On the Hybrid Cramér Rao Bound and Its Application to Dynamical Phase Estimation

Stéphanie Bay, Benoit Geller, Alexandre Renaux, *Member, IEEE*, Jean-Pierre Barbot, and Jean-Marc Brossier

***Abstract*—This letter deals with the Cramér–Rao bound for the estimation of a hybrid vector with both random and deterministic parameters. We point out the specificity of the case when the deterministic and the random vectors of parameters are statistically dependent. The relevance of this expression is illustrated by studying a practical phase estimation problem in a non-data-aided communication context.**

***Index Terms*—Cramér–Rao bounds, synchronization parameters estimation.**

I. Introduction

A natural problematic when designing an estimator is the evaluation of its performance. Lower bounds on the mean square error (MSE) mainly answer this question and the well-known Cramér–Rao bound (CRB) is widely used by the signal processing community. Depending on assumptions on the parameters, the CRB has different expressions. When the vector of parameters is assumed to be deterministic, we obtain the standard CRB and when the vector of parameters is assumed to be random with an *a priori* probability density function (pdf), we obtain the so-called Bayesian CRB [1].

At the end of the 1980s, an extension combining both the standard and the Bayesian CRBs was proposed [2]. Indeed, in some practical scenarios, it is natural to represent the parameter vector by a deterministic part and by a random part. This bound has thus been called the hybrid CRB (HCRB). Until now, results available in the literature essentially focused on the case where the deterministic part and the random part of the parameter vector are assumed to be statistically independent (see, e.g., [2, eq. (5)], [3, eq. (13)] and [4, eq. (13)]). To the best of our knowledge, a closed-form expression of the HCRB with a statistical dependence between the deterministic and the random parameters has never been reported in the literature. The goal of this letter is then twofold. First, in Section II, we remind the structure of the HCRB and we point out the specificity of the case when the deterministic part and the random part of the parameter vector are statistically dependent. Second, in Section III, motivated by this analysis, we give a closed-form expression of the proposed bound in the practical case of a dynamical phase subject to a linear drift in a non-data-aided communication context.

Manuscript received December 13, 2007; revised February 10, 2008. This work was supported in part by the French ANR (Agence Nationale de la Recherche), LURGA project and in part by the European network of excellence NEWCOM++. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Xiaoli Ma.

S. Bay and J.-P. Barbot are with the Ecole Normale Supérieure de Cachan, SATIE Laboratory, 94235 Cachan, France (e-mail: bay@satie.ens-cachan.fr; barbot@satie.ens-cachan.fr).

B. Geller is with Ecole Nationale Supérieure des Techniques Avancées, Laboratory of Electronics and Computer Engineering (LEI), 75015 Paris, France (e-mail: geller@ensta.fr).

A. Renaux is with University Paris-Sud 11, Laboratory of Signals and Systems, Supélec, 91192 Gif-sur-Yvette cedex, France (e-mail: renaux@lss.supelec.fr).

J.-M. Brossier is with Grenoble Institute of Technology, GIPSA Laboratory, BP 46 38402 Saint Martin d'Hères cedex, France (e-mail: jean-marc.brossier@gipsa-lab.inpg.fr).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/LSP.2008.921461

II. Hybrid Cramér-Rao Bound

A. Background

Let $\boldsymbol{\mu} = \left(\boldsymbol{\mu}_r^T \ \boldsymbol{\mu}_d^T\right)^T \in \mathbb{R}^n$ be the parameter vector that we have to estimate. This vector is split into two sub-vectors $\boldsymbol{\mu}_d$ and $\boldsymbol{\mu}_r$, where $\boldsymbol{\mu}_d$ is assumed to be a $(n-m)\times 1$ deterministic vector and $\boldsymbol{\mu}_r$ is assumed to be a $m \times 1$ random vector with an *a priori* pdf $p(\boldsymbol{\mu}_r)$. The true value of $\boldsymbol{\mu}_d$ will be denoted $\boldsymbol{\mu}_d^\star$. We consider $\hat{\boldsymbol{\mu}}(\mathbf{y})$ as an estimator of $\boldsymbol{\mu}$, where $\mathbf{y}$ is the observation vector. The HCRB satisfies the following inequality on the MSE:

$$\mathbb{E}_{\mathbf{y},\boldsymbol{\mu}_r|\boldsymbol{\mu}_d^\star}\left[\left(\hat{\boldsymbol{\mu}}(\mathbf{y})-\boldsymbol{\mu}\right)\left(\hat{\boldsymbol{\mu}}(\mathbf{y})-\boldsymbol{\mu}\right)^T|_{\boldsymbol{\mu}_d^\star}\right] \geq \mathbf{H}^{-1}\left(\boldsymbol{\mu}_d^\star\right) \quad (1)$$

