# Approximate Expressions for Cramer-Rao Bounds of Code Aided QAM Dynamical Phase Estimation

Jianxiao Yang, Benoit Geller, Anne Wei

# Approximate Expressions for Cramer-Rao Bounds of Code Aided QAM Dynamical Phase Estimation

J. Yang, B. Geller, and A. Wei[1]

***Abstract*—In this paper, we study Bayesian and hybrid Cramer-Rao bounds (BCRB and HCRB) for the code-aided (CA) dynamical phase estimation of QAM modulated signals. In order to avoid the calculus of the inverse of the Bayesian information matrix and of the hybrid information matrix, we present some analytical expressions for the various CRBs, which greatly reduce the computation complexity.**

## I. Introduction

Classicaly there are three ways of performing estimation in a telecommunication system: data aided (DA), code aided (CA) and non data aided (NDA) estimations [1]. Earlier attempts of signal synchronization in the low-SNR regime focused either on the DA or NDA synchronization mode [2],[3]. On one hand, DA estimation techniques achieve the better performance but may lead to unacceptable losses in power and spectral efficiency. On the other hand, NDA estimation algorithms drop some information about the transmitted data and may lead to poor results at the benefit of transmission efficiency. However, with the developments of channel coding techniques [4],[5], more and more attention has focused on CA synchronization [6],[7] which uses the decoding gain to improve the estimation performance.

A natural question which arises when designing estimators is the ultimate accuracy that one can achieve in the estimation operation. The lower bounds answer this question by providing a minimum mean square error (MMSE). Although there exists many lower bounds, the Cramer-Rao bound (CRB) family is the most commonly used and the easiest to determine [8]-[11]. There are several works concerning the CA CRBs for carrier phase and frequency estimation. In [11], the CRB for the CA carrier phase estimation has been expressed in terms of the marginal a posteriori probabilities (APPs) of the coded symbols, allowing the numerical evaluation of the CA bounds. This method has been applied to the evaluation of the CRB for the Turbo code aided [12] and the convolution code aided [13] scenarios; the CA CRBs in [11]-[13] were derived from the first derivative of the joint probability function between the observations and parameters. All these papers refer to an idealized situation in which the phase offset is constant. However, in modern communications, it is common to take into account a time-varying phase noise mostly due to oscillator instabilities [14]-[17]. [18] considered the DA CRB for the phase estimation with some noise variance. [19] has derived a Bayesian CRB restricted to the case of NDA BPSK signals. When a deterministic parameter such as the frequency offset is taken into consideration, the hybrid CRB becomes relevant. HCRBs were derived in the case of NDA signals in [20],[21] and were applied as benchmarks in [22],[23].

Differently from [11]-[13],[19],[20], the goal and the contribution of the paper is to give for the dynamic time-varying phase, both the BCRB and the HCRB in the case of coded and QAM modulated signals. In addition, we give analytical expressions to the various CRBs which greatly reduce the computation complexity. The rest of the paper is organized as follows. In section II, we recall the various kinds of CRBs. After describing the system model in section III, we derive the BCRBs and the HCRBs for both on-line and off-line estimations for the CA scenario in section IV. The derivation for the CA case is different than the one used in [11]-[13]. Moreover, we also present analytical expressions for the various CRBs which do not require to compute the inversion of any information matrix. The various results are finally illustrated in section V.[2]

## II. Cramer-Rao Bounds (CRBs) Review

In the following, we briefly review the links between the HCRB, the standard CRB, and the BCRB. In the most general case, the parameters to be estimated include both deterministic and random parameters. Denote this parameter vector as $\mathbf{u} = \left(\mathbf{u}_r^T, \mathbf{u}_d^T\right)^T$, where $\mathbf{u}_d$ is assumed to be a $(n-m)\times 1$ deterministic vector and $\mathbf{u}_r$ is assumed to be a $m \times 1$ random vector with an a priori probability density function (pdf) $p(\mathbf{u}_r)$. The true value of $\mathbf{u}_d$ will be denoted $\mathbf{u}_d^{\Delta}$. We consider $\hat{\mathbf{u}}(\mathbf{y})$ as an estimator of $\mathbf{u}$ where $\mathbf{y}$ is the observation vector. The HCRB satisfies the following inequality [20] on the MSE:

$$E_{\mathbf{y},\mathbf{u}_r|\mathbf{u}_d=\mathbf{u}_d^{\Delta}}\left[\left(\hat{\mathbf{u}}(\mathbf{y})-\mathbf{u}\right)\left(\hat{\mathbf{u}}(\mathbf{y})-\mathbf{u}\right)^T_{|\mathbf{u}_d=\mathbf{u}_d^{\Delta}}\right] \geq \mathbf{H}^{-1}\left(\mathbf{u}_d^{\Delta}\right), \qquad (1)$$

[1] This work was partially funded by the ANR LURGA program.

J. Yang is with SATIE, ENS Cachan (e-mail: shawn.yang@satie.ens-cachan.fr)

B. Geller is with LEI, ENSTA ParisTech (e-mail: geller@ensta.fr)

A. Wei is with LATTIS, Université Toulouse II (e-mail: anne.wei@lattis.Univ-toulouse.fr)

[2] The notational convention adopted is as follows: italic indicates a scalar quantity, as in $a$; boldface indicates a vector quantity, as in $\mathbf{a}$ and capital boldface indicates a matrix quantity as in $\mathbf{A}$. The $(m,n)^{th}$ entry of matrix $\mathbf{A}$ is denoted as $[A]_{m,n}$. The transpose matrix of $\mathbf{A}$ is indicated by a superscript $\mathbf{A}^T$, and $|\mathbf{A}|$ is the determinant of $\mathbf{A}$. $\mathbf{a}_m^n$ represents the vector $[a_m,\cdots,a_n]^T$, where $m$ and $n$ are positive integers ($m<n$). $\mathrm{Re}\{a\}$ and $\mathrm{Im}\{a\}$ are respectively the real and imaginary parts of $a$. $E_{xy}[\,]$ denotes the expectation over $x$ and $y$. $\nabla_{\mathbf{u}} = [\partial/\partial u_1 \ldots \partial/\partial u_n]^T$ and $\Delta_{\mathbf{u}}^{\mathbf{v}} = \nabla_{\mathbf{u}}\nabla_{\mathbf{v}}^T$ represent the first and second order derivative operators.

