

Recursive hybrid Cramer-Rao bound for discrete-time Markovian dynamic systems

Chengfang Ren, Jérôme Galy, Eric Chaumette, François Vincent, Pascal Larzabal, Alexandre Renaux

▶ **To cite this version:**

Chengfang Ren, Jérôme Galy, Eric Chaumette, François Vincent, Pascal Larzabal, et al.. Recursive hybrid Cramer-Rao bound for discrete-time Markovian dynamic systems. IEEE Signal Processing Letters, 2015, 22 (10), pp.1543-1547. 10.1109/LSP.2015.2412173 . hal-01226835

HAL Id: hal-01226835
https://hal.science/hal-01226835

Submitted on 10 Nov 2015

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Recursive hybrid Cramér-Rao bound for discrete-time Markovian dynamic systems

Chengfang Ren, Jerome Galy, Eric Chaumette, François Vincent, Pascal Larzabal and Alexandre Renaux

***Abstract*—In statistical signal processing, hybrid parameter estimation refers to the case where the parameters vector to estimate contains both non-random and random parameters. As a contribution to the hybrid estimation framework, we introduce a recursive hybrid Cramér Rao lower bounds for discrete-time Markovian dynamic systems depending on unknown deterministic parameters. Additionnally, the regularity conditions required for its existence and its use are clarified.**

***Index Terms*—Parameter estimation, dynamic Markovian systems, estimation error lower bound**

I. Introduction

Since its introduction in the context of array shape calibration [1], hybrid parameter estimation has given rise to a growing interest as both random and nonrandom parameters occur simultaneously in miscellaneous estimation problems [2]-[9]. However, the hybrid estimation framework is not just the simple concatenation of the Bayesian and non-Bayesian techniques and new estimators have to be derived [10, §1.1]. Similarly, performance analysis methods of such estimators has to be modified accordingly, which is the aim of hybrid lower bounds on the mean square error (MSE).

The first hybrid lower bound, the so-called hybrid Cramér-Rao bound (HCRB), has been introduced in [1] where the random parameters have a prior probability density function (pdf) independent from deterministic parameters. This initial characterization of hybrid estimation has been generalized by Reuven and Messer [2] who introduced the first "large-error" hybrid bound, the so-called hybrid Barankin Bound (HBB), in order to handle the threshold phenomena and of which one limiting form yields the HCRB. This seminal work [2] has been lately extended to new "large-error" hybrid bounds [7][11][12] in order to improve the estimation of the transition region where the threshold phenomena occurs. Unfortunately, the computational cost of hybrid "large-error" bounds is prohibitive in most applications when the number of unknown parameters increases. Concurrently, an extension of the HCRB where the prior pdf of the random parameters depends on deterministic parameters was proposed in [6] and its asymptotic tightness was further analyzed in [13]. All these works have shown that, like the deterministic CRB and Bayesian CRB (BCRB), the HCRB is valid in the asymptotic region only, i.e., when signal to noise ratio is high or the number of observations is large.

In the Bayesian estimation framework, discrete-time Markovian dynamic systems (MDS) arises in various applications such as adaptive control, analysis, and prediction of non-stationary time series [14]. As is well known, the optimal estimator for this problem cannot be built in general, and it is necessary to turn to one of the large number of existing suboptimal filtering techniques [14]. Assessing the achievable performance may be difficult, and we have to resort to simulations and comparing proximity to bayesian lower bounds corresponding to optimum performance [10][15][16]. Actually, most discrete-time MDS incorporate some deterministic parameters which can be either known [10] or unknown [6] according to the experimental conditions. Even when the deterministic parameters are known, some of the true values may originate from a prior calibration process which accuracy impacts on the optimum performance of random parameter estimates. In both cases, there is a need for computationally tractable hybrid lower bounds for discrete-time MDS depending on unknown deterministic parameters.

As a contribution, we introduce the first recursive form of an hybrid lower bound for discrete-time MDS, namely the recursive HCRB, which, provided that one keeps in mind its limitations, is a lower bound of great interest for system analysis and design in the asymptotic region. Additionally we discuss the regularity conditions required for the existence and the use of the recursive HCRB, which are critical to understand not only the applicability limit of the recursive HCRB but also why, in most case, the posterior BCRB cannot be transformed into the recursive HCRB as it is misleadingly suggested in [15][16].

II. Recursive HCRB for Discrete-Time MDS

In hybrid parameter estimation one wishes to estimate an unknown hybrid parameter vector $(\mathbf{x};\boldsymbol{\theta})$[1] from a random observation vector $\mathbf{y} \in \mathbb{R}^{N'}$. Some prior knowledge is available on random parameter $\mathbf{x} \in \mathbb{R}^{P'}$ that is incorporated by an *a priori* pdf $p(\mathbf{x})$ which support is a subset Π_r of $\mathbb{R}^{P'}$. No such knowledge is available on $\boldsymbol{\theta} \in \Pi_d \subset \mathbb{R}^{D'}$ and

[1]For L column vectors $\mathbf{a}_l$, $(\mathbf{a}_1;\mathbf{a}_2;\ldots;\mathbf{a}_L) \triangleq (\mathbf{a}_1^T,\mathbf{a}_2^T,\ldots,\mathbf{a}_L^T)^T$ denotes the vertical concatenation

Chengfang Ren and Alexandre Renaux are with Université Paris-Sud/LSS 3, Rue Joliot-Curie, 91192 Gif-sur-Yvette, France. Emails: cren@lss.supelec.fr, renaux@lss.supelec.fr

Jérome Galy is with Université de Montpellier 2/LIRMM, 161 rue Ada 34392 Montpellier Cedex 5, France. Email: galy@lirmm.fr

Eric Chaumette and François Vincent are with University of Toulouse-ISAE, DEOS, 10 Avenue Edouard Belin, 31055 Toulouse, France. Emails: eric.chaumette@isae.fr, francois.vincent@isae.fr

Pascal Larzabal is with Université Paris-Sud/SATIE, 61 av. du President Wilson, 94235 Cachan, France. Email: pascal.larzabal@satie.ens-cachan.fr

This work has been partially supported by the European Network of excellence NEWCOM#, by the iCODE institute, research project of the IDEX Paris-Saclay, by the DGA/DGCIS and by the Display-Mastodons project of CNRS.

thus it is considered deterministic. In the general case, $p(\mathbf{x})$ may depend on the unknown parameter $\boldsymbol{\theta}$, and it is denoted $p(\mathbf{x}|\boldsymbol{\theta})$. The conditional pdf of $\mathbf{y}$ given $\mathbf{x}$ parameterized by $\boldsymbol{\theta}$ is $p(\mathbf{y}|\mathbf{x},\boldsymbol{\theta})$ and their joint pdf parameterized by $\boldsymbol{\theta}$ is given by $p(\mathbf{y},\mathbf{x}|\boldsymbol{\theta}) = p(\mathbf{y}|\mathbf{x},\boldsymbol{\theta})\,p(\mathbf{x}|\boldsymbol{\theta})$. Then, for any estimators $\widehat{\boldsymbol{\theta}}(\mathbf{y})$ of $\boldsymbol{\theta}$ and $\widehat{\mathbf{x}}(\mathbf{y})$ of $\mathbf{x}$, one of the possible lower bounds [10][11][12] deriving from the covariance inequality principle (16) is the HCRB which usual form is given by [10]:

$$E_{\mathbf{y},\mathbf{x}|\boldsymbol{\theta}}\left[\mathbf{e}(\mathbf{y})\,\mathbf{e}(\mathbf{y})^T\right] \succeq \mathbf{HCRB}_{\mathbf{x},\boldsymbol{\theta}} = \mathbf{J}_{\mathbf{x},\boldsymbol{\theta}}^{-1}, \tag{1}$$

