underline{\boldsymbol{\theta}}^0) &= \left(\frac{\partial \underline{\mathbf{g}}(\underline{\boldsymbol{\theta}}^0)^T}{\partial \underline{\boldsymbol{\theta}}}\right)^H \mathbf{CRB}_{\underline{\boldsymbol{\theta}}|\underline{\boldsymbol{\theta}}}(\underline{\boldsymbol{\theta}}^0) \frac{\partial \underline{\mathbf{g}}(\underline{\boldsymbol{\theta}}^0)^T}{\partial \underline{\boldsymbol{\theta}}} \\ \mathbf{CRB}_{\underline{\boldsymbol{\theta}}|\underline{\boldsymbol{\theta}}}(\underline{\boldsymbol{\theta}}^0) &= \mathbf{F}_{\underline{\boldsymbol{\theta}}^0}^{-1} \end{aligned} \tag{8}$$

and consider the problem of estimating $\mathbf{h}(\underline{\boldsymbol{\omega}}) = \mathbf{g}(\underline{\boldsymbol{\theta}}(\underline{\boldsymbol{\omega}}))$, where $\underline{\boldsymbol{\theta}}(\underline{\boldsymbol{\omega}})$ is an injective reparameterization of the p.d.f. $p(\mathbf{x}; \underline{\boldsymbol{\theta}})$ for the unknown parameters $\underline{\boldsymbol{\theta}}$ ($\dim\{\underline{\boldsymbol{\theta}}\} = P$):

$$\underline{\boldsymbol{\theta}} = \underline{\boldsymbol{\theta}}(\underline{\boldsymbol{\omega}}), \ \dim\{\underline{\boldsymbol{\omega}}\} = P' = rank\left(\frac{\partial \underline{\boldsymbol{\theta}}(\underline{\boldsymbol{\omega}}^0)}{\partial \underline{\boldsymbol{\omega}}^T}\right), \ P' \leq P.$$

According to (7), the CRB associated to any locally unbiased estimator of $\mathbf{h}(\underline{\boldsymbol{\omega}}^0) = \mathbf{g}(\underline{\boldsymbol{\theta}}(\underline{\boldsymbol{\omega}}^0))$ is given by:

$$\begin{aligned} \mathbf{CRB}_{\underline{\mathbf{h}}|\underline{\boldsymbol{\omega}}}(\underline{\boldsymbol{\omega}}^0) &= \left(\frac{\partial \underline{\mathbf{h}}(\underline{\boldsymbol{\omega}}^0)^T}{\partial \underline{\boldsymbol{\omega}}}\right)^H \mathbf{CRB}_{\underline{\boldsymbol{\omega}}|\underline{\boldsymbol{\omega}}}(\underline{\boldsymbol{\omega}}^0) \frac{\partial \underline{\mathbf{h}}(\underline{\boldsymbol{\omega}}^0)^T}{\partial \underline{\boldsymbol{\omega}}} \\ \mathbf{CRB}_{\underline{\boldsymbol{\omega}}|\underline{\boldsymbol{\omega}}}(\underline{\boldsymbol{\omega}}^0) &= \mathbf{F}_{\underline{\boldsymbol{\omega}}^0}^{-1} \end{aligned}$$

or equivalently by (by use of the derivation chain rule identity):