where $\mathbf{H}\left(\mathbf{u}_d^{\Delta}\right)$ is the so-called hybrid information matrix (HIM) and is defined as:

$$\mathbf{H}\left(\mathbf{u}_d^{\Delta}\right) = E_{\mathbf{y},\mathbf{u}_r|\mathbf{u}_d=\mathbf{u}_d^{\Delta}}\left[-\Delta_{\mathbf{u}}^{\mathbf{u}} \log p\left(\mathbf{y},\mathbf{u}_r \mid \mathbf{u}_d\right)_{|\mathbf{u}_d=\mathbf{u}_d^{\Delta}}\right]. \quad (2)$$

It is shown in [20] that inequality (1) is still respected when the deterministic and the random parts of the parameter vector are dependent. The HIM can be rewritten as:

$$\mathbf{H}\left(\mathbf{u}_d^{\Delta}\right) = E_{\mathbf{u}_r|\mathbf{u}_d=\mathbf{u}_d^{\Delta}}\left[\mathbf{F}\left(\mathbf{u}_d^{\Delta},\mathbf{u}_r\right)\right] + E_{\mathbf{u}_r|\mathbf{u}_d=\mathbf{u}_d^{\Delta}}\left[-\Delta_{\mathbf{u}}^{\mathbf{u}} \log p\left(\mathbf{u}_r \mid \mathbf{u}_d\right)_{\mathbf{u}_d=\mathbf{u}_d^{\Delta}}\right] \quad (3)$$

where $\mathbf{F}\left(\mathbf{u}_d^{\Delta},\mathbf{u}_r\right) = E_{\mathbf{y}|\mathbf{u}_r,\mathbf{u}_d=\mathbf{u}_d^{\Delta}}\left[-\Delta_{\mathbf{u}}^{\mathbf{u}} \log p\left(\mathbf{y} \mid \mathbf{u}_d,\mathbf{u}_r\right)_{|\mathbf{u}_d=\mathbf{u}_d^{\Delta}}\right]$ (4)

is the Fisher information matrix (FIM).

If in particular $\mathbf{u} = \mathbf{u}_d$, (3) reduces to:

$$\mathbf{H}\left(\mathbf{u}_d^{\Delta}\right) = \mathbf{F}\left(\mathbf{u}_d^{\Delta}\right) = E_{\mathbf{y}|\mathbf{u}_d=\mathbf{u}_d^{\Delta}}\left[-\Delta_{\mathbf{u}_d}^{\mathbf{u}_d} \log p\left(\mathbf{y} \mid \mathbf{u}_d\right)_{|\mathbf{u}_d=\mathbf{u}_d^{\Delta}}\right]. \quad (5)$$

Then, the inverse of (5) is just the standard CRB [8].

On the contrary, if $\mathbf{u} = \mathbf{u}_r$, then (3) becomes:

$$\mathbf{H} = E_{\mathbf{u}_r}\left[\mathbf{F}\left(\mathbf{u}_r\right)\right] + E_{\mathbf{u}_r}\left[-\Delta_{\mathbf{u}_r}^{\mathbf{u}_r} \log p\left(\mathbf{u}_r\right)\right] \quad (6)$$

where $\mathbf{F}\left(\mathbf{u}_r\right) = E_{\mathbf{y}|\mathbf{u}_r}\left[-\Delta_{\mathbf{u}_r}^{\mathbf{u}_r} \log p\left(\mathbf{y} \mid \mathbf{u}_r\right)\right]$. (7)

In this case, the inverse of (6) is the Bayesian CRB [8].

## III. System Model

We consider the transmission of a modulated sequence $\mathbf{s} = [s_1,\cdots,s_L]^T$ from a constellation set $\mathbf{s}_M$ (M-QAM, M-PSK, etc.) over an additive white Gaussian noise (AWGN) channel affected by some carrier phase offsets vector $\boldsymbol{\theta} = [\theta_1,\cdots,\theta_L]^T$. Without any inter-symbol interference (ISI), the received baseband symbols $\mathbf{y} = [y_1,\cdots,y_L]^T$ can be written as:

$$y_l = s_l e^{j\theta_l} + n_l = \left(a_l + jb_l\right)e^{j\theta_l} + n_l, \quad (8)$$

where $s_l$, $\theta_l$ and $n_l$ are respectively the $l$-th transmitted complex symbol, the residual phase distortion and the zero mean circular Gaussian noise with known variance $\sigma_n^2$.

We assume that the channel coding maps a message of $K$ bits to a codeword of $N$ bits denoted as $\mathbf{c} = \{c_1,c_2,\cdots,c_N\}$ with $c_n \in \{0,1\}$ $(1 \le n \le N)$ and $\mathbf{c} \in \left\{\tilde{\mathbf{c}}_v \mid v = 1,\cdots,2^K\right\}$; moreover, the constellation vector corresponding to code $\tilde{\mathbf{c}}_v$ is denoted as $\tilde{\mathbf{s}}\left(\tilde{\mathbf{c}}_v\right) = \left\{\tilde{s}_1\left(\tilde{\mathbf{c}}_v\right),\cdots,\tilde{s}_L\left(\tilde{\mathbf{c}}_v\right)\right\}$. Hence, the conditional probability based on the known vector phase is:

$$\begin{aligned} p\left(\mathbf{y} \mid \boldsymbol{\theta}\right) &= \sum_{v=1}^{2^K} p\left(\mathbf{y} \mid \mathbf{c} = \tilde{\mathbf{c}}_v, \boldsymbol{\theta}\right) p\left(\mathbf{c} = \tilde{\mathbf{c}}_v\right) = \sum_{v=1}^{2^K} p\left(\mathbf{y} \mid \tilde{\mathbf{s}} = \tilde{\mathbf{s}}\left(\tilde{\mathbf{c}}_v\right), \boldsymbol{\theta}\right) p\left(\tilde{\mathbf{s}} = \tilde{\mathbf{s}}\left(\tilde{\mathbf{c}}_v\right)\right) \\ &= \left(\frac{1}{\pi\sigma_n^2}\right)^L \sum_{v=1}^{2^K} \frac{1}{2^K} \prod_{l=1}^{L} \exp\left\{-\frac{\left|\tilde{s}_l\left(\tilde{\mathbf{c}}_v\right)\right|^2 + \left|y_l\right|^2}{\sigma_n^2} + 2\frac{\mathrm{Re}\left\{y_l \tilde{s}_l^*\left(\tilde{\mathbf{c}}_v\right)e^{-j\theta_l}\right\}}{\sigma_n^2}\right\}. \end{aligned} \quad (9)$$

In practice, clocks are never perfect, and oscillators suffer from jitters. This results in a Brownian phase model with a linear drift:

$$\theta_l = \theta_{l-1} + \xi + w_l, \quad (10)$$

where $\theta_l$ is the unknown phase offset at time $l$, $\xi$ is the unknown constant frequency offset (linear drift), $w_l$ is a white Gaussian noise with zero mean and variance $\sigma_w^2$. This model is commonly used [14]-[17] in order to describe the behavior of practical oscillators for which the frequency is randomly perturbed. The corresponding conditional probability is:

$$p\left(\theta_l \mid \theta_{l-1},\xi\right) = \frac{1}{\sqrt{2\pi}\sigma_w}\exp\left[-\frac{\left(\theta_l - \theta_{l-1} - \xi\right)^2}{2\sigma_w^2}\right]. \quad (11)$$

## IV. CRBs for the Dynamical Phase Estimation

In practical receivers, phase estimation can actually be considered following two main scenarios:

- *Off-line synchronization*: the receiver waits until the whole observation frame $\mathbf{y}$ has been received. Then afterwards, it processes all the observations to compute the estimates of the carrier phase offsets $\boldsymbol{\theta}$.
- *On-line synchronization*: the receiver estimates $\theta_l$ upon the arrival of the $l$-th observation, i.e. $y_l$. The phase estimate is then computed based on the current and previous observations only, i.e. $\mathbf{y}_1^l = [y_1,\cdots,y_l]^T$.

In this section, we derive some analytical expressions for both the on-line and the off-line CRBs. The parameters of the phase model in (10) include some random parameters $\boldsymbol{\theta} = [\theta_1,\cdots,\theta_L]^T$ (i.e. the dynamical phase) and a deterministic parameter $\xi$ (i.e. the linear drift). So the parameters vector can be written as:

$$\mathbf{u} = \begin{bmatrix}\mathbf{u}_r \\ \mathbf{u}_d\end{bmatrix} = \begin{bmatrix}\boldsymbol{\theta} \\ \xi\end{bmatrix}. \quad (12)$$

Equation (3) thus becomes:

$$\begin{aligned} \mathbf{H}\left(\xi^{\Delta}\right) &= E_{\boldsymbol{\theta}|\xi^{\Delta}}\left[\mathbf{F}\left(\xi^{\Delta},\boldsymbol{\theta}\right)\right] \\ &+ E_{\boldsymbol{\theta}|\xi=\xi^{\Delta}}\begin{bmatrix} -\Delta_{\boldsymbol{\theta}}^{\boldsymbol{\theta}} \log p\left(\boldsymbol{\theta} \mid \xi\right)_{|\xi=\xi^{\Delta}} & -\Delta_{\boldsymbol{\theta}}^{\xi} \log p\left(\boldsymbol{\theta} \mid \xi\right)_{|\xi=\xi^{\Delta}} \\ \left(-\Delta_{\boldsymbol{\theta}}^{\xi} \log p\left(\boldsymbol{\theta} \mid \xi\right)_{|\xi=\xi^{\Delta}}\right)^T & -\Delta_{\xi}^{\xi} \log p\left(\boldsymbol{\theta} \mid \xi\right)_{|\xi=\xi^{\Delta}} \end{bmatrix} \end{aligned} \quad (13)$$

where $\mathbf{F}\left(\xi^{\Delta},\boldsymbol{\theta}\right) = E_{\mathbf{y}|\boldsymbol{\theta},\xi=\xi^{\Delta}}\left[-\Delta_{\boldsymbol{\theta}}^{\boldsymbol{\theta}} \log p\left(\mathbf{y} \mid \xi,\boldsymbol{\theta}\right)_{|\xi=\xi^{\Delta}}\right]$.

We then decompose the HIM into several sub-matrices that will be useful in the sequel:

$$\mathbf{H} = \begin{bmatrix}\mathbf{H}_{11} & \mathbf{H}_{12} \\ \mathbf{H}_{21} & \mathbf{H}_{22}\end{bmatrix} = \begin{bmatrix}\mathbf{H}_{11} & \mathbf{H}_{12} \\ \mathbf{H}_{12}^T & \mathbf{H}_{22}\end{bmatrix}, \quad (14)$$