$$\mathbf{J}_{\mathbf{x},\boldsymbol{\theta}} = E_{\mathbf{y},\mathbf{x}|\boldsymbol{\theta}}\left[\frac{\partial \ln p(\mathbf{y},\mathbf{x}|\boldsymbol{\theta})}{\partial(\mathbf{x};\boldsymbol{\theta})}\frac{\partial \ln p(\mathbf{y},\mathbf{x}|\boldsymbol{\theta})}{\partial(\mathbf{x};\boldsymbol{\theta})^T}\right], \tag{2}$$

where $\mathbf{e}(\mathbf{y}) = (\widehat{\mathbf{x}}(\mathbf{y}) - \mathbf{x}; \widehat{\boldsymbol{\theta}}(\mathbf{y}) - \boldsymbol{\theta})$, $E_{\mathbf{y},\mathbf{x}|\boldsymbol{\theta}}[\mathbf{g}(\mathbf{y},\mathbf{x})]$ is the statistical expectation of the vector of functions $\mathbf{g}(\,)$ with respect to $\mathbf{y}$ and $\mathbf{x}$ parameterized by $\boldsymbol{\theta}$, and for two matrices, $\mathbf{A} \succeq \mathbf{B}$ means that $\mathbf{A} - \mathbf{B}$ is positive semi-definite. The regularity conditions for the hybrid Fisher information matrix (HFIM) $\mathbf{J}_{\mathbf{x},\boldsymbol{\theta}}$ to be of the usual form (2) are (see Section III):
R1): $\Pi_r = \mathbb{R}^P$,
(R2): $E_{\mathbf{y},\mathbf{x}|\boldsymbol{\theta}}\left[\frac{\partial \ln p(\mathbf{y},\mathbf{x}|\boldsymbol{\theta})}{\partial x_p}^2\right], E_{\mathbf{y},\mathbf{x}|\boldsymbol{\theta}}\left[\frac{\partial \ln p(\mathbf{y},\mathbf{x}|\boldsymbol{\theta})}{\partial \theta_d}^2\right] < \infty$.
Moreover, under its usual form (1), the HCRB is a lower bound for the class of estimates satisfying (R3): $E_{\mathbf{y},\mathbf{x}|\boldsymbol{\theta}}\left[\left(\widehat{\mathbf{x}}(\mathbf{y}) - \mathbf{x}; \widehat{\boldsymbol{\theta}}(\mathbf{y}) - \boldsymbol{\theta}\right)\right] = (\boldsymbol{\mu}; \mathbf{0})$.
Our main concern is the derivation of a computationally tractable HFIM (2) for hybrid discrete-time MDS represented with the state and measurement equations:

$$\mathbf{x}_k = \mathbf{f}_{k-1}(\mathbf{x}_{k-1}, \mathbf{w}_{k-1}, \boldsymbol{\alpha}),\ \mathbf{y}_k = \mathbf{h}_k(\mathbf{x}_k, \mathbf{v}_k, \boldsymbol{\lambda}) \tag{3}$$

where $k \geq 1$ is a time index, $\mathbf{x}_k$ is the P-dimensional state vector, $\mathbf{y}_k$ is the N-dimensional measurement vector, $f_k(\ ,\ ,\boldsymbol{\alpha})$ and $\mathbf{h}_k(\ ,\ ,\boldsymbol{\lambda})$ are known parametric vector functions depending on an unknown deterministic parameter vector ($\boldsymbol{\alpha}$ and $\boldsymbol{\lambda}$ respectively). The process noise sequence $\{\mathbf{w}_k\}$ and the measurement noise sequence $\{\mathbf{v}_k\}$ are mutually independent white sequences described by known pdfs $p(\mathbf{w}_k|\boldsymbol{\beta})$ and $p(\mathbf{v}_k|\boldsymbol{\mu})$, respectively, depending on an unknown deterministic parameter vector ($\boldsymbol{\beta}$ and $\boldsymbol{\mu}$ respectively). The noises are independent of the initial state $\mathbf{x}_0$ described by the known pdf $p(\mathbf{x}_0|\boldsymbol{\alpha})$. Let $\boldsymbol{\theta} = (\boldsymbol{\lambda};\boldsymbol{\mu};\boldsymbol{\alpha};\boldsymbol{\beta})$ be the vector gathering all the unknown deterministic parameters. The state transition and the measurement pdfs depend on unknown deterministic parameters:

$$p(\mathbf{x}_k|\mathbf{x}_{k-1}) \triangleq p(\mathbf{x}_k|\mathbf{x}_{k-1},\boldsymbol{\alpha},\boldsymbol{\beta}), p(\mathbf{y}_k|\mathbf{x}_k) \triangleq p(\mathbf{y}_k|\mathbf{x}_k,\boldsymbol{\theta}),$$

and we suppose that both $p(\mathbf{x}_k|\mathbf{x}_{k-1})$ and $p(\mathbf{y}_k|\mathbf{x}_k)$ are twice differentiable with respect to all their arguments. We adopt the notational convention: $\forall l \leq k, \overline{\boldsymbol{\delta}}_{l:k} = (\boldsymbol{\delta}_l;\ldots;\boldsymbol{\delta}_k)$ where all vectors are of same dimension. Since (3) is a MDS:

$$p(\overline{\mathbf{y}}_{1:k}, \overline{\mathbf{x}}_{0:k}|\boldsymbol{\theta}) = p(\mathbf{x}_0|\boldsymbol{\alpha}) \prod_{l=1}^{k} p(\mathbf{y}_l|\mathbf{x}_l,\boldsymbol{\theta})\,p(\mathbf{x}_l|\mathbf{x}_{l-1},\boldsymbol{\alpha},\boldsymbol{\beta}) \tag{4}$$

From a theoretical point of view, we are primarily interested in the HFIM (2) on $(\mathbf{x};\boldsymbol{\theta}) \triangleq (\mathbf{x}_k;\boldsymbol{\theta})$ associated to the observation vector $\mathbf{y} \triangleq \overline{\mathbf{y}}_{1:k}$ resulting from the set of k measurements $\mathbf{y}_1$,...,$\mathbf{y}_k$:

$$\mathbf{J}_{\mathbf{x}_k,\boldsymbol{\theta}} = E_{\overline{\mathbf{y}}_{1:k},\mathbf{x}_k|\boldsymbol{\theta}}\left[\frac{\partial \ln p(\overline{\mathbf{y}}_{1:k},\mathbf{x}_k|\boldsymbol{\theta})}{\partial(\mathbf{x}_k;\boldsymbol{\theta})}\frac{\partial \ln p(\overline{\mathbf{y}}_{1:k},\mathbf{x}_k|\boldsymbol{\theta})}{\partial(\mathbf{x}_k;\boldsymbol{\theta})^T}\right]$$

which alternative formula is [1][2][10]:

$$\mathbf{J}_{\mathbf{x}_k,\boldsymbol{\theta}} = E_{\overline{\mathbf{y}}_{1:k},\mathbf{x}_k|\boldsymbol{\theta}}\left[\frac{-\partial^2 \ln p(\overline{\mathbf{y}}_{1:k},\mathbf{x}_k|\boldsymbol{\theta})}{\partial(\mathbf{x}_k;\boldsymbol{\theta})\,\partial(\mathbf{x}_k;\boldsymbol{\theta})^T}\right] \tag{5}$$