$$\begin{aligned} \mathbf{CRB}_{\underline{\mathbf{h}}|\underline{\boldsymbol{\omega}}}(\underline{\boldsymbol{\omega}}^0) &= \left(\frac{\partial \underline{\mathbf{g}}(\underline{\boldsymbol{\theta}}(\underline{\boldsymbol{\omega}}^0))^T}{\partial \underline{\boldsymbol{\theta}}}\right)^H \mathbf{CRB}_{\underline{\boldsymbol{\theta}}|\underline{\boldsymbol{\omega}}}(\underline{\boldsymbol{\omega}}^0) \frac{\partial \underline{\mathbf{g}}(\underline{\boldsymbol{\theta}}(\underline{\boldsymbol{\omega}}^0))^T}{\partial \underline{\boldsymbol{\theta}}} \\ \mathbf{CRB}_{\underline{\boldsymbol{\theta}}|\underline{\boldsymbol{\omega}}}(\underline{\boldsymbol{\omega}}^0) &= \left(\frac{\partial \underline{\boldsymbol{\theta}}(\underline{\boldsymbol{\omega}}^0)^T}{\partial \underline{\boldsymbol{\omega}}}\right)^H \mathbf{F}_{\underline{\boldsymbol{\omega}}^0}^{-1} \left(\frac{\partial \underline{\boldsymbol{\theta}}(\underline{\boldsymbol{\omega}}^0)^T}{\partial \underline{\boldsymbol{\omega}}}\right) \\ \mathbf{F}_{\underline{\boldsymbol{\omega}}^0} &= \frac{\partial \underline{\boldsymbol{\theta}}(\underline{\boldsymbol{\omega}}^0)^T}{\partial \underline{\boldsymbol{\omega}}} \mathbf{F}_{\underline{\boldsymbol{\theta}}(\underline{\boldsymbol{\omega}}^0)} \left(\frac{\partial \underline{\boldsymbol{\theta}}(\underline{\boldsymbol{\omega}}^0)^T}{\partial \underline{\boldsymbol{\omega}}}\right)^H \end{aligned} \tag{9}$$

and satisfies the reparameterization inequality introduced in [4]:

$$\mathbf{CRB}_{\underline{\mathbf{g}}|\underline{\boldsymbol{\theta}}}(\underline{\boldsymbol{\theta}}(\underline{\boldsymbol{\omega}}^0)) \geq \mathbf{CRB}_{\underline{\mathbf{h}}|\underline{\boldsymbol{\omega}}}(\underline{\boldsymbol{\omega}}^0) = \mathbf{CRB}_{\underline{\mathbf{g}}(\underline{\boldsymbol{\theta}})|\underline{\boldsymbol{\omega}}}(\underline{\boldsymbol{\omega}}^0) \tag{10}$$

The reparameterization inequality (10) expresses analytically a quite intuitive estimation principle: when the total number of unknown parameters decreases in an observation model ($P' < P$), the asymptotic quality of estimation increases (or remain unchanged), in the sense that the CRB decreases (or remain equal), whatever the function $\mathbf{g}\left(\,\right)$ of the unknown parameters considered. If $P' = P$, then the reparameterization has no effect on the asymptotic quality of estimation since then $\mathbf{CRB}_{\underline{\boldsymbol{\theta}}|\underline{\boldsymbol{\theta}}}\left(\underline{\boldsymbol{\theta}}\left(\underline{\boldsymbol{\omega}}^0\right)\right) = \mathbf{CRB}_{\underline{\boldsymbol{\theta}}|\underline{\boldsymbol{\omega}}}\left(\underline{\boldsymbol{\theta}}\left(\underline{\boldsymbol{\omega}}^0\right)\right)$. Additionally, as shown in [4], regarding the computation of the CRB, equality constraints on parameters: $\underline{\mathbf{f}}\left(\underline{\boldsymbol{\theta}}\right) = \mathbf{0} \in \mathbb{C}^K$, $1 \leq K \leq P$, is a particular case of the reparameterization of the unknown parameters $\underline{\boldsymbol{\theta}}$, provided that the set of constraints is not redundant. Then, the reparameterization inequality (10) holds provided that in (9) $\frac{\partial \underline{\boldsymbol{\theta}}\left(\underline{\boldsymbol{\omega}}^0\right)}{\partial \underline{\boldsymbol{\omega}}^T} \equiv \mathbf{U}_{\underline{\boldsymbol{\theta}}^0}$ where $\mathbf{U}_{\underline{\boldsymbol{\theta}}^0} \in \mathcal{M}_{\mathbb{C}}\left(P, P-K\right)$ is a basis of $\ker\left\{\frac{\partial \underline{\mathbf{f}}\left(\underline{\boldsymbol{\theta}}^0\right)}{\partial \underline{\boldsymbol{\theta}}^T}\right\}$. $\mathbf{U}_{\underline{\boldsymbol{\theta}}^0}$ can always be computed - after rearrangement of $\underline{\boldsymbol{\theta}}$ - as:

$$\mathbf{U}_{\underline{\boldsymbol{\theta}}^0} = \left[\begin{array}{c} \mathbf{I}_{P-K} \\ -\left(\frac{\partial \underline{\mathbf{f}}\left(\underline{\boldsymbol{\theta}}^0\right)}{\partial \underline{\boldsymbol{\varepsilon}}^T}\right)^{-1} \frac{\partial \underline{\mathbf{f}}\left(\underline{\boldsymbol{\theta}}^0\right)}{\partial \underline{\boldsymbol{\omega}}^T} \end{array}\right], \quad \underline{\boldsymbol{\theta}} = \left(\begin{array}{c} \underline{\boldsymbol{\omega}} \\ \underline{\boldsymbol{\varepsilon}} \end{array}\right) \qquad (11)$$

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

3. RADAR ESTIMATION OF DELAY AND VELOCITY

3.1. CRB for active radar

For sake of legibility in the following, the dependency of vectors and matrices of $L^2\left(\Omega\right)$, e.g. $\mathbf{s}\left(f;\underline{\boldsymbol{\theta}}_{\mathbf{s}}\right)$, $\mathbf{b}\left(f;\boldsymbol{\varepsilon}_m\right)$, $\mathbf{B}\left(f;\boldsymbol{\Xi}\right)$ ..., on frequency f will be omitted wherever this omission is unambiguous. In radar, and many other practical problems of interest (radar, sonar, communication, ...), the complex observation vector $\mathbf{x}$ consists of a bandpass signal with bandwidth B $\left(f \in \left[-\frac{B}{2}, \frac{B}{2}\right]\right)$, 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 in the frequency domain:

$$\begin{gathered} \mathbf{x}\left(f;\underline{\boldsymbol{\theta}}\right) = \mathbf{s}\left(f;\underline{\boldsymbol{\theta}}_{\mathbf{s}}\right) + \mathbf{n}\left(f;\underline{\boldsymbol{\theta}}_{\mathbf{n}}\right), \quad \underline{\boldsymbol{\theta}}^T = \left(\underline{\boldsymbol{\theta}}_{\mathbf{s}}^T, \underline{\boldsymbol{\theta}}_{\mathbf{n}}^T\right) \\ \mathbf{s}\left(f;\underline{\boldsymbol{\theta}}_{\mathbf{s}}\right) = \sum_{m=1}^{M} \mathbf{b}\left(f;\boldsymbol{\varepsilon}_m\right)\sigma_m = \mathbf{B}\left(f;\boldsymbol{\Xi}\right)\boldsymbol{\sigma} \end{gathered} \qquad (12)$$

where $\underline{\boldsymbol{\theta}}_{\mathbf{s}}^T = \left(\boldsymbol{\sigma}^T, \boldsymbol{\sigma}^H, \boldsymbol{\Xi}^T\right)$, $\boldsymbol{\Xi}^T = \left(\boldsymbol{\varepsilon}_1^T, \ldots, \boldsymbol{\varepsilon}_M^T\right)$, $\boldsymbol{\sigma} = \left(\sigma_1, \ldots, \sigma_M\right)^T$ and:

• $\mathbf{s}\left(f;\underline{\boldsymbol{\theta}}_{\mathbf{s}}\right)$ is the spectrum of the signal of interest consisting of M backscattered signals function of a parametric propagation model $\mathbf{b}\left(f;\boldsymbol{\varepsilon}_m\right)$ of finite duration T depending on K real parameters $\boldsymbol{\varepsilon}_m^T = \left(\varepsilon_{1,m}, \ldots, \varepsilon_{K,m}\right)$, and of a complex backscattered amplitude σ_m constant during duration T,

• $\mathbf{n}\left(f;\underline{\boldsymbol{\theta}}_{\mathbf{n}}\right)$ is the spectrum of the nuisance signal consisting of noise plus interference contribution depending on the parameters $\underline{\boldsymbol{\theta}}_{\mathbf{n}}$.