where

$$\begin{cases} \mathbf{H}_{11} = E_{\mathbf{y},\boldsymbol{\theta}|\xi=\xi^{\Delta}}\left[-\Delta_{\boldsymbol{\theta}}^{\boldsymbol{\theta}} \log p\left(\mathbf{y},\boldsymbol{\theta} \mid \xi\right)_{|\xi=\xi^{\Delta}}\right] + E_{\boldsymbol{\theta}|\xi=\xi^{\Delta}}\left[-\Delta_{\boldsymbol{\theta}}^{\boldsymbol{\theta}} \log p\left(\boldsymbol{\theta} \mid \xi\right)_{|\xi=\xi^{\Delta}}\right] \\ \mathbf{H}_{12} = E_{\mathbf{y},\boldsymbol{\theta}|\xi=\xi^{\Delta}}\left[-\Delta_{\boldsymbol{\theta}}^{\xi} \log p\left(\mathbf{y},\boldsymbol{\theta} \mid \xi\right)_{|\xi=\xi^{\Delta}}\right] + E_{\boldsymbol{\theta}|\xi=\xi^{\Delta}}\left[-\Delta_{\boldsymbol{\theta}}^{\xi} \log p\left(\boldsymbol{\theta} \mid \xi\right)_{|\xi=\xi^{\Delta}}\right] \\ \mathbf{H}_{22} = E_{\mathbf{y},\boldsymbol{\theta}|\xi=\xi^{\Delta}}\left[-\Delta_{\xi}^{\xi} \log p\left(\mathbf{y},\boldsymbol{\theta} \mid \xi\right)_{|\xi=\xi^{\Delta}}\right] + E_{\boldsymbol{\theta}|\xi=\xi^{\Delta}}\left[-\Delta_{\xi}^{\xi} \log p\left(\boldsymbol{\theta} \mid \xi\right)_{|\xi=\xi^{\Delta}}\right]. \end{cases} \quad (15)$$

### A. Computation of $E_{\boldsymbol{\theta}|\xi^{\Delta}}\left[\mathbf{F}\left(\xi^{\Delta},\boldsymbol{\theta}\right)\right]$

From (9), one obtains after some calculation the first derivative with respect to $\theta_m$:

$$\frac{\partial \ln p\left(\mathbf{y} \mid \boldsymbol{\theta},\xi\right)}{\partial \theta_m} = \sum_{v=1}^{2^K} \Pr\left(\mathbf{c} = \tilde{\mathbf{c}}_v \mid \mathbf{y},\boldsymbol{\theta},\xi\right) \frac{\partial \ln\left(y_m \mid s_m = \tilde{s}_m\left(\tilde{\mathbf{c}}_v\right),\theta_m,\xi\right)}{\partial \theta_m} \quad (16)$$

where $\Pr\left(\mathbf{c} = \tilde{\mathbf{c}}_v \mid \mathbf{y},\boldsymbol{\theta},\xi\right)$ is the a posteriori probability of a code word. Discriminating in (16) the code words according to the M

possible values of the m[th] received constellation symbol, and noting $\Pr(s_m = \tilde{s}_i \mid \{\tilde{\mathbf{c}}_v\}, \mathbf{y}, \boldsymbol{\theta}, \xi)$ the APP when all possible codewords are known, one further obtains:

$$\frac{\partial \ln p(\mathbf{y} \mid \boldsymbol{\theta}, \xi)}{\partial \theta_m} = \frac{2}{\sigma_n^2} \sum_{i=1}^{M} \Pr(s_m = \tilde{s}_i \mid \{\tilde{\mathbf{c}}_v\}, \mathbf{y}, \boldsymbol{\theta}, \xi) \operatorname{Im}\left\{ y_l \tilde{s}_i^* e^{-j\theta_l} \right\}\Big|_{\tilde{s}_i \in \mathbf{S}_M} \quad (17)$$

Taking the second derivative with respect to $\theta_k$ and $\theta_m$, respectively for $k \neq m$ and for $k = m$, one obtains:

$$\begin{aligned} \frac{\partial^2 \ln p(\mathbf{y} \mid \boldsymbol{\theta}, \xi)}{\partial \theta_k \partial \theta_m} &= -\frac{\partial \ln p(\mathbf{y} \mid \boldsymbol{\theta}, \xi)}{\partial \theta_k} \cdot \frac{\partial \ln p(\mathbf{y} \mid \boldsymbol{\theta}, \xi)}{\partial \theta_m} \\ &+ \sum_{i_1=1}^{M} \sum_{i_2=1}^{M} \frac{2 \operatorname{Im}\left\{ y_k s_{i_1}^* e^{-j\theta_k} \right\}}{\sigma_n^2} \frac{2 \operatorname{Im}\left\{ y_m s_{i_2}^* e^{-j\theta_m} \right\}}{\sigma_n^2} \Pr(s_k = \tilde{s}_{i_1}, s_m = \tilde{s}_{i_2} \mid \{\tilde{\mathbf{c}}_v\}, \mathbf{y}, \boldsymbol{\theta}, \xi), \end{aligned} \quad (18)$$

and:

$$\begin{aligned} \frac{\partial^2 \ln p(\mathbf{y} \mid \boldsymbol{\theta}, \xi)}{\partial \theta_k^2} &= -\left( \frac{\partial \ln p(\mathbf{y} \mid \boldsymbol{\theta}, \xi)}{\partial \theta_k} \right)^2 \\ &+ \sum_{i_1=1}^{M} \Pr(s_k = s_{i_1}^* \mid \{\tilde{\mathbf{c}}_v\}, \mathbf{y}, \boldsymbol{\theta}, \xi) \left( -\frac{2 \operatorname{Re}\left\{ y_k s_{i_1}^* e^{-j\theta_k} \right\}}{\sigma_n^2} + \left( \frac{2 \operatorname{Im}\left\{ y_k s_{i_1}^* e^{-j\theta_k} \right\}}{\sigma_n^2} \right)^2 \right). \end{aligned} \quad (19)$$