Unfortunately the computation of (5) requires the derivation of the marginal pdf $p(\overline{\mathbf{y}}_{1:k},\mathbf{x}_k|\boldsymbol{\theta})$ from (4) which is generally mathematically intractable [1][10][15]. However an upper bound of (5) can be derived from the HFIM (2) on $(\mathbf{x};\boldsymbol{\theta}) \triangleq (\overline{\mathbf{x}}_{0:k};\boldsymbol{\theta})$ associated to the observation vector $\mathbf{y} \triangleq \overline{\mathbf{y}}_{1:k}$:

$$\mathbf{J}_{\overline{\mathbf{x}}_{0:k},\boldsymbol{\theta}} = E_{\overline{\mathbf{y}}_{1:k},\overline{\mathbf{x}}_{0:k}|\boldsymbol{\theta}}\left[\frac{-\partial^2 \ln p(\overline{\mathbf{y}}_{1:k},\overline{\mathbf{x}}_{0:k}|\boldsymbol{\theta})}{\partial(\overline{\mathbf{x}}_{0:k};\boldsymbol{\theta})\,\partial(\overline{\mathbf{x}}_{0:k};\boldsymbol{\theta})^T}\right]. \tag{6}$$

Indeed, if we decompose $\mathbf{J}_{\overline{\mathbf{x}}_{0:k},\boldsymbol{\theta}}$ as:

$$\mathbf{J}_{\overline{\mathbf{x}}_{0:k},\boldsymbol{\theta}} = \begin{bmatrix} \mathbf{J}_k^{11} & \mathbf{B}_k \\ \mathbf{B}_k^T & \mathbf{C}_k \end{bmatrix}, \begin{array}{l} \mathbf{J}_k^{11} = E_{\overline{\mathbf{y}}_{1:k},\overline{\mathbf{x}}_{0:k}|\boldsymbol{\theta}}\left[-\frac{\partial^2 \ln p(\overline{\mathbf{y}}_{1:k},\overline{\mathbf{x}}_{0:k}|\boldsymbol{\theta})}{\partial \overline{\mathbf{x}}_{0:k-1}\partial \overline{\mathbf{x}}_{0:k-1}^T}\right] \\ \mathbf{B}_k = E_{\overline{\mathbf{y}}_{1:k},\overline{\mathbf{x}}_{0:k}|\boldsymbol{\theta}}\left[-\frac{\partial^2 \ln p(\overline{\mathbf{y}}_{1:k},\overline{\mathbf{x}}_{0:k}|\boldsymbol{\theta})}{\partial \overline{\mathbf{x}}_{0:k-1}\partial(\mathbf{x}_k;\boldsymbol{\theta})^T}\right] \\ \mathbf{C}_k = E_{\overline{\mathbf{y}}_{1:k},\overline{\mathbf{x}}_{0:k}|\boldsymbol{\theta}}\left[-\frac{\partial^2 \ln p(\overline{\mathbf{y}}_{1:k},\mathbf{x}_k|\boldsymbol{\theta})}{\partial(\mathbf{x}_k;\boldsymbol{\theta})\partial(\mathbf{x}_k;\boldsymbol{\theta})^T}\right] \end{array},$$

then a generalization of Proposition 1 in [17] (which proof follows lines similar to the proof of Theorem 1 in [13]) to hybrid estimation yields:

$$\widetilde{\mathbf{J}}_{\mathbf{x}_k,\boldsymbol{\theta}} \succeq \mathbf{J}_{\mathbf{x}_k,\boldsymbol{\theta}} \ \Leftrightarrow\ \mathbf{HCRB}_{\mathbf{x}_k,\boldsymbol{\theta}} \succeq \widetilde{\mathbf{J}}_{\mathbf{x}_k,\boldsymbol{\theta}}^{-1}, \tag{7}$$

where:

$$\widetilde{\mathbf{J}}_{\mathbf{x}_k,\boldsymbol{\theta}} = \mathbf{C}_k - \mathbf{B}_k^T\left(\mathbf{J}_k^{11}\right)^{-1}\mathbf{B}_k, \tag{8}$$

and leading to:

$$E_{\mathbf{y},\mathbf{x}|\boldsymbol{\theta}}\left[\mathbf{e}(\mathbf{y})\,\mathbf{e}(\mathbf{y})^T\right] \succeq \widetilde{\mathbf{J}}_{\mathbf{x}_k,\boldsymbol{\theta}}^{-1}. \tag{9}$$

$\widetilde{\mathbf{J}}_{\mathbf{x}_k,\boldsymbol{\theta}}^{-1}$ defines a looser (in comparison with $\mathbf{HCRB}_{\mathbf{x}_k,\boldsymbol{\theta}}$) but general computable hybrid bound for discrete-time MDS pdf (4). Additionally $\widetilde{\mathbf{J}}_{\mathbf{x}_k,\boldsymbol{\theta}}$ is computationally tractable: it can be assessed without computing the inverse of large matrices such as $(PK \times PK)$ matrix $\mathbf{J}_k^{11}$. Indeed $\widetilde{\mathbf{J}}_{\mathbf{x}_k,\boldsymbol{\theta}}$ (8) can be decomposed into block matrices:

$$\widetilde{\mathbf{J}}_{\mathbf{x}_k,\boldsymbol{\theta}} = \begin{bmatrix} \widetilde{\mathbf{J}}_k^{\mathbf{x}_k,\mathbf{x}_k} & \widetilde{\mathbf{J}}_k^{\mathbf{x}_k,\boldsymbol{\theta}} \\ \widetilde{\mathbf{J}}_k^{\boldsymbol{\theta},\mathbf{x}_k} & \widetilde{\mathbf{J}}_k^{\boldsymbol{\theta},\boldsymbol{\theta}} \end{bmatrix} \tag{10}$$

which obey the recursion (see the Appendix for details):

$$\begin{aligned} \widetilde{\mathbf{J}}_k^{\mathbf{x}_k,\mathbf{x}_k} &= \mathbf{D}_{k-1}^{22} - \left(\mathbf{D}_{k-1}^{12}\right)^T\left(\mathbf{D}_{k-1}^{11} + \widetilde{\mathbf{J}}_{k-1}^{\mathbf{x}_{k-1},\mathbf{x}_{k-1}}\right)^{-1}\mathbf{D}_{k-1}^{12} \\ \widetilde{\mathbf{J}}_k^{\mathbf{x}_k,\boldsymbol{\theta}} &= \mathbf{D}_{k-1}^{23} - \left(\mathbf{D}_{k-1}^{12}\right)^T\left(\mathbf{D}_{k-1}^{11} + \widetilde{\mathbf{J}}_{k-1}^{\mathbf{x}_{k-1},\mathbf{x}_{k-1}}\right)^{-1} \\ &\qquad \times\left(\mathbf{D}_{k-1}^{13} + \widetilde{\mathbf{J}}_{k-1}^{\mathbf{x}_{k-1},\boldsymbol{\theta}}\right) \\ \widetilde{\mathbf{J}}_k^{\boldsymbol{\theta},\boldsymbol{\theta}} &= \mathbf{D}_{k-1}^{33} + \mathbf{J}_{k-1}^{\boldsymbol{\theta},\boldsymbol{\theta}} - \left(\mathbf{D}_{k-1}^{13} + \widetilde{\mathbf{J}}_{k-1}^{\mathbf{x}_{k-1},\boldsymbol{\theta}}\right)^T \\ &\qquad \times\left(\mathbf{D}_{k-1}^{11} + \widetilde{\mathbf{J}}_{k-1}^{\mathbf{x}_{k-1},\mathbf{x}_{k-1}}\right)^{-1}\left(\mathbf{D}_{k-1}^{13} + \widetilde{\mathbf{J}}_{k-1}^{\mathbf{x}_{k-1},\boldsymbol{\theta}}\right) \end{aligned} \tag{11}$$