Under the assumption of Gaussian centred nuisance and unknown a priori p.d.f. $p\left(\underline{\boldsymbol{\sigma}}\right)$, (12) belongs to the set of deterministic observation models. Additionally, if $\mathbf{n}\left(f;\underline{\boldsymbol{\theta}}_{\mathbf{n}}\right)$ is a wide sense stationary (WSS) band limited process with spectral density matrix $\boldsymbol{\Gamma}\left(f,\underline{\boldsymbol{\theta}}_{\mathbf{n}}\right)$, then for L independent observation of (12):

$$. \qquad \mathbf{x}^l\left(f;\underline{\boldsymbol{\theta}}\right) = \mathbf{B}\left(f;\boldsymbol{\Xi}\right)\boldsymbol{\sigma}^l + \mathbf{n}^l\left(f;\underline{\boldsymbol{\theta}}_{\mathbf{n}}\right)$$

the CRB of the parameters $\boldsymbol{\Xi}$ of the M signals backscattered by the M targets, whatever $\underline{\boldsymbol{\theta}}_{\mathbf{n}}$ is known or unknown, is given by [7][8]:

$$\mathbf{BCR}_{\Xi|\underline{\boldsymbol{\theta}}}^{-1}\left(\underline{\boldsymbol{\theta}}\right) = 2\,\mathrm{Re}\left\{\mathbf{H}_{\Xi|\underline{\boldsymbol{\theta}}}\left(\underline{\boldsymbol{\theta}}\right) \odot \left(\boldsymbol{\Sigma}_s^T \otimes \mathbf{1}_{K\times K}\right)\right\} \qquad (13)$$

$$\mathbf{H}_{\Xi|\underline{\boldsymbol{\theta}}}\left(\underline{\boldsymbol{\theta}}\right) = \left[\begin{array}{ccc} \mathbf{H}\left(\underline{\boldsymbol{\theta}}\right)_{1,1} & \ldots & \mathbf{H}\left(\underline{\boldsymbol{\theta}}\right)_{1,M} \\ \vdots & \ddots & \vdots \\ \mathbf{H}\left(\underline{\boldsymbol{\theta}}\right)_{M,1} & \ldots & \mathbf{H}\left(\underline{\boldsymbol{\theta}}\right)_{M,M} \end{array}\right], \; \boldsymbol{\Sigma}_s = \frac{1}{L}\sum_{l=1}^{L}\boldsymbol{\sigma}^l\left(\boldsymbol{\sigma}^l\right)^H$$

$$\mathbf{H}\left(\underline{\boldsymbol{\theta}}\right)_{m_1,m_2} = \mathbf{G}_{\underline{\boldsymbol{\theta}}_{\mathbf{n}}}\left(\Pi_{\{\mathbf{B}(\boldsymbol{\Xi})\}}^{\perp}\left(\left\{\frac{\partial \mathbf{b}\left(\boldsymbol{\varepsilon}_{m_2}\right)}{\partial \boldsymbol{\varepsilon}^T}\right\}\right), \Pi_{\{\mathbf{B}(\boldsymbol{\Xi})\}}^{\perp}\left(\left\{\frac{\partial \mathbf{b}\left(\boldsymbol{\varepsilon}_{m_1}\right)}{\partial \boldsymbol{\varepsilon}^T}\right\}\right)\right)$$

where $\mathbf{1}_{P\times P}$ is a $P \times P$ matrix of ones and the Hermitian product is: $\left\langle \mathbf{x} \mid \mathbf{y}\right\rangle_{\underline{\boldsymbol{\theta}}_{\mathbf{n}}} = \int_{-\frac{B}{2}}^{\frac{B}{2}} \mathbf{y}\left(f\right)^H \boldsymbol{\Gamma}\left(f,\underline{\boldsymbol{\theta}}_{\mathbf{n}}\right)^{-1}\mathbf{x}\left(f\right)df$.