If the codeword has been sufficiently interleaved before the mapping from bits into symbols, then:

$$\Pr(s_k = \tilde{s}_{i_1}, s_m = \tilde{s}_{i_2} \mid \{\tilde{\mathbf{c}}_v\}, \mathbf{y}, \boldsymbol{\theta}, \xi) \approx \Pr(s_k = \tilde{s}_{i_1} \mid \{\tilde{\mathbf{c}}_v\}, \mathbf{y}, \boldsymbol{\theta}, \xi) \Pr(s_m = \tilde{s}_{i_2} \mid \{\tilde{\mathbf{c}}_v\}, \mathbf{y}, \boldsymbol{\theta}, \xi), \quad (20)$$

so that taking into account (17), (18) just becomes :

$$\frac{\partial^2 \ln p(\mathbf{y} \mid \boldsymbol{\theta}, \xi)}{\partial \theta_k \partial \theta_m} \approx 0. \quad (21)$$

It is important to note that differently from [11]-[13], we take the second derivative to calculate the FIM so as to reduce the complexity of the computation with the help of equation (21). Then the FIM can be approximately written as a diagonal matrix and one has:

$$E_{\boldsymbol{\theta}}\left[ \mathbf{F}\left( \xi^{\Delta}, \boldsymbol{\theta} \right) \right] = J_D \mathbf{I}_L, \quad (22)$$

where $\mathbf{I}_L$ is the $L \times L$ identity matrix and $J_D$ is defined as:

$$J_D \triangleq E_{\mathbf{y},\boldsymbol{\theta}}\left[ -\frac{\partial^2 \ln p(\mathbf{y} \mid \boldsymbol{\theta}, \xi)}{\partial \theta_k^2} \right]. \quad (23)$$

### *B. Computation of the HIM*

Like detailed in [19],[20], we now calculate the different sub-matrices of $\mathbf{H}$ (see (14)). From (15), due to the model of Section III and under the assumption that we have no priori knowledge of $\theta_1$, i.e. $E_{\theta_1}\left[ \Delta_{\theta_1}^{\theta_1} \log p(\theta_1) \right] = 0$, one obtains $\mathbf{H}_{11}$ which has a particular mathematical structure just like in [19]:

$$\mathbf{H}_{11} = b \begin{bmatrix} A+1, & 1 & 0 & \cdots & 0 \\ 1 & A & 1 & 0 & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & 0 & 1 & A & 1 \\ 0 & \cdots & 0 & 1 & A+1 \end{bmatrix}, \quad (24)$$

where $A = -\sigma_w^2 J_D - 2$ and $b = -1/\sigma_w^2$. (25)

From (9), one can see that $\log p(y_l \mid \xi, \theta_l)$ is independent of $\xi$, i.e., the partial derivatives $\Delta_{\boldsymbol{\theta}}^{\xi} \log p(\mathbf{y}, \boldsymbol{\theta} \mid \xi)_{|\xi = \xi^{\Delta}}$ and $\Delta_{\xi}^{\xi} \log p(\mathbf{y}, \boldsymbol{\theta} \mid \xi)_{|\xi = \xi^{\Delta}}$ are equal to 0. $\mathbf{H}_{12}$ and $\mathbf{H}_{22}$ thus become:

$$\mathbf{H}_{12} = \left[ 1/\sigma_w^2, \quad \mathbf{0}_{1\times(L-2)}, \quad -1/\sigma_w^2 \right]^T \quad (26)$$

$$\mathbf{H}_{22} = (L-1)/\sigma_w^2. \quad (27)$$

We now inverse the HIM so as to obtain the analytical HCRB.

### *C. Analytical Expressions of HCRBs*

Similarly to Appendix I of [19], from (24)-(25) one obtains:

$$\left[ \mathbf{H}_{11}^{-1} \right]_{1,l} = \frac{b^{l-1}}{|\mathbf{H}_{11}|} \left( \rho_1 r_1^{L-l-1} (r_1 + b) + \rho_2 r_2^{L-l-1} (r_2 + b) \right), \quad (28)$$

$$\left[ \mathbf{H}_{11}^{-1} \right]_{l,l} = \frac{1}{|\mathbf{H}_{11}|} \left\{ \begin{array}{l} \rho_1^2 (b + r_1)^2 r_1^{L-3} + \rho_2^2 (b + r_2)^2 r_2^{L-3} \\ \quad - b^2 \left( r_1^{L-2} r_2^{L-l-1} + r_1^{L-l-1} r_2^{L-2} \right) (A-2)^{-1} \end{array} \right\}, \quad (29)$$

where $\left\{ \begin{array}{l} r_1 = 1/\sigma_w^2 + \left( 1 - \sqrt{1 + 4 \left( J_D \sigma_w^2 \right)^{-1}} \right) J_D / 2 \\ r_2 = 1/\sigma_w^2 + \left( 1 + \sqrt{1 + 4 \left( J_D \sigma_w^2 \right)^{-1}} \right) J_D / 2 \end{array} \right.$, (30)

and $\left\{ \begin{array}{l} \rho_1 \triangleq \left[ \sqrt{1 + 4 \left( J_D \sigma_w^2 \right)^{-1}} - 1 - 2 \left( J_D \sigma_w^2 \right)^{-1} \right] \Big/ \left[ 2 \sqrt{1 + 4 \left( J_D \sigma_w^2 \right)^{-1}} \right] \\ \rho_2 \triangleq \left[ \sqrt{1 + 4 \left( J_D \sigma_w^2 \right)^{-1}} + 1 + 2 \left( J_D \sigma_w^2 \right)^{-1} \right] \Big/ \left[ 2 \sqrt{1 + 4 \left( J_D \sigma_w^2 \right)^{-1}} \right] \end{array} \right.$. (31)

Thanks to the block-matrix inversion formula [10], we have:

$$\mathbf{H}^{-1} = \begin{bmatrix} \mathbf{H}_{11}^{-1} + \mathbf{V}_L & -\lambda^{-1} \mathbf{H}_{11}^{-1} \mathbf{H}_{12} \\ -\lambda^{-1} \mathbf{H}_{12}^T \mathbf{H}_{11}^{-1} & \lambda^{-1} \end{bmatrix}, \quad (32)$$

where we define $\lambda \triangleq \frac{L-1}{\sigma_w^2} - \mathbf{H}_{12}^T \mathbf{H}_{11}^{-1} \mathbf{H}_{12}$ and $\mathbf{V}_L \triangleq \lambda^{-1} \mathbf{H}_{11}^{-1} \mathbf{H}_{12} \mathbf{H}_{12}^T \mathbf{H}_{11}^{-1}$.