where $\mathbf{H}\left(\boldsymbol{\mu}_d^\star\right) \in \mathbb{R}^{n\times n}$ is the so-called Hybrid Information Matrix (HIM) defined as [2]

$$\mathbf{H}\left(\boldsymbol{\mu}_d^\star\right) = \mathbb{E}_{\mathbf{y},\boldsymbol{\mu}_r|\boldsymbol{\mu}_d^\star}\left[-\Delta_{\boldsymbol{\mu}}^{\boldsymbol{\mu}}\log p(\mathbf{y},\boldsymbol{\mu}_r|\,\boldsymbol{\mu}_d)\Big|_{\boldsymbol{\mu}_d^\star}\right] \quad (2)$$

where $\left[\Delta_{\boldsymbol{\eta}}^{\boldsymbol{\nu}}\right]_{k,l} = \partial^2/\partial\left[\boldsymbol{\eta}\right]_k\partial\left[\boldsymbol{\nu}\right]_l$.

When the deterministic and the random parts of the parameter vector are assumed to be independent, and after some algebraic manipulations, the HIM can be rewritten as (see [3, eq. (18)])

$$\mathbf{H}\left(\boldsymbol{\mu}_d^\star\right) = \mathbb{E}_{\boldsymbol{\mu}_r}\left[\mathbf{F}(\boldsymbol{\mu}_d^\star,\boldsymbol{\mu}_r)\right] + \begin{pmatrix} \mathbb{E}_{\boldsymbol{\mu}_r}\left[-\Delta_{\boldsymbol{\mu}_r}^{\boldsymbol{\mu}_r}\log p\left(\boldsymbol{\mu}_r\right)\right] & \mathbf{0}_{m\times(n-m)} \\ \mathbf{0}_{(n-m)\times m} & \mathbf{0}_{(n-m)\times(n-m)} \end{pmatrix} \quad (3)$$

where

$$\mathbf{F}(\boldsymbol{\mu}_d^\star,\boldsymbol{\mu}_r) = \mathbb{E}_{\mathbf{y}|\boldsymbol{\mu}_d^\star,\boldsymbol{\mu}_r}\left[-\Delta_{\boldsymbol{\mu}}^{\boldsymbol{\mu}}\log p(\mathbf{y}|\,\boldsymbol{\mu}_d,\boldsymbol{\mu}_r)\Big|_{\boldsymbol{\mu}_d^\star}\right]. \quad (4)$$

With this aforementioned structure, it is straightforward to reobtain the standard and the Bayesian CRBs. Indeed, if $\boldsymbol{\mu}=\boldsymbol{\mu}_d$, we have

$$\mathbf{H}^{-1}\left(\boldsymbol{\mu}_d^\star\right) = \left(\mathbb{E}_{\mathbf{y}|\boldsymbol{\mu}_d^\star}\left[-\Delta_{\boldsymbol{\mu}_d}^{\boldsymbol{\mu}_d}\log p(\mathbf{y}|\,\boldsymbol{\mu}_d)\Big|_{\boldsymbol{\mu}_d^\star}\right]\right)^{-1} \quad (5)$$

which is the standard CRB, and, if $\boldsymbol{\mu} = \boldsymbol{\mu}_r$, we have

$$\mathbf{H}^{-1} = \left(\mathbb{E}_{\mathbf{y},\boldsymbol{\mu}_r}\left[-\Delta_{\boldsymbol{\mu}_r}^{\boldsymbol{\mu}_r}\log p(\mathbf{y}|\,\boldsymbol{\mu}_r)\right] + \mathbb{E}_{\boldsymbol{\mu}_r}\left[-\Delta_{\boldsymbol{\mu}_r}^{\boldsymbol{\mu}_r}\log p\left(\boldsymbol{\mu}_r\right)\right]\right)^{-1} \quad (6)$$

which is the Bayesian CRB.

B. Extension When $\boldsymbol{\mu}_r$ and $\boldsymbol{\mu}_d$ Are Statistically Dependent

We now assume a possible statistical dependence between $\boldsymbol{\mu}_r$ and $\boldsymbol{\mu}_d$. In other words, $\boldsymbol{\mu}_r$ is now assumed to be a $m \times 1$ random vector with an *a priori* pdf $p(\boldsymbol{\mu}_r | \boldsymbol{\mu}_d^\star) \neq p(\boldsymbol{\mu}_r)$.