$\Pi_{\{\mathbf{B}(\boldsymbol{\Xi})\}}^{\perp}$ denotes the orthonormal projector on the orthogonal complement of $span\left\{\mathbf{B}\left(f;\boldsymbol{\Xi}\right)\right\}$: $\Pi_{\{\mathbf{B}(\boldsymbol{\Xi})\}}\left(\mathbf{a}\right) + \Pi_{\{\mathbf{B}(\boldsymbol{\Xi})\}}^{\perp}\left(\mathbf{a}\right) = \mathbf{a}$ where $\Pi_{\{\mathbf{B}(\boldsymbol{\Xi})\}}\left(\mathbf{a}\right) = \mathbf{B}\left(\boldsymbol{\Xi}\right)\mathbf{G}_{\underline{\boldsymbol{\theta}}_{\mathbf{n}}}^{-1}\left(\left\{\mathbf{B}\left(\boldsymbol{\Xi}\right)\right\}\right)\mathbf{G}_{\boldsymbol{\theta}_{\mathbf{n}}}\left(\mathbf{a}, \left\{\mathbf{B}\left(\boldsymbol{\Xi}\right)\right\}\right)$.

3.2. Estimation of delay and velocity

In the present paper we consider a radar system consisting of a 1-element antenna array receiving scaled, timedelayed, and Doppler-shifted echoes of a known complex bandpass signal $e_T\left(t\right)e^{-j2\pi f_c t}$, where f_c is the carrier frequency. The antenna receives a pulse train (burst) of L pulses of width T_0 and bandwidth B, with a pulse repetition interval (PRI) T. Under the usual approximation in radar of "slow" moving targets in comparison with $e_T\left(t\right)$, i.e. [5] :

• $\left|2v_m\left(L-1\right)T\right| << \frac{c}{B}$ (no range migration),

• $\frac{2v_m}{\lambda_c}T_0 << 1$ (Doppler effect on $e_T\left(t\right)$ is negligible),

the standard hypothesis of temporally white nuisance signal (thermal noise) of power σ_n^2 and non fluctuating targets during the burst duration, then (12) can be significantly simplified and becomes [5]:

$$x^l\left(t;\underline{\boldsymbol{\theta}}\right) = \sum_{m=1}^{M} e_T\left(t-\tau_m\right)\sigma_m^l + n^l\left(t\right) \qquad (14)$$

$$\sigma_m^l = \sigma_m e^{j2\pi\omega_m(l-1)T}, \quad \omega_m = \frac{-2v_m}{\lambda_c} \qquad (15)$$

Then:

$$\boldsymbol{\Sigma}_s = \frac{1}{L}\sum_{l=1}^{L}\boldsymbol{\sigma}^l\left(\boldsymbol{\sigma}^l\right)^H = \left(\boldsymbol{\sigma}\boldsymbol{\sigma}^H\right)\odot\left(\boldsymbol{\Psi}^H\boldsymbol{\Psi}\right) \qquad (16)$$

$$\boldsymbol{\Psi} = \left[\ldots\; \boldsymbol{\psi}\left(\omega_m\right)\; \ldots\right], \quad \boldsymbol{\psi}\left(\omega\right) = \left(1, \ldots, e^{j2\pi\omega(L-1)T}\right)^T$$

Observation model (14) allows for a dual modeling of unknown target parameters according to whether the Doppler information is a parameter of interest or not:

• the noncoherent observation model where the inter-observation relationship containing the Doppler information (14) is not taken into account. In this approach, the set of unknown parameters is $\underline{\boldsymbol{\theta}}^T = \left(\boldsymbol{\sigma}_1^T, \boldsymbol{\sigma}_1^H, \ldots, \boldsymbol{\sigma}_L^T, \boldsymbol{\sigma}_L^H, \boldsymbol{\tau}^T, \sigma_n^2\right)$ and the targets parameters of interest reduce to the delays $\boldsymbol{\tau}$. Then, by application of (13):

$$\mathbf{BCR}_{\boldsymbol{\tau}|\underline{\boldsymbol{\theta}}}\left(\underline{\boldsymbol{\theta}}\right) = \mathrm{Re}\left\{2\mathbf{H}_{\boldsymbol{\tau}|\underline{\boldsymbol{\theta}}}\left(\underline{\boldsymbol{\theta}}\right)\odot\left(\boldsymbol{\Psi}^T\boldsymbol{\Psi}^*\right)\odot\left(\boldsymbol{\sigma}^T\boldsymbol{\sigma}^*\right)\right\}^{-1} \qquad (17)$$