where:

$$\left|\begin{array}{l}\mathbf{D}_{k-1}^{11}=E_{\mathbf{x}_k,\mathbf{x}_{k-1}|\boldsymbol{\alpha},\boldsymbol{\beta}}\left[-\frac{\partial^2\ln p(\mathbf{x}_k|\mathbf{x}_{k-1},\boldsymbol{\alpha},\boldsymbol{\beta})}{\partial\mathbf{x}_{k-1}\partial\mathbf{x}_{k-1}^T}\right]\\ \mathbf{D}_{k-1}^{12}=E_{\mathbf{x}_k,\mathbf{x}_{k-1}|\boldsymbol{\alpha},\boldsymbol{\beta}}\left[-\frac{\partial^2\ln p((\mathbf{x}_k|\mathbf{x}_{k-1},\boldsymbol{\alpha},\boldsymbol{\beta}))}{\partial\mathbf{x}_{k-1}\partial\mathbf{x}_k^T}\right]\\ \mathbf{D}_{k-1}^{22}=E_{\mathbf{y}_k,\mathbf{x}_k|\boldsymbol{\theta}}\left[-\frac{\partial^2\ln p(\mathbf{y}_k|\mathbf{x}_k,\boldsymbol{\theta})}{\partial\mathbf{x}_k\partial\mathbf{x}_k^T}\right]\\ \qquad\qquad+E_{\mathbf{x}_k,\mathbf{x}_{k-1}|\boldsymbol{\alpha},\boldsymbol{\beta}}\left[-\frac{\partial^2\ln p((\mathbf{x}_k|\mathbf{x}_{k-1},\boldsymbol{\alpha},\boldsymbol{\beta}))}{\partial\mathbf{x}_k\partial\mathbf{x}_k^T}\right]\end{array}\right. \tag{12}$$

$$\left|\begin{array}{l}\mathbf{D}_{k-1}^{13}=E_{\mathbf{x}_k,\mathbf{x}_{k-1}|\boldsymbol{\alpha},\boldsymbol{\beta}}\left[-\frac{\partial^2\ln p((\mathbf{x}_k|\mathbf{x}_{k-1},\boldsymbol{\alpha},\boldsymbol{\beta}))}{\partial\mathbf{x}_{k-1}\partial\boldsymbol{\theta}^T}\right]\\ \mathbf{D}_{k-1}^{23}=E_{\mathbf{y}_k,\mathbf{x}_k|\boldsymbol{\theta}}\left[-\frac{\partial^2\ln p(\mathbf{y}_k|\mathbf{x}_k,\boldsymbol{\theta})}{\partial\mathbf{x}_k\partial\boldsymbol{\theta}^T}\right]\\ \qquad\qquad+E_{\mathbf{x}_k,\mathbf{x}_{k-1}|\boldsymbol{\alpha},\boldsymbol{\beta}}\left[-\frac{\partial^2\ln p((\mathbf{x}_k|\mathbf{x}_{k-1},\boldsymbol{\alpha},\boldsymbol{\beta}))}{\partial\mathbf{x}_k\partial\boldsymbol{\theta}^T}\right]\\ \mathbf{D}_{k-1}^{33}=E_{\mathbf{y}_k,\mathbf{x}_k|\boldsymbol{\theta}}\left[-\frac{\partial^2\ln p(\mathbf{y}_k|\mathbf{x}_k,\boldsymbol{\theta})}{\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}^T}\right]\\ \qquad\qquad+E_{\mathbf{x}_k,\mathbf{x}_{k-1}|\boldsymbol{\alpha},\boldsymbol{\beta}}\left[-\frac{\partial^2\ln p((\mathbf{x}_k|\mathbf{x}_{k-1},\boldsymbol{\alpha},\boldsymbol{\beta}))}{\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}^T}\right]\end{array}\right. \tag{13}$$

$$\widetilde{\mathbf{J}}_{\mathbf{x}_0,\boldsymbol{\theta}}=E_{\mathbf{x}_0|\boldsymbol{\theta}}\left[\frac{-\partial^2\ln p\left(\mathbf{x}_0|\boldsymbol{\alpha}\right)}{\partial\left(\mathbf{x}_0;\boldsymbol{\theta}\right)\partial\left(\mathbf{x}_0^T,\boldsymbol{\theta}^T\right)}\right] \tag{14}$$

Using definition (10) of $\widetilde{\mathbf{J}}_{\mathbf{x}_k,\boldsymbol{\theta}}$ and the associated recursion (11) only involve computations with matrix of dimension $(\max\{D,P\}\times\max\{D,P\})$. Therefore it seems adequate to name $\widetilde{\mathbf{J}}_{\mathbf{x}_k,\boldsymbol{\theta}}$ and $\widetilde{\mathbf{J}}_{\mathbf{x}_k,\boldsymbol{\theta}}^{-1}$ the "recursive" HFIM and HCRB (for discrete-time MDS), respectively. Some special cases of interest can be easily derived by updating the definitions of $\boldsymbol{\theta}$, $p\left(\mathbf{x}_k|\mathbf{x}_{k-1}\right)$, $p\left(\mathbf{y}_k|\mathbf{x}_k\right)$ and $p\left(\mathbf{x}_0|\boldsymbol{\alpha}\right)$ accordingly. For instance, if $\boldsymbol{\theta}=\varnothing$, i.e. there is no unknown deterministic parameter, then (11)(12) reduce to:

$$\left|\begin{array}{l}\widetilde{\mathbf{J}}_k^{\mathbf{x}_k,\mathbf{x}_k}=\mathbf{D}_{k-1}^{22}-\left(\mathbf{D}_{k-1}^{12}\right)^T\left(\mathbf{D}_{k-1}^{11}+\widetilde{\mathbf{J}}_{k-1}^{\mathbf{x}_{k-1},\mathbf{x}_{k-1}}\right)^{-1}\mathbf{D}_{k-1}^{12}\\ \mathbf{D}_{k-1}^{11}=E_{\mathbf{x}_k,\mathbf{x}_{k-1}}\left[\frac{-\partial^2\ln p((\mathbf{x}_k|\mathbf{x}_{k-1}))}{\partial\mathbf{x}_{k-1}\partial\mathbf{x}_{k-1}^T}\right]\\ \mathbf{D}_{k-1}^{12}=E_{\mathbf{x}_k,\mathbf{x}_{k-1}}\left[\frac{-\partial^2\ln p(\mathbf{x}_k|\mathbf{x}_{k-1})}{\partial\mathbf{x}_{k-1}\partial\mathbf{x}_k^T}\right]\\ \mathbf{D}_{k-1}^{22}=E_{\mathbf{y}_k,\mathbf{x}_k}\left[\frac{-\partial^2\ln p(\mathbf{y}_k|\mathbf{x}_k)}{\partial\mathbf{x}_k\partial\mathbf{x}_k^T}\right]\\ \qquad\qquad+E_{\mathbf{x}_k,\mathbf{x}_{k-1}}\left[\frac{-\partial^2\ln p(\mathbf{x}_k|\mathbf{x}_{k-1})}{\partial\mathbf{x}_k\partial\mathbf{x}_k^T}\right]\end{array}\right. \tag{15}$$

and $\widetilde{\mathbf{J}}_{\mathbf{x}_0,\boldsymbol{\theta}}\triangleq\mathbf{J}_{\mathbf{x}_0}$, which are (21-25) in [15]. If (3) reduces to:

$$\mathbf{x}_k=\mathbf{f}_{k-1}\left(\mathbf{x}_{k-1},\mathbf{w}_{k-1}\right),\ \mathbf{y}_k=\mathbf{h}_k\left(\mathbf{x}_k,\mathbf{v}_k,\boldsymbol{\lambda}\right),$$