Based on the HIM definition given by (2) and expending the log-likelihood as $\log p(\mathbf{y}, \boldsymbol{\mu}_r | \boldsymbol{\mu}_d^\star) = \log p(\mathbf{y} | \boldsymbol{\mu}_d^\star, \boldsymbol{\mu}_r) + \log p(\boldsymbol{\mu}_r | \boldsymbol{\mu}_d^\star)$, we obtain the following HIM:

$$\mathbf{H}(\boldsymbol{\mu}_d^\star) = \mathbb{E}_{\boldsymbol{\mu}_r|\boldsymbol{\mu}_d^\star}[\mathbf{F}(\boldsymbol{\mu}_d^\star, \boldsymbol{\mu}_r)] + \mathbb{E}_{\boldsymbol{\mu}_r|\boldsymbol{\mu}_d^\star}\left[- \Delta_{\boldsymbol{\mu}}^{\boldsymbol{\mu}} \log p(\boldsymbol{\mu}_r | \boldsymbol{\mu}_d)\Big|_{\boldsymbol{\mu}_d^\star}\right] \quad (7)$$

where $\mathbf{F}(\boldsymbol{\mu}_d^\star, \boldsymbol{\mu}_r)$ is given by (4).

In order to explicitly show the modification in comparison with the HIM given by (3), $\mathbf{H}(\boldsymbol{\mu}_d^\star)$ can be rewritten as (8) at the bottom of the page.

Obviously, if we assume $p(\boldsymbol{\mu}_r | \boldsymbol{\mu}_d) = p(\boldsymbol{\mu}_r)$ in this expression, we straightforwardly reobtain (3).

Based on this structure, one now has to prove that there is still an inequality, i.e., a lower bound on the MSE

$$\mathbb{E}_{\mathbf{y},\boldsymbol{\mu}_r|\boldsymbol{\mu}_d^\star}\left[(\hat{\boldsymbol{\mu}}(\mathbf{y}) - \boldsymbol{\mu})(\hat{\boldsymbol{\mu}}(\mathbf{y}) - \boldsymbol{\mu})^T\Big|_{\boldsymbol{\mu}_d^\star}\right] \geq \mathbf{H}^{-1}(\boldsymbol{\mu}_d^\star) \quad (9)$$

when $\mathbf{H}(\boldsymbol{\mu}_d^\star)$ is given by (8).

Proof: Following the idea of [3] to prove the inequality (1), one defines a vector $\mathbf{h}$ such that $\mathbf{h} = \begin{pmatrix} \nabla_{\boldsymbol{\mu}} \log p(\mathbf{y}, \boldsymbol{\mu}_r | \boldsymbol{\mu}_d)|_{\boldsymbol{\mu}_d^\star} \\ \hat{\boldsymbol{\mu}}(\mathbf{y}) - \boldsymbol{\mu}|_{\boldsymbol{\mu}_d^\star} \end{pmatrix}$, where $\nabla_{\boldsymbol{\mu}} = (\partial/\partial[\boldsymbol{\mu}]_1 \ \ldots \ \partial/\partial[\boldsymbol{\mu}]_n)^T$.

Consequently, the nonnegative definite matrix $\mathbf{G}(\boldsymbol{\mu}_d^\star) = \mathbb{E}_{\mathbf{y},\boldsymbol{\mu}_r|\boldsymbol{\mu}_d^\star}[\mathbf{h}\,\mathbf{h}^T]$ can be decomposed as the following block matrix: $\mathbf{G}(\boldsymbol{\mu}_d^\star) = \begin{pmatrix} \mathbf{H}(\boldsymbol{\mu}_d^\star) & \mathbf{L}(\boldsymbol{\mu}_d^\star) \\ \mathbf{L}^T(\boldsymbol{\mu}_d^\star) & \mathbf{R}(\boldsymbol{\mu}_d^\star) \end{pmatrix}$, where $\mathbf{R}(\boldsymbol{\mu}_d^\star)$ is the covariance matrix of $\hat{\boldsymbol{\mu}}(\mathbf{y})$, i.e., $\mathbf{R}(\boldsymbol{\mu}_d^\star) = \mathbb{E}_{\mathbf{y},\boldsymbol{\mu}_r|\boldsymbol{\mu}_d^\star}\left[(\hat{\boldsymbol{\mu}}(\mathbf{y}) - \boldsymbol{\mu})(\hat{\boldsymbol{\mu}}(\mathbf{y}) - \boldsymbol{\mu})^T\Big|_{\boldsymbol{\mu}_d^\star}\right]$ and, where $\mathbf{L}(\boldsymbol{\mu}_d^\star)$ is given by$\mathbf{L}(\boldsymbol{\mu}_d^\star) = \mathbb{E}_{\mathbf{y},\boldsymbol{\mu}_r|\boldsymbol{\mu}_d^\star} \times \left[\nabla_{\boldsymbol{\mu}} \log p(\mathbf{y}, \boldsymbol{\mu}_r | \boldsymbol{\mu}_d)|_{\boldsymbol{\mu}_d^\star}\left(\hat{\boldsymbol{\mu}}(\mathbf{y}) - \boldsymbol{\mu}|_{\boldsymbol{\mu}_d^\star}\right)^T\right]$.