• the coherent observation model where the inter-observation relationship containing the Doppler information (14) is taken into account. Then the set of unknown parameters is $\underline{\widetilde{\boldsymbol{\theta}}}^T = \left(\boldsymbol{\sigma}^T, \boldsymbol{\sigma}^H, \boldsymbol{\tau}^T, \boldsymbol{\omega}^T, \sigma_n^2\right)$ and the targets parameters of interest are the delays $\boldsymbol{\tau}$ and the Doppler frequencies $\boldsymbol{\omega}$. Then, by application of (13):

$$\mathbf{BCR}_{(\boldsymbol{\tau},\boldsymbol{\omega})|\underline{\widetilde{\boldsymbol{\theta}}}}\left(\underline{\widetilde{\boldsymbol{\theta}}}\right) = \mathrm{Re}\left\{2\mathbf{H}_{(\boldsymbol{\tau},\boldsymbol{\omega})|\underline{\widetilde{\boldsymbol{\theta}}}}\left(\underline{\widetilde{\boldsymbol{\theta}}}\right) \odot \left(\left(\boldsymbol{\sigma}^T\boldsymbol{\sigma}^*\right) \otimes \mathbf{1}_{2\times 2}\right)\right\}^{-1} \tag{18}$$

Expressions of $\mathbf{H}_{\boldsymbol{\tau}|\underline{\boldsymbol{\theta}}}\left(\underline{\boldsymbol{\theta}}\right)$ and $\mathbf{H}_{(\boldsymbol{\tau},\boldsymbol{\omega})|\underline{\widetilde{\boldsymbol{\theta}}}}\left(\underline{\widetilde{\boldsymbol{\theta}}}\right)$ are more legible if the sampled version of (14) is considered:

$$x^l\left(i;\underline{\boldsymbol{\theta}}\right) = \sum_{m=1}^{M} e_T\left(\tfrac{i}{B} - \tau_m\right)\sigma_m^l + n^l\left(\tfrac{i}{B}\right)$$

We assume that the snapshots taken at $i = 1\ldots I$ cover the whole observation interval including all the timedelayed echoes $e_T\left(t-\tau_m\right)$, and let:

$$\begin{aligned}&\mathbf{a}\left(\tau\right) = \left[\ldots, e_T\left(\tfrac{i}{B}-\tau\right), \ldots\right]^T, \quad \mathbf{A} = \left[\mathbf{a}\left(\tau_1\right) \ \cdots \ \mathbf{a}\left(\tau_M\right)\right]\\ &\mathbf{D_a} = \left[\tfrac{\partial \mathbf{a}(\tau_1)}{\partial \tau} \ \cdots \ \tfrac{\partial \mathbf{a}(\tau_M)}{\partial \tau}\right], \quad \mathbf{D}_{\boldsymbol{\psi}} = \left[\tfrac{\partial \boldsymbol{\psi}(\omega_1)}{\partial \omega} \ \cdots \ \tfrac{\partial \boldsymbol{\psi}(\omega_M)}{\partial \omega}\right]\end{aligned}$$

Then a few lines of straightforward calculus lead to:

$$\begin{aligned}\mathbf{H}_{\boldsymbol{\tau}|\underline{\boldsymbol{\theta}}}\left(\underline{\boldsymbol{\theta}}\right) &= \mathbf{D_a}^H\mathbf{D_a} - \mathbf{D_a}^H\mathbf{A}\left(\mathbf{A}^H\mathbf{A}\right)^{-1}\mathbf{A}^H\mathbf{D_a}\\ \mathbf{H}_{(\boldsymbol{\tau},\boldsymbol{\omega})|\underline{\widetilde{\boldsymbol{\theta}}}}\left(\underline{\widetilde{\boldsymbol{\theta}}}\right) &= \begin{bmatrix}\mathbf{H}_{\boldsymbol{\tau}|\underline{\widetilde{\boldsymbol{\theta}}}} & \mathbf{H}_{\boldsymbol{\tau},\boldsymbol{\omega}|\underline{\widetilde{\boldsymbol{\theta}}}}\\ \mathbf{H}^H_{(\boldsymbol{\tau},\boldsymbol{\omega})|\underline{\widetilde{\boldsymbol{\theta}}}} & \mathbf{H}_{\boldsymbol{\omega}|\underline{\widetilde{\boldsymbol{\theta}}}}\end{bmatrix}\end{aligned} \tag{19}$$

where:

$$\begin{aligned}\mathbf{H}_{\boldsymbol{\tau}|\underline{\widetilde{\boldsymbol{\theta}}}} &= \boldsymbol{\Psi}^H\boldsymbol{\Psi} \odot \mathbf{D_a}^H\mathbf{D_a}\\ &\quad - \left(\boldsymbol{\Psi}^H\boldsymbol{\Psi} \odot \mathbf{D_a}^H\mathbf{A}\right)\left(\boldsymbol{\Psi}^H\boldsymbol{\Psi} \odot \mathbf{A}^H\mathbf{A}\right)^{-1}\left(\boldsymbol{\Psi}^H\boldsymbol{\Psi} \odot \mathbf{A}^H\mathbf{D_a}\right)\\ \mathbf{H}_{\boldsymbol{\tau},\boldsymbol{\omega}|\underline{\widetilde{\boldsymbol{\theta}}}} &= \boldsymbol{\Psi}^H\mathbf{D}_{\boldsymbol{\psi}} \odot \mathbf{D_a}^H\mathbf{A}\\ &\quad - \left(\boldsymbol{\Psi}^H\boldsymbol{\Psi} \odot \mathbf{D_a}^H\mathbf{A}\right)\left(\boldsymbol{\Psi}^H\boldsymbol{\Psi} \odot \mathbf{A}^H\mathbf{A}\right)^{-1}\left(\boldsymbol{\Psi}^H\mathbf{D}_{\boldsymbol{\psi}} \odot \mathbf{A}^H\mathbf{A}\right)\\ \mathbf{H}_{\boldsymbol{\omega}|\underline{\widetilde{\boldsymbol{\theta}}}} &= \mathbf{D}_{\boldsymbol{\psi}}^H\mathbf{D}_{\boldsymbol{\psi}} \odot \mathbf{A}^H\mathbf{A}\\ &\quad - \left(\mathbf{D}_{\boldsymbol{\psi}}^H\boldsymbol{\Psi} \odot \mathbf{A}^H\mathbf{A}\right)\left(\boldsymbol{\Psi}^H\boldsymbol{\Psi} \odot \mathbf{A}^H\mathbf{A}\right)^{-1}\left(\boldsymbol{\Psi}^H\mathbf{D}_{\boldsymbol{\psi}} \odot \mathbf{A}^H\mathbf{A}\right)\end{aligned}$$

Due to the presence of the Hadamard and Kroneker products, it is impossible to compare (with respect to the Löwner ordering) directly $\mathbf{BCR}_{\boldsymbol{\tau}|\underline{\boldsymbol{\theta}}}\left(\underline{\boldsymbol{\theta}}\right)$ and $\mathbf{BCR}_{\boldsymbol{\tau}|\underline{\widetilde{\boldsymbol{\theta}}}}\left(\underline{\widetilde{\boldsymbol{\theta}}}\right)$ (obtained by matrix inversion in block form). Therefore, so far (to the best of our knowledge) it has been impossible to answer to the following question:
if one is only interested in the delay estimation, what is the best (in terms of asymptotic MSE) observation model to take into account? The coherent or the noncoherent one?
The answer is given by the reparameterization inequality (10) as one can note that the coherent observation model is obtained from the noncoherent observation model by introducing $(L-1)\times M$ equality constraints on the $\boldsymbol{\sigma}_l$:

$$\sigma_m^l - \sigma_m e^{j2\pi\omega_m(l-1)T} = 0, \quad 2 \leq l \leq L, \quad 1 \leq m \leq M,$$

or by introducing an injective reparameterization of the $\boldsymbol{\sigma}_l$:

$$\boldsymbol{\sigma}_l = \boldsymbol{\sigma}_l\left(\boldsymbol{\sigma}, \boldsymbol{\omega}\right), \quad 1 \leq l \leq L.$$