To obtain the analytical expression of $\lambda$ and of the diagonal terms of $\mathbf{V}_L$, we just exploit that $\mathbf{H}_{12}$ in (26) has only two non-zero terms and that $\mathbf{H}_{11}$ has a particular structure (see (24)); substituting (26),(28),(29) into the definition of $\lambda$ and $\mathbf{V}_L$, we directly have:

$$\lambda = (L-1)/\sigma_w^2 - 2b^2 \left\{ \rho_1 r_1^{L-2} (r_1 + b) + \rho_2 r_2^{L-2} (r_2 + b) + b^{L-1} \right\} \Big/ |\mathbf{H}_{11}|, \quad (33)$$

and the diagonal elements $[\mathbf{V}_L]_{l,l}$ written as:

$$[\mathbf{V}_L]_{l,l} = \frac{1}{\lambda \sigma_w^4} \left( \left[ \mathbf{H}_{11}^{-1} \right]_{1,l} - \left[ \mathbf{H}_{11}^{-1} \right]_{1,L+1-l} \right)^2. \quad (34)$$

We now derive an analytical expression of the diagonal elements of $\mathbf{H}_{11}^{-1} + \mathbf{V}_L$ corresponding to the minimum bound on $\boldsymbol{\theta}$ (see (32)). Using (29) and (34), one can also get the analytical expression of the upper diagonal elements $\left[ \mathbf{H}^{-1} \right]_{l,l}$ in (32) i.e. the off-line HCRB associated to the estimation of $\theta_k$ :

$$\begin{aligned} \left[ \mathbf{H}^{-1} \right]_{l,l} &= \frac{1}{|\mathbf{H}_{11}|} \left\{ \begin{array}{l} \rho_1^2 (b + r_1)^2 r_1^{L-3} + \rho_2^2 (b + r_2)^2 r_2^{L-3} \\ \quad - b^2 \left( r_1^{L-2} r_2^{L-l-1} + r_1^{L-l-1} r_2^{L-2} \right) (A-2)^{-1} \end{array} \right\} \\ &+ \frac{b^2}{\lambda |\mathbf{H}_{11}|^2} \left\{ \begin{array}{l} b^{l-1} \left( \rho_1 r_1^{L-l-1} (b + r_1) + \rho_2 r_2^{L-l-1} (b + r_2) \right) \\ \quad + b^{L-l} \left( \rho_1 r_1^{l-2} (b + r_1) + \rho_2 r_2^{l-2} (b + r_2) \right) \end{array} \right\}^2. \end{aligned} \quad (35)$$

Finally, note that replacing $|\mathbf{H}_{11}|$ by $|\mathbf{H}_{11}(l)|$ in (35), where $\mathbf{H}_{11}(l)$ is the upper left $l \times l$ sub-matrix of $\mathbf{H}_{11}$ (see (24)), one also readily obtains the analytical expression of the on-line HCRB associated to the estimation of $\theta_l$ ($l \geq 3$):

$$\begin{aligned} C_{\mathbf{H}_l} &= \frac{1}{|\mathbf{H}_{11}(l)|} \left\{ \begin{array}{l} \rho_1^2 (b + r_1)^2 r_1^{l-3} + \rho_2^2 (b + r_2)^2 r_2^{l-3} \\ \quad - b^2 \left( r_1^{l-2} r_2^{-1} + r_1^{-1} r_2^{l-2} \right) (A-2)^{-1} \end{array} \right\} \\ &+ \frac{b^2}{\lambda |\mathbf{H}_{11}(l)|^2} \left\{ \begin{array}{l} b^{l-1} \left( \rho_1 r_1^{-1} (b + r_1) + \rho_2 r_2^{-1} (b + r_2) \right) \\ \quad + \rho_1 r_1^{l-2} (b + r_1) + \rho_2 r_2^{l-2} (b + r_2) \end{array} \right\}^2. \end{aligned} \quad (36)$$

One can notice that (35) and (36) do not depend on the value of the parameter $\xi$ .

### D. The Bayesian Cramer-Rao Bounds (BCRBs)

When there is no linear drift i.e. $\xi = 0$ , the parameter vector $\mathbf{u}$ contains only random parameters $\boldsymbol{\theta}$ , i.e. $\mathbf{u} = \mathbf{u}_r = \boldsymbol{\theta}$ . In this scenario, the BCRB is the lower bound of the MSE. Moreover, the Bayesian information matrix (BIM) $\mathbf{B}_L$ is equal to the upper left sub-matrix of the hybrid information matrix (HIM):

$$\mathbf{B}_L = \mathbf{H}_{11} \tag{37}$$

The diagonal element $\left[\mathbf{B}_L^{-1}\right]_{l,l}$ of the inverse of matrix $\mathbf{B}_L$ is the off-line BCRB associated to the estimation of $\theta_l$ . Consequently, from the previous sections, the corresponding analytical expressions associated with the off-line and the on-line BCRB are respectively:

$$\left[\mathbf{B}_L^{-1}\right]_{l,l} = \frac{1}{|\mathbf{B}_L|}\left\{\begin{array}{l}\rho_1 (b + r_1)^2 r_1^{L-3} + \rho_2 (b + r_2)^2 r_2^{L-3} \\ \quad - b^2 \left(r_1^{L-2} r_2^{L-l-1} + r_1^{L-l-1} r_2^{L-2}\right)(A-2)^{-1}\end{array}\right\}, \tag{38}$$

$$\text{and } C_{\mathbf{B}_l} = \frac{1}{|\mathbf{B}_l|}\left\{\begin{array}{l}\rho_1 (b + r_1)^2 r_1^{l-3} + \rho_2 (b + r_2)^2 r_2^{l-3} \\ \quad - b^2 \left(r_1^{l-2} r_2^{-1} + r_1^{-1} r_2^{l-2}\right)(A-2)^{-1}\end{array}\right\}, \tag{39}$$

where $r_1$ , $r_2$ , $\rho_1$ and $\rho_2$ are given by (30) and (31). Note that (38) (resp. (39)) is the first term on the right side of (35) (resp. (36)). The second terms in (35) and (36) represent the additional positive uncertainty brought by $\xi$ so that the HCRB is always lower bounded by the BCRB.