Since $\mathbf{G}(\boldsymbol{\mu}_d^\star) \geq 0$, its Schur complement satisfies$\mathbf{R}(\boldsymbol{\mu}_d^\star) \geq \mathbf{L}^T(\boldsymbol{\mu}_d^\star)\mathbf{H}^{-1}(\boldsymbol{\mu}_d^\star)\mathbf{L}(\boldsymbol{\mu}_d^\star)$. ■

It is straightforward to show that, for an unbiased estimator w.r.t. the pdf $p(\mathbf{y}, \boldsymbol{\mu}_r | \boldsymbol{\mu}_d^\star)$, $\mathbf{L}(\boldsymbol{\mu}_d^\star) = \mathbf{I}_{n \times n}$.

Consequently, the inequality (9) is proved and $\mathbf{H}^{-1}(\boldsymbol{\mu}_d^\star)$ is a lower bound on the MSE.

III. HCRB for a Dynamical Phase Estimation Problem

In [4], we have proposed a closed-form expression of the Bayesian CRB for the estimation of the phase offset for a BPSK transmission in a non-data-aided context. In this section, we extend these previous results by providing a closed-form expression of the HCRB for the estimation of the phase offset and also of the linear drift.

A. Observation and State Models

We consider a linearly modulated signal, obtained by applying to a square-root Nyquist transmit filter an unknown symbol sequence $\mathbf{a} = (a_1 \cdots a_K)^T$ taken from a unit energy BPSK constellation. The signal is transmitted over an additive white Gaussian noise channel. The output signal is sampled at the symbol rate which yields to the observations

$$y_k = a_k e^{j\theta_k} + n_k \text{ with } k = 1 \ldots K \quad (10)$$

where $\{n_k\}$ is a sequence of i.i.d., circular, zero-mean complex Gaussian noise variables with variance σ_n^2. We consider that the system operates in a non-data-aided synchronization mode, i.e., the transmitted symbols are i.i.d. with $P_r(a_k = \pm 1) = 1/2$.

In practice, several sources of distortions affect the phase. An efficient model representing these effects is the so-called Brownian phase with a linear drift widely studied in the literature. The Brownian phase model with a linear drift is given as follows:

$$\theta_k = \theta_{k-1} + \xi + w_k \text{ with } k = 2 \ldots K \quad (11)$$

where, for any index k, $\{\theta_k\}$ is the sequence of phases to be estimated, ξ represents the deterministic unknown linear drift with true value $\xi^\star$, and where $\{w_k\}$ is an i.i.d. sequence of centered Gaussian random variables with known variance σ_w^2.

The parameter vector of interest is then made up of both random and deterministic parameters $\boldsymbol{\mu} = (\boldsymbol{\mu}_r^T \ \mu_d)^T$, where $\boldsymbol{\mu}_r = \boldsymbol{\theta} = (\theta_1 \ldots \theta_K)^T$ and $\mu_d = \xi$. Moreover, from (16), it is clear that $p(\boldsymbol{\theta} | \xi^\star) \neq p(\boldsymbol{\theta})$.

B. Derivation of the HCRB

For notational convenience, we drop the dependence of the different matrices on $\mu_d^\star = \xi^\star$ in the remainder of this letter. From (8), the HIM $\mathbf{H}$ can be rewritten into a block matrix $\mathbf{H} = \begin{pmatrix} \mathbf{H}_{11} & \mathbf{h}_{12} \\ \mathbf{h}_{21} & H_{22} \end{pmatrix}$, where we see (12), shown at the bottom of the next page.

These blocks only depend on the log-likelihoods $\log p(\mathbf{y} | \boldsymbol{\theta}, \xi^\star)$ and $\log p(\boldsymbol{\theta} | \xi^\star)$. Let us set $\mathbf{y} = (y_1 \cdots y_K)^T$ and assume that the initial phase θ_1 does not depend on ξ, i.e., $p(\theta_1 | \xi^\star) = p(\theta_1)$. Using (10) and (11), i.e., the Gaussian nature of the noise and the equiprobability of the symbols, one has to see (13), shown at the bottom of the page.