with known pdfs $p\left(\mathbf{w}_k\right)$, $p\left(\mathbf{v}_k\right)$, $p\left(\mathbf{x}_0\right)$, then one only needs to set $\boldsymbol{\theta}=\boldsymbol{\lambda}$, $p\left(\mathbf{x}_k|\mathbf{x}_{k-1}\right)\triangleq p\left(\mathbf{x}_k|\mathbf{x}_{k-1}\right)$ and $\widetilde{\mathbf{J}}_{\mathbf{x}_0,\boldsymbol{\theta}}\triangleq\mathbf{J}_{\mathbf{x}_0}$, yielding: $\mathbf{D}_{k-1}^{13}=\mathbf{0}$, $\mathbf{D}_{k-1}^{23}=E_{\mathbf{y}_k,\mathbf{x}_k|\boldsymbol{\theta}}\left[-\frac{\partial^2\ln p(\mathbf{y}_k|\mathbf{x}_k,\boldsymbol{\theta})}{\partial\mathbf{x}_k\partial\boldsymbol{\theta}^T}\right]$, $\mathbf{D}_{k-1}^{33}=E_{\mathbf{y}_k,\mathbf{x}_k|\boldsymbol{\theta}}\left[-\frac{\partial^2\ln p(\mathbf{y}_k|\mathbf{x}_k,\boldsymbol{\theta})}{\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}^T}\right]$. If (3) reduces to:

$$\mathbf{x}_k=\mathbf{f}_{k-1}\left(\mathbf{x}_{k-1},\mathbf{w}_{k-1},\boldsymbol{\alpha}\right),\ \mathbf{y}_k=\mathbf{h}_k\left(\mathbf{x}_k,\mathbf{v}_k\right),$$

with known pdfs $p\left(\mathbf{w}_k\right)$, $p\left(\mathbf{v}_k\right)$, $p\left(\mathbf{x}_0\right)$, then one only needs to set $\boldsymbol{\theta}=\boldsymbol{\alpha}$, $p\left(\mathbf{x}_k|\mathbf{x}_{k-1}\right)\triangleq p\left(\mathbf{x}_k|\mathbf{x}_{k-1},\boldsymbol{\theta}\right)$ and $\widetilde{\mathbf{J}}_{\mathbf{x}_0,\boldsymbol{\theta}}\triangleq\mathbf{J}_{\mathbf{x}_0}$. Interestingly enough, the application considered in [6] addresses this case with measurement and state equations given by:

$$x_k=x_{k-1}+\alpha+w_{k-1},\ y_k=a_ke^{jx_k}+v_k,$$

where a_k are i.i.d. discrete random variable with $P\left(a_k=\pm1\right)=1/2$. Although authors in [6] were primarily interested in the direct computation of $\mathbf{J}_{\widetilde{\mathbf{x}}_{0:k},\boldsymbol{\theta}}^{-1}$ (6), a careful examination of the derivation reveals that $\left[\mathbf{H}^{-1}\right]_{K,K}$ in [6, (22)] is another possible recursive form for $\left[\widetilde{\mathbf{J}}_{x_K,\alpha}^{-1}\right]_{11}$ (10) allowed by the specific discrete-time MDS considered. Thus [6, (22)] provides an example of the behavior of the recursive HCRB (9) for a practical phase estimation problem.

III. Regularity conditions for the recursive HCRB

We discuss in this section the regularity condition (stated in the previous section) required for the existence and the use of the recursive HCRB (9) deriving from the usual form of the HCRB (1). This discussion clarifies some previous results on HCRB [2][6][13] and is also helpful to understand why, in most case, the posterior BCRB cannot be transformed into the recursive HCRB as it is misleadingly suggested in [15][16].

Let $\Lambda_{\boldsymbol{\theta}}=\left\{\mathbf{h}_d\in\mathbb{R}^D\mid\boldsymbol{\theta}+\mathbf{h}_d\in\Pi_d\right\}$, $\Lambda_{\mathbf{x}}=\left\{\mathbf{h}_r\in\mathbb{R}^P\mid\mathbf{x}+\mathbf{h}_r\in\Pi_r\right\}$ and $\Lambda=\Lambda_{\boldsymbol{\theta}}\times\Lambda_{\mathbf{x}}=\left\{\mathbf{h}\in\mathbb{R}^{D+P}\mid(\mathbf{x};\boldsymbol{\theta})+\mathbf{h}\in\Pi_d\times\Pi_r\right\}$. We assume that for any couple set $(S,T)\subset\Omega\times\Pi_r$ not empty, $\int_S\int_T p\left(\mathbf{y},\mathbf{x}|\boldsymbol{\theta}\right)d\mathbf{y}d\mathbf{x}$ exists. Then, for any L-dimensional real-valued vector $\mathbf{v}\left(\mathbf{y},\mathbf{x},\boldsymbol{\theta}\right)$ with finite second order moment, the covariance inequality principle yields [18]:

$$E_{\mathbf{y},\mathbf{x}|\boldsymbol{\theta}}\left[\mathbf{e}\left(\mathbf{y}\right)\mathbf{e}\left(\mathbf{y}\right)^T\right]\succeq\mathbf{C}\mathbf{V}^{-1}\mathbf{C}^T, \tag{16}$$

$$\mathbf{V}=E_{\mathbf{y},\mathbf{x}|\boldsymbol{\theta}}\left[\mathbf{v}\left(\mathbf{y},\mathbf{x},\boldsymbol{\theta}\right)\mathbf{v}\left(\mathbf{y},\mathbf{x},\boldsymbol{\theta}\right)^T\right], \tag{17}$$

$$\mathbf{C}=E_{\mathbf{y},\mathbf{x}|\boldsymbol{\theta}}\left[\mathbf{e}\left(\mathbf{y}\right)\mathbf{v}\left(\mathbf{y},\mathbf{x},\boldsymbol{\theta}\right)^T\right]. \tag{18}$$

Note that $\mathbf{C}$ (18) depends on the estimation scheme $\left(\widehat{\mathbf{x}}\left(\mathbf{y}\right);\widehat{\boldsymbol{\theta}}\left(\mathbf{y}\right)\right)$ in general; however, some judicious choices of $\mathbf{v}\left(\mathbf{y},\mathbf{x},\boldsymbol{\theta}\right)$ lead to various lower bounds [10][11][12]. Thus, the $\mathbf{v}_{CRB}\left(\mathbf{y},\mathbf{x},\boldsymbol{\theta}\right)$ leading to HCRB is a limiting form of the $\mathbf{v}\left(\mathbf{y},\mathbf{x},\boldsymbol{\theta}\right)$ leading to the HBB which may be derived from the McAulay-Seidman bound (MSB) [19]. The MSB is the usual approximation of the Barankin Bound (BB) [20], the greatest lower bound on the MSE on deterministic parameters $\boldsymbol{\theta}$, resulting from a discretization of the uniform unbiasedness definition[2]:

$$E_{\mathbf{y}|\boldsymbol{\theta}}\left[\widehat{\boldsymbol{\theta}}\left(\mathbf{y}\right)\right]=\int_{\Omega}\widehat{\boldsymbol{\theta}}\left(\mathbf{y}\right)p\left(\mathbf{y}|\boldsymbol{\theta}\right)d\mathbf{y}=\boldsymbol{\theta},\forall\boldsymbol{\theta}\in\Pi_d,$$

expressed for a subset of I test points [19][21][22]:

$$E_{\mathbf{y}|\boldsymbol{\theta}+\mathbf{h}_{di}}\left[\widehat{\boldsymbol{\theta}}\left(\mathbf{y}\right)\right]=\boldsymbol{\theta}+\mathbf{h}_{di},\ \boldsymbol{\theta}+\mathbf{h}_{di}\in\Pi_d,1\leq i\leq I. \tag{19}$$