## V. Simulation and Discussion

We assume the transmission of Gray mapped symbols. We display results for the rate $1/2$ and 64 states non-recursive convolution code adopted in DVB-T [24] with a pseudo random interleaver. For codes that are described by means of a trellis, the marginal symbol APPs can be computed from the trellis state APPs and state transition APPs, which in turn can be determined efficiently by the famous BCJR algorithm. Constrained by the paper size, only the HCRB will be discussed in the following. First like in [19], one readily sees on Fig. 1 the superiority of the off-line approach compared to the on-line approach in the different positions of the block. Also, there is little improvement for the CA scenario (compared to the NDA scenario) when using a BPSK modulation but the gain becomes obvious for a larger constellation. Moreover, as the observation number increases, both the on-line and off-line CRBs decrease and tend to reach the corresponding asymptote values.

We now illustrate the behavior of the HCRB on $\theta_l$ as a function of the SNR (Fig. 2 to Fig. 4).

- At high SNR (above 30dB), we notice that the various CRBs logically merge independently of the constellations, on-line/off-line and DA/CA/NDA schemes. The received symbols are reliable enough to make a correct decision and the additive noise can be neglected. Hence, the estimation problem tends to a deterministic phase estimation problem where we estimate $L$ independent phases $\theta_l$ with $L$ independent observations.

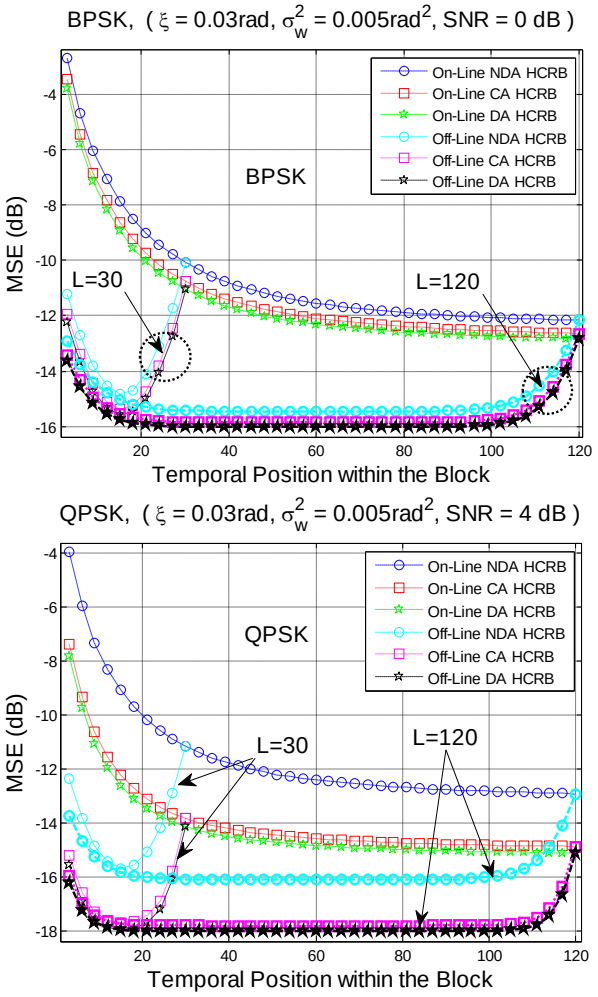


Fig. 1 HCRBs in the various block positions for two constellations (BPSK, QPSK) and two block lengths ( $L = 30$ , $L = 120$ ).

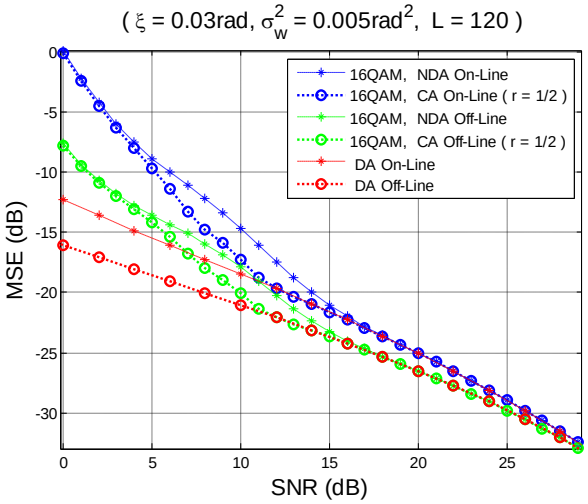


Fig. 2 HCRBs in the center of the block ( $l = 60$ ) for a 16QAM.

- In mid-range SNRs, the on-line HCRBs leave their corresponding off-line HCRBs, because one observation is not sufficient to estimate the phase offset and a block of observations can improve the estimation performance. This also explains why the NDA CRBs do not merge anymore with the CA and the DA CRBs. Moreover, we note on Fig.

4 that every time that the constellation size is increased by a factor of 4, the thresholds where the CA bounds leave the DA bound are increased by 6dB.

- At low SNR, the lack of symbol knowledge directly affects the estimation on $\theta_l$. This illustrates why the NDA CRBs increase quicker at low SNR than at high SNR. Note that the CRBs of the BPSK do not merge with the CRBs of the other two-dimensions constellations.