- Expression of $\mathbf{H}_{11}$: assuming that we have no prior knowledge, i.e., $\mathbb{E}_{\theta_1}\left[\Delta_{\theta_1}^{\theta_1} \log p(\theta_1)\right] = 0$, it is shown in [4] (due

$$\mathbf{H}(\boldsymbol{\mu}_d^\star) = \mathbb{E}_{\boldsymbol{\mu}_r|\boldsymbol{\mu}_d^\star}[\mathbf{F}(\boldsymbol{\mu}_d^\star, \boldsymbol{\mu}_r)] + \mathbb{E}_{\boldsymbol{\mu}_r|\boldsymbol{\mu}_d^\star} \times \left[\begin{pmatrix} -\Delta_{\boldsymbol{\mu}_r}^{\boldsymbol{\mu}_r} \log p(\boldsymbol{\mu}_r | \boldsymbol{\mu}_d^\star) & -\Delta_{\boldsymbol{\mu}_d}^{\boldsymbol{\mu}_r} \log p(\boldsymbol{\mu}_r | \boldsymbol{\mu}_d)\Big|_{\boldsymbol{\mu}_d^\star} \\ \left(-\Delta_{\boldsymbol{\mu}_d}^{\boldsymbol{\mu}_r} \log p(\boldsymbol{\mu}_r | \boldsymbol{\mu}_d)\Big|_{\boldsymbol{\mu}_d^\star}\right)^T & -\Delta_{\boldsymbol{\mu}_d}^{\boldsymbol{\mu}_d} \log p(\boldsymbol{\mu}_r | \boldsymbol{\mu}_d)\Big|_{\boldsymbol{\mu}_d^\star} \end{pmatrix}\right] \quad (8)$$

to the order one Markov structure exhibited by (11)) that $\mathbf{H}_{11}$ takes the following tridiagonal structure:

$$\mathbf{H}_{11} = b \begin{pmatrix} A+1 & 1 & 0 & \dots & 0 \\ 1 & A & 1 & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & 1 & A & 1 \\ 0 & \dots & 0 & 1 & A+1 \end{pmatrix} \quad (14)$$

where $b = -1/\sigma_w^2$, and where $A = -\sigma_w^2 J_D - 2$ with $J_D = \mathbb{E}_{\mathbf{y},\boldsymbol{\theta}|\xi^\star}\left[-\Delta_{\theta_k}^{\theta_k} \log p(y_k|\theta_k, \xi^\star)\right]$.

- Expression of $\mathbf{h}_{12}$: since, from (18), $\log p(\mathbf{y}|\boldsymbol{\theta}, \xi^\star)$ is independent of $\xi^\star$, $\Delta_\theta^\xi \log p(\mathbf{y}|\boldsymbol{\theta}, \xi)\Big|_{\xi^\star} = 0$. Consequently

$$\mathbf{h}_{12} = \mathbb{E}_{\boldsymbol{\theta}|\xi^\star}\left[-\Delta_{\boldsymbol{\theta}}^\xi \log p(\boldsymbol{\theta}|\xi)\Big|_{\xi^\star}\right]. \quad (15)$$

Using the state model, we have

$$\begin{cases} \Delta_\xi^{\theta_1} \log p(\boldsymbol{\theta}|\xi)\Big|_{\xi^\star} = -\frac{1}{\sigma_w^2} \\ \Delta_\xi^{\theta_K} \log p(\boldsymbol{\theta}|\xi)\Big|_{\xi^\star} = \frac{1}{\sigma_w^2} \\ \Delta_\xi^{\theta_k} \log p(\boldsymbol{\theta}|\xi)\Big|_{\xi^\star} = 0 \text{ for } k \in \{2, \dots, K-1\}. \end{cases}$$

Applying the expectation operator $\mathbb{E}_{\boldsymbol{\theta}|\xi^\star}[.]$, we obtain

$$\mathbf{h}_{12} = \begin{pmatrix} \frac{1}{\sigma_w^2} & \mathbf{0}_{1\times K-2} & -\frac{1}{\sigma_w^2} \end{pmatrix}^T. \quad (16)$$

- Expression of H_{22}: since, from (13), $\log p(\mathbf{y}|\boldsymbol{\theta}, \xi^\star)$ is independent of $\xi^\star$, $\Delta_\xi^\xi \log p(\mathbf{y}|\boldsymbol{\theta}, \xi)\Big|_{\xi^\star} = 0$. Consequently

$$H_{22} = \mathbb{E}_{\boldsymbol{\theta}|\xi^\star}\left[-\Delta_\xi^\xi \log p(\boldsymbol{\theta}|\xi)\Big|_{\xi^\star}\right] = \frac{K-1}{\sigma_w^2}. \quad (17)$$

- **Expression of the HCRB**: we now give the expression of $\mathbf{H}^{-1}$ which bounds the MSE. Thanks to the block-matrix inversion formula, we have

$$\mathbf{H}^{-1} = \begin{pmatrix} \mathbf{H}_{11}^{-1} + \mathbf{V}_K & -\frac{1}{\lambda}\mathbf{H}_{11}^{-1}\mathbf{h}_{12} \\ -\frac{1}{\lambda}\mathbf{h}_{12}^T\mathbf{H}_{11}^{-1} & \frac{1}{\lambda} \end{pmatrix} \quad (18)$$

where $\lambda = K - 1/\sigma_w^2 - \mathbf{h}_{12}^T\mathbf{H}_{11}^{-1}\mathbf{h}_{12}$ and $\mathbf{V}_K = 1/\lambda \mathbf{H}_{11}^{-1}\mathbf{h}_{12}\mathbf{h}_{12}^T\mathbf{H}_{11}^{-1}$.