Then (10) allow to state that:

$$\mathbf{BCR}_{\boldsymbol{\tau}|\underline{\widetilde{\boldsymbol{\theta}}}}\left(\underline{\widetilde{\boldsymbol{\theta}}}\right) \leq \mathbf{BCR}_{\boldsymbol{\tau}|\underline{\boldsymbol{\theta}}}\left(\underline{\boldsymbol{\theta}}\right) \tag{20}$$

Additionally, if all the targets have the same velocity: $\omega_m = \omega, 1 \leq m \leq M$, then $\mathbf{H}_{\boldsymbol{\tau},\boldsymbol{\omega}|\underline{\widetilde{\boldsymbol{\theta}}}} = \mathbf{0}$ and equality is reached $\mathbf{BCR}_{\boldsymbol{\tau}|\underline{\widetilde{\boldsymbol{\theta}}}}\left(\underline{\widetilde{\boldsymbol{\theta}}}\right) = \mathbf{BCR}_{\boldsymbol{\tau}|\underline{\boldsymbol{\theta}}}\left(\underline{\boldsymbol{\theta}}\right)$.
As a consequence, the reparameterization inequality is a key tool for system analysis and design in the asymptotic region since it allows to state the following design principles for delay and Doppler estimation:
• if your main requirement is performance estimation, then the Doppler information must always be taken into account (has an unknown parameter) when you estimate delay, at the expense of a more complex ML algorithm,
• if your main requirement is delay estimation of a set of point scatterers having the same velocity $\boldsymbol{\omega} = \omega\mathbf{1}_{M\times 1}$ (point scatterers of the body of a target for example), then the best processing (same asymptotic performance but less computations) is the noncoherent MLE:

$$\widehat{\boldsymbol{\tau}} = \arg\max_{\boldsymbol{\tau}}\left\{\sum_{l=1}^{L}\left\|\boldsymbol{\Pi}_{\mathbf{A}}\mathbf{x}_l\right\|^2\right\}, \quad \boldsymbol{\Pi}_{\mathbf{A}} = \mathbf{A}\left(\mathbf{A}^H\mathbf{A}\right)^{-1}\mathbf{A}^H$$

in comparison with the coherent MLE:

$$\left(\widehat{\boldsymbol{\tau}},\widehat{\omega}\right) = \arg\max_{(\boldsymbol{\tau},\omega)}\left\{\left\|\boldsymbol{\Pi}_{\mathbf{A}}\left(\left[\mathbf{x}_1 \ \cdots \ \mathbf{x}_L\right]\boldsymbol{\psi}^*\right)\right\|^2\right\}.$$

4. CONCLUSION

For system design and optimization it is worth knowing the general reparameterization inequality (10) derived in [4]. 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).

5. REFERENCES

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

[2] 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, vol. 56, no. 11, pp. 5319-5333, Nov. 2008

[3] 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, vol. 56, no. 10, Oct. 2010

[4] T. Menni, E. Chaumette, P. Larzabal and J. P. Barbot, "New results on Deterministic Cramér-Rao bounds for real and complex parameters", accepted for publication with minor revision to IEEE Trans. on SP, jan 2011

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

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

[7] S. F. Yau and Y. Bresler, "A Compact Cram&-Rao Bound Expression for Parametric Estimation of Superimposed Signals", IEEE Trans. on SP, vol. 40, no. 5, pp 1226-1230, May 1992.

[8] T. Menni, E. Chaumette, P. Larzabal and J. P. Barbot, "Crb for Active Radar", in Proc European Signal Processing Conference, 2011 Barcelona Spain