Unfortunately, in hybrid estimation the MSB can hardly ever be computed to lower bound the MSE of unbiased estimators of $\widehat{\boldsymbol{\theta}}$ since a closed-form expression of $p\left(\mathbf{y}|\boldsymbol{\theta}\right)$ hardly ever exists. Interestingly enough, this major stumbling block can be bypassed for a general class of pdf. Let $1_A\left(\mathbf{y}\right)$ denote the indicator function of subset A of $\mathbb{R}^P$. Then, some easy integral calculuses show that, for any $\mathbf{h}_{ri}\in\mathbb{R}^P$ for which:

$$1_{\Pi_r}\left(\mathbf{x}+\mathbf{h}_{ri}\right)=1_{\Pi_r}\left(\mathbf{x}\right),\ \forall\mathbf{x}\in\mathbb{R}^P, \tag{20}$$

[2]For a straightforward extension to biased estimators see [21][22].

subject to (20), (19) can be recasted as:

$$E_{\mathbf{y},\mathbf{x}|\boldsymbol{\theta}}\left[\left(\widehat{\boldsymbol{\theta}}\left(\mathbf{y}\right)-\boldsymbol{\theta}\right)v\left(\mathbf{y},\boldsymbol{\theta},\mathbf{h}_{ri},\mathbf{h}_{di}\right)\right]=\mathbf{h}_{di}, \quad (21)$$

$$v\left(\mathbf{y},\boldsymbol{\theta},\mathbf{h}_r,\mathbf{h}_d\right)=\left|\begin{array}{l}\frac{p(\mathbf{y},\mathbf{x}+\mathbf{h}_r|\boldsymbol{\theta}+\mathbf{h}_d)}{p(\mathbf{y},\mathbf{x}|\boldsymbol{\theta})}-1,\ (\mathbf{x},\boldsymbol{\theta})\in\boldsymbol{\Theta},\\ 0,\text{ otherwise.}\end{array}\right.$$

Additionally, if $\mathbf{h}_{ri}$ also satisfies:

$$1_{\Pi_r}\left(\mathbf{x}-\mathbf{h}_{ri}\right)=1_{\Pi_r}\left(\mathbf{x}\right),\ \forall\mathbf{x}\in\mathbb{R}^P, \quad (22)$$

then, some other easy integral calculuses show that, subject to (20) and (22), any estimator $\widehat{\mathbf{x}}$ of $\mathbf{x}$ satisfies:

$$E_{\mathbf{y},\mathbf{x}|\boldsymbol{\theta}}\left[\left(\widehat{\mathbf{x}}\left(\mathbf{y}\right)-\mathbf{x}\right)v\left(\mathbf{y},\boldsymbol{\theta},\mathbf{h}_{ri},\mathbf{h}_{di}\right)\right]=\mathbf{h}_{ri}+E_{\mathbf{y},\mathbf{x}|\boldsymbol{\theta}+\mathbf{h}_{di}}\left[\widehat{\mathbf{x}}\left(\mathbf{y}\right)-\mathbf{x}\right]-E_{\mathbf{y},\mathbf{x}|\boldsymbol{\theta}}\left[\widehat{\mathbf{x}}\left(\mathbf{y}\right)-\mathbf{x}\right]. \quad (23)$$

Unlike what is stated in [2][6][13], first, (23) proves that for each class of estimates satisfying:

$$E_{\mathbf{y},\mathbf{x}|\boldsymbol{\theta}}\left[\left(\widehat{\mathbf{x}}\left(\mathbf{y}\right)-\mathbf{x};\widehat{\boldsymbol{\theta}}\left(\mathbf{y}\right)-\boldsymbol{\theta}\right)\right]=\left(\boldsymbol{\mu}\left(\boldsymbol{\theta}\right);\mathbf{0}\right),$$

a specific HBB (16) (and a specific HCRB) can be computed. The class of wide-sense unbiased estimates [13, (11)] is only the particular case where $\boldsymbol{\mu}\left(\boldsymbol{\theta}\right)=\mathbf{0}$. Second, the regularity condition (20)(22) required for the existence of the HBB only imposes on $1_{\Pi_r}\left(\mathbf{x}\right)$ to be of the following form:

$$1_{\Pi_r}\left(\mathbf{x}\right)=\left|\begin{array}{l}0\text{ if }\sum\limits_{\mathbf{h}_r\in A}\left(\sum\limits_{l\in\mathbb{Z}}1_{\Pi_r^0}\left(\mathbf{x}+l\mathbf{h}_r\right)\right)=0,\\ 1,\text{ otherwise,}\end{array}\right. \quad (24)$$

where A and Π_r^0 are subsets of $\mathbb{R}^P$, what means that Π_r may be a discrete subset of $\mathbb{R}^P$ or a subset of intervals of $\mathbb{R}^P$. Since the HCRB is the limiting case of the HBB obtained for:

$$\mathbf{v}\left(\mathbf{y},\mathbf{x},\boldsymbol{\theta}\right)=\left(v\left(\mathbf{y},\boldsymbol{\theta},\mathbf{u}_1h_{r1},\mathbf{0}\right),\ldots,v\left(\mathbf{y},\boldsymbol{\theta},\mathbf{u}_Ph_{rP},\mathbf{0}\right),\right.$$
$$\left.v\left(\mathbf{y},\boldsymbol{\theta},\mathbf{0},\mathbf{u}_{P+1}h_{d1}\right),\ldots,v\left(\mathbf{y},\boldsymbol{\theta},\mathbf{0},\mathbf{u}_{P+D}h_{dD}\right)\right)$$

where $\mathbf{u}_i$ is the ith column of the identity matrix, and by letting $\left(h_{r1},\ldots,h_{rP},h_{d1},\ldots,h_{dD}\right)$ be infinitesimally small, therefore (24) reduces to: $\forall\mathbf{x}\in\mathbb{R}^P, 1_{\Pi_r}\left(\mathbf{x}\right)=1$, that is $\Pi_r=\mathbb{R}^P$ (R1). Last, the simplest form of $\mathbf{C}$ (18), i.e. $\mathbf{C}=\mathbf{I}$, is obtained for the class of estimates satisfying (R3):

$$E_{\mathbf{y},\mathbf{x}|\boldsymbol{\theta}}\left[\left(\widehat{\mathbf{x}}\left(\mathbf{y}\right)-\mathbf{x};\widehat{\boldsymbol{\theta}}\left(\mathbf{y}\right)-\boldsymbol{\theta}\right)\right]=\left(\boldsymbol{\mu};\mathbf{0}\right), \quad (25)$$

leading to the usual form of the HCRB (16):

$$\mathbf{HCRB}_{\mathbf{x},\boldsymbol{\theta}}=E_{\mathbf{y},\mathbf{x}|\boldsymbol{\theta}}\left[\frac{\partial\ln p\left(\mathbf{y},\mathbf{x}|\boldsymbol{\theta}\right)}{\partial\left(\mathbf{x};\boldsymbol{\theta}\right)}\frac{\partial\ln p\left(\mathbf{y},\mathbf{x}|\boldsymbol{\theta}\right)}{\partial\left(\mathbf{x};\boldsymbol{\theta}\right)^T}\right]^{-1} \quad (26)$$