We start to compute λ corresponding to the inverse of the minimal bound on the MSE of ξ. Due to the particular structure of matrices $\mathbf{H}_{11}$ and $\mathbf{h}_{12}$ (14), (16), we obtain $\lambda = \frac{K-1}{\sigma_w^2} - \frac{2}{\sigma_w^4}\left(\left[\mathbf{H}_{11}^{-1}\right]_{1,1} - \left[\mathbf{H}_{11}^{-1}\right]_{1,K}\right)$.

From (14), thanks to the cofactor expression in the matrix inversion formula, we have for any index k, $\left[\mathbf{H}_{11}^{-1}\right]_{1,k} = b^{k-1}/|\mathbf{H}_{11}|\,(d_{K-k} + b\, d_{K-k-1})$, where d_k is the determinant of a $k \times k$ matrix $\mathbf{D}_k$. equal to the matrix of (14) without the plus one on each corner.

The sequence $\{d_k\}$ satisfies the following recursive equation: $d_k = A\, b\, d_{k-1} - b^2\, d_{k-2}$ with $d_0 = 1$ and $d_1 = bA$. d_k can thus be written as $d_k = \rho_1 (r_1)^k + \rho_2 (r_2)^k$, where r_1, r_2, ρ_1, and ρ_2 are given by

$$\begin{cases} r_1 = \frac{b}{2}\left(A + \sqrt{A^2-4}\right), & r_2 = \frac{b}{2}\left(A - \sqrt{A^2-4}\right) \\ \rho_1 = \frac{\sqrt{A^2-4}+A}{2\sqrt{A^2-4}}, & \rho_2 = \frac{\sqrt{A^2-4}-A}{2\sqrt{A^2-4}}. \end{cases} \quad (19)$$

Consequently

$$\left[\mathbf{H}_{11}^{-1}\right]_{1,k} = \frac{b^{k-1}}{|\mathbf{H}_{11}|} \times \left(\rho_1 r_1^{K-k-1}(r_1 + b) + \rho_2 r_2^{K-k-1}(r_2 + b)\right) \quad (20)$$

and $\lambda = \frac{K-1}{\sigma_w^2} - \frac{2}{\sigma_w^4|\mathbf{H}_{11}|} \times \left(\rho_1 r_1^{K-2}(r_1 + b) + \rho_2 r_2^{K-2}(r_2 + b) - b^{K-1}\right)$.

From the definition of $\mathbf{V}_K$, we have

$$[\mathbf{V}_K]_{k,k} = \frac{1}{\lambda\sigma_w^4}\left(\left[\mathbf{H}_{11}^{-1}\right]_{1,k} - \left[\mathbf{H}_{11}^{-1}\right]_{1,K+1-k}\right)^2. \quad (21)$$

Using (18), (20), and (21), we obtain, for any index k, the analytical expression of the HCRB diagonal elements in (22), shown at the bottom of the next page.

Remark: Note that, if (3) was used instead of (8), the HIM would not be invertible.

C. Simulation Results

We now illustrate the behavior of the HCRB versus the signal-to-noise ratio (SNR) defined by $1/\sigma_n^2$. We consider a block of $K = 40$ BPSK transmitted symbols. For two distinct phase-noise variances ($\sigma_w^2 = 0.1$ rad^2 and $\sigma_w^2 \to 0$ rad^2), Fig. 1 superimposes on one side the HCRB [see (30)], the data-aided HCRB $\left(J_D = 2/\sigma_n^2\right)$, and the BCRB (see [4, eq.

$$\begin{cases} \mathbf{H}_{11} = \mathbb{E}_{\mathbf{y},\boldsymbol{\theta}|\xi^\star}\left[-\Delta_{\boldsymbol{\theta}}^{\boldsymbol{\theta}} \log p(\mathbf{y}|\boldsymbol{\theta}, \xi)\Big|_{\xi^\star}\right] + \mathbb{E}_{\boldsymbol{\theta}|\xi^\star}\left[-\Delta_{\boldsymbol{\theta}}^{\boldsymbol{\theta}} \log p(\boldsymbol{\theta}|\xi^\star)\right] \\ \mathbf{h}_{12} = \mathbf{h}_{21}^T = \mathbb{E}_{\mathbf{y},\boldsymbol{\theta}|\xi^\star}\left[-\Delta_{\boldsymbol{\theta}}^\xi \log p(\mathbf{y}|\boldsymbol{\theta}, \xi)\Big|_{\xi^\star}\right] + \mathbb{E}_{\boldsymbol{\theta}|\xi^\star}\left[-\Delta_{\boldsymbol{\theta}}^\xi \log p(\boldsymbol{\theta}|\xi)\Big|_{\xi^\star}\right] \\ H_{22} = \mathbb{E}_{\mathbf{y},\boldsymbol{\theta}|\xi^\star}\left[-\Delta_\xi^\xi \log, p(\mathbf{y}|\boldsymbol{\theta}, \xi)\Big|_{\xi^\star}\right] + \mathbb{E}_{\boldsymbol{\theta}|\xi^\star}\left[-\Delta_\xi^\xi \log p(\boldsymbol{\theta}|\xi)\Big|_{\xi^\star}\right] \end{cases} \quad (12)$$

$$\begin{cases} \log p(\mathbf{y}|\boldsymbol{\theta}, \xi^\star) = \sum_{k=1}^{K}\left(-\log(\pi\sigma_n^2) - \frac{1+\|y_k\|^2}{\sigma_n^2} + \log\left(\cosh\left(\frac{2}{\sigma_n^2}\Re\{y_k e^{-j\theta_k}\}\right)\right)\right) \\ \log p(\boldsymbol{\theta}|\xi^\star) = \log p(\theta_1) + (K-1)\log\left(\frac{1}{\sqrt{2\pi}\sigma_w}\right) - \sum_{k=2}^{K}\frac{(\theta_k - \theta_{k-1} - \xi^\star)^2}{2\sigma_w^2} \end{cases} \quad (13)$$

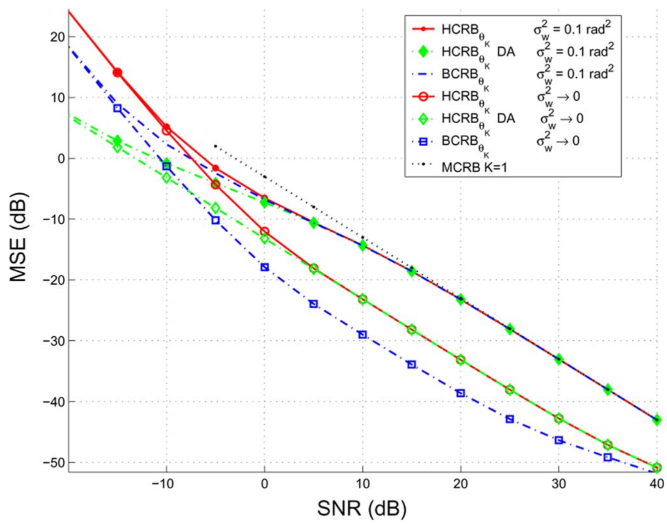

Fig. 1. Bounds on θ_K versus the SNR ($K = 40$ observations, $\sigma_w^2 = 0.1\,\mathrm{rad}^2$, and $\sigma_w^2 \to 0\,\mathrm{rad}^2$, J_D evaluated over 10^8 Monte Carlo trials).

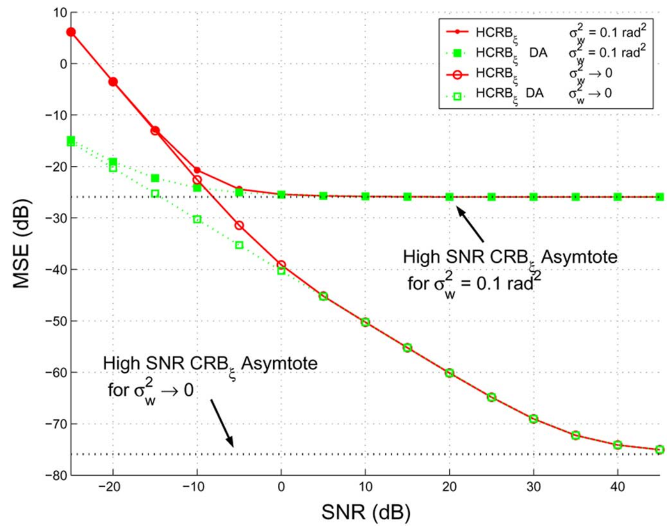

Fig. 2. Bounds on ξ versus the SNR ($K = 40$ observations, $\sigma_w^2 = 0.1\,\mathrm{rad}^2$ and $\sigma_w^2 \to 0\,\mathrm{rad}^2$, J_D evaluated over 10^8 Monte Carlo trials).