which contains elements with finite modulus provided that (R2): $E_{\mathbf{y},\mathbf{x}|\boldsymbol{\theta}}\left[\frac{\partial\ln p(\mathbf{y},\mathbf{x}|\boldsymbol{\theta})}{\partial x_p}^2\right], E_{\mathbf{y},\mathbf{x}|\boldsymbol{\theta}}\left[\frac{\partial\ln p(\mathbf{y},\mathbf{x}|\boldsymbol{\theta})}{\partial\theta_d}^2\right]<\infty$. Since the standard form of the HFIM (2)(6) [6][10][13] shares a common analytical form with the posterior BFIM [10]:

$$\mathbf{J}_{\overline{\mathbf{x}}_{0:k}}=E_{\overline{\mathbf{y}}_{1:k},\overline{\mathbf{x}}_{0:k}}\left[\frac{-\partial^2\ln p\left(\overline{\mathbf{y}}_{1:k},\overline{\mathbf{x}}_{0:k}\right)}{\partial\overline{\mathbf{x}}_{0:k}\partial\overline{\mathbf{x}}_{0:k}^T}\right],$$

one might think that the posterior BFIM becomes naturally a HFIM (6) when some random parameters are transformed into deterministic parameters by discarding their prior information as mentioned in [15] and [16]. For example, if we consider that $p\left(\mathbf{x}_0\right)$ is unknown, then setting $\mathbf{J}_{\mathbf{x}_0}=\mathbf{0}$ in (15) as suggested in [15][16], is equivalent to transform $p\left(\overline{\mathbf{y}}_{1:k},\overline{\mathbf{x}}_{0:k}\right)$ into:

$$p\left(\overline{\mathbf{y}}_{1:k},\overline{\mathbf{x}}_{1:k}|\mathbf{x}_0\right)=\prod_{l=1}^{k}p\left(\mathbf{y}_l|\mathbf{x}_l\right)p\left(\mathbf{x}_l|\mathbf{x}_{l-1}\right). \quad (27)$$

Even if the recursion (15) still holds when computed with the conditional pdf (27), the block matrices obtained are valid components of the recursive HFIM (10) if and only if (R1) is satisfied. Thus, if any of $p\left(\mathbf{x}_k|\mathbf{x}_{k-1}\right)$, $p\left(\mathbf{y}_k|\mathbf{x}_k\right)$ or $p\left(\mathbf{x}_0\right)$ have a support which is not, $\mathbb{R}^P$, $\mathbb{R}^N$ or $\mathbb{R}^P$, respectively [23], then $\widetilde{\mathbf{J}}_k^{\mathbf{x}_k,\mathbf{x}_k}$ in (15) is not a HFIM. Additionally, even if pdfs support are $\mathbb{R}^P$, $\mathbb{R}^N$ or $\mathbb{R}^P$, respectively, the HFIM obtained is valid only for estimates satisfying (R3) and are no longer valid for any realizable estimate as it is the case for the BFIM [23]. Therefore it is critical to understand that discarding prior information in the computation of the BCRB change thoroughly the estimation problem under consideration and that the correct rationale to address the computation of a HCRB for MDS problem is the one introduced.

IV. Appendix

Let $E\left[\ \right]\triangleq E_{\overline{\mathbf{y}}_{1:k},\overline{\mathbf{x}}_{0:k}|\boldsymbol{\theta}}\left[\ \right]$. First $\mathbf{J}_{\overline{\mathbf{x}}_{0:k},\boldsymbol{\theta}}$ (6) can be broken down as:

$$\left[\begin{array}{ccc}\mathbf{J}_k^{11}&\mathbf{J}_k^{12}&\mathbf{J}_k^{13}\\ \mathbf{J}_k^{21}&\mathbf{J}_k^{22}&\mathbf{J}_k^{23}\\ \mathbf{J}_k^{31}&\mathbf{J}_k^{32}&\mathbf{J}_k^{33}\end{array}\right],\ \mathbf{J}_k^{11}=E\left[-\frac{\partial^2\ln p(\overline{\mathbf{y}}_{1:k},\overline{\mathbf{x}}_{0:k}|\boldsymbol{\theta})}{\partial\overline{\mathbf{x}}_{0:k-1}\partial\overline{\mathbf{x}}_{0:k-1}^T}\right],$$
$$\mathbf{J}_k^{12}=E\left[-\frac{\partial^2\ln p(\overline{\mathbf{y}}_{1:k},\overline{\mathbf{x}}_{0:k}|\boldsymbol{\theta})}{\partial\overline{\mathbf{x}}_{0:k-1}\partial\mathbf{x}_k^T}\right],$$
$$\mathbf{J}_k^{22}=E\left[-\frac{\partial^2\ln p(\overline{\mathbf{y}}_{1:k},\overline{\mathbf{x}}_{0:k}|\boldsymbol{\theta})}{\partial\mathbf{x}_k\partial\mathbf{x}_k^T}\right],\mathbf{J}_k^{13}=E\left[-\frac{\partial^2\ln p(\overline{\mathbf{y}}_{1:k},\overline{\mathbf{x}}_{0:k}|\boldsymbol{\theta})}{\partial\overline{\mathbf{x}}_{0:k-1}\partial\boldsymbol{\theta}^T}\right],$$
$$\mathbf{J}_k^{23}=E\left[-\frac{\partial^2\ln p(\overline{\mathbf{y}}_{1:k},\overline{\mathbf{x}}_{0:k}|\boldsymbol{\theta})}{\partial\mathbf{x}_k\partial\boldsymbol{\theta}^T}\right],\mathbf{J}_k^{33}=E\left[-\frac{\partial^2\ln p(\overline{\mathbf{y}}_{1:k},\overline{\mathbf{x}}_{0:k}|\boldsymbol{\theta})}{\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}^T}\right].$$

Therefore, using block matrix inversion [24, p293]:

$$\begin{array}{l}\widetilde{\mathbf{J}}_k^{\mathbf{x}_k,\mathbf{x}_k}=\mathbf{J}_k^{22}-\mathbf{J}_k^{21}\left(\mathbf{J}_{k-1}^{11}\right)^{-1}\mathbf{J}_k^{12}\\ \widetilde{\mathbf{J}}_k^{\boldsymbol{\theta},\mathbf{x}_k}=\mathbf{J}_k^{23}-\mathbf{J}_k^{21}\left(\mathbf{J}_{k-1}^{11}\right)^{-1}\mathbf{J}_k^{13}\\ \widetilde{\mathbf{J}}_k^{\boldsymbol{\theta},\boldsymbol{\theta}}=\mathbf{J}_k^{33}-\mathbf{J}_k^{31}\left(\mathbf{J}_{k-1}^{11}\right)^{-1}\mathbf{J}_k^{13}\end{array} \quad (28)$$