(21)]) on θ_K. For the same scenario, Fig. 2 superimposes on one side the HCRB and the data-aided HCRB on ξ.

At high SNR, we first notice that $HCRB_\xi$ converges to its horizontal asymptote given by $\sigma_w^2/K-1$ which is the standard CRB when $\boldsymbol{\theta}$ is assumed to be known. The observation noise compared to the phase noise is then not significant enough to disturb the estimation of ξ; consequently $HCRB_\xi$ depends only on the phase noise and on the number of observations. Concerning the bounds on θ_K, $HCRB_{\theta_k}$ and $BCRB_{\theta_k}$ both have the same asymptote given by $\sigma_n^2/2$ which is the modified CRB (MCRB) for one observation (see [6]). It means that, at high SNR, the observation y_K is self-sufficient to estimate θ_K, and the error on ξ does not disturb the performance on θ_K. Moreover, the HCRB logically tends to the data-aided HCRB.

For median SNR, $HCRB_{\theta_K}$ and $HCRB_\xi$ leave their respective asymptote. $HCRB_{\theta_K}$ is still lower bounded by the BCRB and upper bounded by the high-SNR asymptote. This stems from the fact that taking into account a block of observations instead of one observation necessarily improves the performance. However, for large σ_w^2 values (e.g., $\sigma_w^2 = 0.1\,\mathrm{rad}^2$), $HCRB_{\theta_K}$ stays close to the MCRB because the correlation between the phase offsets θ_k is less significant than the information brought by the observation y_K. Moreover, when σ_w^2 tends to 0, $HCRB_{\theta_K}$ is above the BCRB because performance is now limited by the accuracy on the parameter ξ.

At low SNR, n_k is preponderant compared to w_k. Both $HCRB_\xi$ and $HCRB_{\theta_K}$ do not depend on σ_w^2: the lack of knowledge on ξ directly affects the estimation on θ_K. As expected, the knowledge of the symbols (data-aided HCRB) leads to a better estimation of $\boldsymbol{\theta}$ and ξ.

IV. Conclusion

In this letter, we have studied the hybrid Cramér–Rao bound when the random and the deterministic parts of the parameter vector are statistically dependent. We have applied this bound in order to evaluate the performance of a dynamical phase estimator where the linear drift is unknown in a non-data-aided context.

References

[1] H. L. Van Trees, *Detection, Estimation and Modulation Theory*. New York: Wiley, 1968, vol. 1.
[2] Y. Rockah and P. Schultheiss, "Array shape calibration using sources in unknown locations-part I: Far-field sources," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. ASSP-35, no. 3, pp. 286–299, Mar. 1987.
[3] H. Messer, "The hybrid Cramér-Rao lower bound—From practice to theory," in *Proc. IEEE Workshop Sensor Array and Multi-Channel Processing (SAM)*, Waltham, MA, Jul. 2006, pp. 304–307.
[4] S. Bay, C. Herzet, J.-M. Brossier, J.-P. Barbot, and B. Geller, "Analytic and asymptotic analysis of Bayesian Cramér-Rao bound for dynamical phase offset estimation," *IEEE Trans. Signal Process.*, vol. 56, no. 1, pp. 61–70, Jan. 2008.
[5] J. A. McNeill, "Jitter in ring oscillators," Ph.D. dissertation, Boston Univ., Boston, MA, 1994.
[6] M. Moeneclaey, "On the true and the modified Cramér-Rao bounds for the estimation of a scalar parameter in the presence of nuisance parameters," *IEEE Trans. Commun.*, vol. 46, no. 11, pp. 1536–1544, Nov. 1998.

$$
\begin{aligned}
\left[\mathbf{H}^{-1}\right]_{k,k} = \frac{1}{|\mathbf{H}_{11}|}\Bigg[&\rho_1^2\left(b+r_1\right)^2 r_1^{K-3} + \rho_2^2(b+r_2)^2 r_2^{K-3} - \frac{b^2}{A-2}\left(r_1^{k-2}r_2^{K-k-1} + r_1^{K-k-1}r_2^{k-2}\right)\Bigg] \\
&+ \frac{1}{\lambda\sigma_w^4|\mathbf{H}_{11}|^2}\Bigg[b^{k-1}\left(\rho_1\left(r_1\right)^{K-k-1}\left(b+r_1\right) + \rho_2\left(r_2\right)^{K-k-1}\left(b+r_2\right)\right) \\
&\qquad + b^{K-k}\left(\rho_1\left(r_1\right)^{k-2}\left(b+r_1\right) + \rho_2\left(r_2\right)^{k-2}\left(b+r_2\right)\right)\Bigg]^2 \qquad (22)
\end{aligned}
$$