Moreover, for MDS, (4) leads to:

$$p\left(\overline{\mathbf{y}}_{1:k},\overline{\mathbf{x}}_{0:k}|\boldsymbol{\theta}\right)=p\left(\mathbf{y}_k|\mathbf{x}_k,\boldsymbol{\theta}\right)p\left(\mathbf{x}_k|\mathbf{x}_{k-1},\boldsymbol{\alpha},\boldsymbol{\beta}\right)p\left(\overline{\mathbf{y}}_{1:k-1},\overline{\mathbf{x}}_{0:k-1}|\boldsymbol{\theta}\right)$$

yielding:

$$\mathbf{J}_k^{11}=\left[\begin{array}{cc}\mathbf{J}_{k-1}^{11}&\mathbf{J}_{k-1}^{12}\\ \mathbf{J}_{k-1}^{21}&\mathbf{J}_{k-1}^{22}+\mathbf{D}_{k-1}^{11}\end{array}\right],\ \mathbf{J}_k^{12}=\left[\begin{array}{c}\mathbf{0}\\ \mathbf{D}_{k-1}^{12}\end{array}\right],\ \mathbf{J}_k^{22}=\mathbf{D}_{k-1}^{22},$$
$$\mathbf{J}_k^{13}=\left[\begin{array}{c}\mathbf{J}_{k-1}^{13}\\ \mathbf{J}_{k-1}^{23}+\mathbf{D}_{k-1}^{13}\end{array}\right],\ \mathbf{J}_k^{23}=\mathbf{D}_{k-1}^{23},\ \mathbf{J}_k^{33}=\mathbf{J}_{k-1}^{33}+\mathbf{D}_{k-1}^{33}.$$

where $\mathbf{D}_{k-1}^{ij}$ are given by (12)(13). Second, using once again block matrix inversion:

$$\left(\mathbf{J}_k^{11}\right)^{-1}=\left[\begin{array}{cc}\boldsymbol{\Phi}_{k-1}&-\boldsymbol{\Gamma}_{k-1}\boldsymbol{\Delta}_{k-1}^{-1}\\ -\boldsymbol{\Delta}_{k-1}^{-1}\boldsymbol{\Gamma}_{k-1}^T&\boldsymbol{\Delta}_{k-1}^{-1}\end{array}\right],$$
$$\boldsymbol{\Delta}_{k-1}=\mathbf{D}_{k-1}^{11}+\mathbf{J}_{k-1}^{22}-\mathbf{J}_{k-1}^{21}\left(\mathbf{J}_{k-1}^{11}\right)^{-1}\mathbf{J}_{k-1}^{12},$$
$$\boldsymbol{\Gamma}_{k-1}=\left(\mathbf{J}_{k-1}^{11}\right)^{-1}\mathbf{J}_{k-1}^{12},$$
$$\boldsymbol{\Phi}_{k-1}=\left(\mathbf{J}_{k-1}^{11}\right)^{-1}+\boldsymbol{\Gamma}_{k-1}\boldsymbol{\Delta}_{k-1}^{-1}\boldsymbol{\Gamma}_{k-1}^T$$

Finally, by noting that $\boldsymbol{\Delta}_{k-1}=\mathbf{D}_{k-1}^{11}+\widetilde{\mathbf{J}}_{k-1}^{\mathbf{x}_{k-1},\mathbf{x}_{k-1}}$, a few additional lines of calculus allows to show that equivalent forms of $\widetilde{\mathbf{J}}_k^{\mathbf{x}_k,\mathbf{x}_k}$, $\widetilde{\mathbf{J}}_k^{\boldsymbol{\theta},\boldsymbol{\theta}}$and $\widetilde{\mathbf{J}}_k^{\boldsymbol{\theta},\mathbf{x}_k}$ in (28) are given by (11).

REFERENCES

[1] Y. Rockah and P. Schultheiss, "Array shape calibration using sources in unknown locations–part I: Far-field sources," IEEE Transactions on Acoustics, Speech, and Signal Processing, vol. 35, no. 3, pp. 286-299, Mar. 1987.

[2] I. Reuven and H. Messer, "A Barankin-type lower bound on the estimation error of a hybrid parameter vector," IEEE Transactions on Information Theory, vol. 43, no. 3, pp. 1084-1093, May 1997.

[3] H. L. Van Trees, *Detection, Estimation and Modulation theory: Optimum Array Processing*. New-York, NY, USA: John Wiley & Sons, Mar. 2002, vol. 4.

[4] P. Tichavsk´y and K. Wong, "Quasi-fluid-mecanics based quasi-Bayesian Cramér-Rao bounds for towed-array direction finding," IEEE Transactions on Signal Processing, vol. 52, no. 1, pp. 36-47, Jan. 2004.

[5] S. Buzzi, M. Lops, and S. Sardellitti, "Further result on Cramér-Rao bounds for parameter estimation in long-code DS/CDMA systems," IEEE Transactions on Signal Processing, vol. 53, no. 3, pp. 1216-1221, Mar. 2005.

[6] S. Bay, B. Geller, A. Renaux, J.-P. Barbot, and J.-M. Brossier, "On the hybrid Cramér-Rao bound and its application to dynamical phase estimation," IEEE Signal Processing Letters, vol. 15, pp. 453-456, 2008.

[7] K. Todros and J. Tabrikian, "Hybrid lower bound via compression of the sampled CLR function," in Proc. of IEEE Workshop on Statistical Signal Processing (SSP), Cardiff, Wales, UK, Aug. 2009, pp. 602-605.

[8] J. Yang, B. Geller, and S. Bay, "Bayesian and hybrid Cramér-Rao bounds for the carrier recovery under dynamic phase uncertain channels," IEEE Transactions on Signal Processing, vol. 59, no. 2, pp. 667-680, 2011.

[9] J. Vila-Valls, L. Ros, and J. M. Brossier, "Joint oversampled carrier and time-delay synchronization in digital communications with large excess bandwidth," ELSEVIER Signal Processing, vol. 92, no. 1, pp. 76-88, Jan. 2012.

[10] H. L. Van Trees and K. L. Bell, Eds., Bayesian Bounds for Parameter Estimation and Nonlinear Filtering/Tracking. New-York, NY, USA: Wiley/IEEE Press, Sep. 2007.

[11] C. Ren, J. Galy, E. Chaumette, P. Larzabal, and A. Renaux, "Hybrid lower bound on the MSE based on the Barankin and Weiss-Weinstein bounds," in Proc. of IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP), Vancouver, Canada, May 2013, pp. 5534-5538.

[12] C. Ren, J. Galy, E. Chaumette, P. Larzabal, and A. Renaux, "A Ziv-Zakaï type bound for hybrid parameter estimation," in Proc. of IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP), Florence, Italy, May 2014, pp. 4663-4667.

[13] Y. Noam and H. Messer, "Notes on the tightness of the hybrid Cramér-Rao lower bound," IEEE Transactions on Signal Processing, vol. 57, no. 6, pp. 2074–2084, 2009.

[14] D. Simon, *Optimal State Estimation: Kalman, H-infinity, and Nonlinear Approaches*, Wiley InterScience, 2006

[15] P. Tichavsky, C. Muravchik and A. Nehorai, "Posterior Cramér-Rao bounds for discrete-time nonlinear Filtering", IEEE Transactions on Signal Processing, vol. 46, no. 5, pp. 1386-1396, 1998

[16] M. Simandl, J. Kralovec and P. Tichavsky, "Filtering, predictive, and smoothing Cramér-Rao bounds for discrete-time nonlinear dynamic systems,", Automatica, vol 37, pp. 1703-1716, 2001

[17] B.Z. Bobrovsky, E. Mayer-Wolf and M. Zakai, "Some Classes of Global Cramer-Rao Bounds", The Annals of Statistics, 15(4): 1421-1438, 1987

[18] E. L. Lehmann and G. Casella, *Theory of Point Estimation (2nd ed.)*. Springer, 1998

[19] R. McAulay, L.P. Seidman, "A useful form of the Barankin Lower Bound and its application to PPM threshold analysis", IEEE Trans. on IT, 15(2): 273-279, 1969

[20] E.W. Barankin, "Locally best unbiased estimates", Ann. Math. Stat., 20(4): 477-501, 1949

[21] 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): 5064-5082, 2010

[22] 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

[23] E. Weinstein and A. J. Weiss, "A general class of lower bounds in parameter estimation," IEEE Trans. Inf. Theory, vol. IT-34, no. 2, pp. 338–342, Mar. 1988.

[24] G.A.F. Seber, *Matrix Handbook forr Statisticians*, Wiley Series in Probability and Statistics, 2008