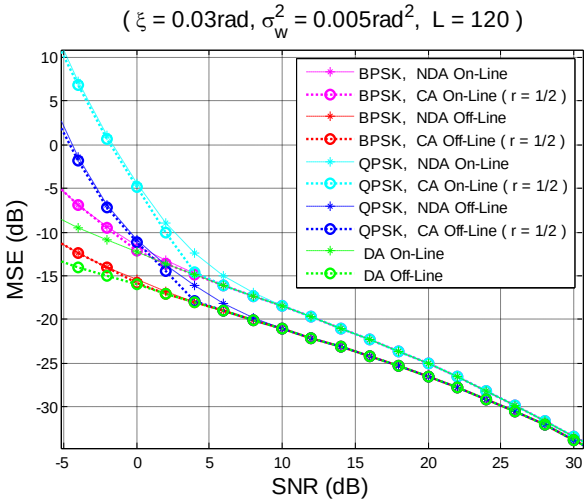


Fig. 3 HCRBs in the center of the block ( $l = 60$ ) for BPSK and QPSK.

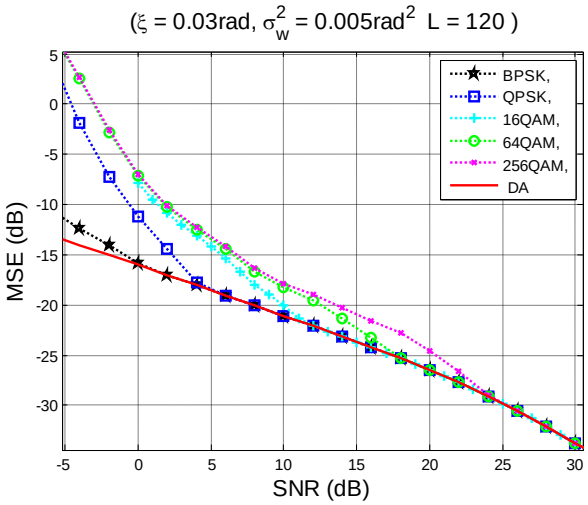


Fig. 4 Off-Line CA HCRBs in the center of the block ( $l = 60$ ).

## VI. Conclusion

In this paper, we have applied the general form of the CA BCRB and of the CA HCRB in order to evaluate the ultimate performance of a dynamical phase estimator. This illustrates the respective possible advantage of decoding and of the off-line scenario on the synchronization performance. In particular, besides the off-line synchronization gain, there is some space for additional CA synchronization gain at low SNR.

## References


[1] C. Herzet, N. Noels, V. Lottici, H. Wymeersch, M. Luise, M. Moeneclaey, L. Vandendorpe "Code-aided turbo synchronization," *Proceedings of the IEEE*, Vol. 95, pp. 1255-1271, June 2007.

[2] H. Meyr, M. Moeneclaey, and S. Fechtel, *Digital Communication Receivers: Synchronization, Channel Estimation and Signal Processing*, New York: Wiley, 1998.

[3] U. Mengali and A. N. D'Andrea, *Synchronization Techniques for Digital Receivers*. New York: Plenum, 1997.

[4] R. Gallager, "Low-Density Parity-Check Codes," *IEEE Trans. Inform. Theory*, vol. 8, pp. 21–28, Jan. 1962.

[5] C. Berrou and A. Glavieux, "Near optimum error correcting coding and decoding: turbo-codes," *IEEE Trans. Commun.*, vol. 44, pp. 1261–1272, Oct. 1996.

[6] C. Berrou, J. Hagenauer, M. Luise, C. Schlegel, L. Vandendorpe: "Turbo-Information Processing: Algorithms, Implementations & Applications," *Editorial, IEEE Proceedings*, vol. 95 n. 6, June 2007.

[7] C. Berrou, C. Langlais, and F. Seguin, "Turbo Processing," *New Directions in Statistical Signal Processing: From Systems to Brain*, MIT Press Cambridge, 2006.

[8] H. L. V. Trees, *Detection, Estimation and Modulation Theory*. New York: Wiley, 1968, vol. 1.

[9] S. M. Kay, *Fundamentals of statistical signal processing: estimation theory*. Upper Saddle River, NJ, USA: Prentice-Hall, Inc., 1993.

[10] B. Z. Bobrovsky, E. Mayer-Wolf, and M. Zakai, "Some classes of global Cramer-Rao bounds," *Ann. Statistics*, vol. 15, pp. 1421 – 1438, 1987.

[11] N. Noels, H. Steendam, M. Moeneclaey, "The Cramer-Rao Bound for Phase Estimation from Coded Linearly Modulated Signals," *IEEE Communication Letters*, vol. 7, No. 5, pp. 207-209, May 2003.

[12] N. Noels, H. Steendam, M. Moeneclaey, "Carrier and Clock recovery in (turbo) coded systems: Cramer-Rao Bound and Synchronizer Performance" , *EURASIP Journal on Applied Signal processing, Special Issue on Turbo Processing*, vol. 2005, Nr. 6, pp. 972-980, May 2005.

[13] N. Noels, H. Steendam, M. Moeneclaey, "The True Cramer-Rao Bound for Estimating the Carrier Phase of a Convolutionally Encoded PSK Signal", *9th Symp. on Vehicular Tech. and Commun. 2002, SCVT'02*, Louvain-la-Neuve, Belgium, pp. 9-14, Oct. 17, 2002.

[14] P. O. Amblard, J. M. Brossier, and E. Moisan, "Phase tracking: what do we gain from optimality? Particle filtering versus phase-locked loops," *Elsevier Signal Processing*, vol. 83, pp. 151-167, Oct. 2003.

[15] J. A. McNeill, *Jitter in ring oscillators*, Ph.D. dissertation, Boston University, 1994.

[16] A. Demir, A. Mehrotra, and J. Roychowdhury, "Phase noise in oscillators: a unifying theory and numerical methods for characterization," *IEEE Trans. Circuits Syst. I*, vol. 47, pp. 655–674, May 2000.

[17] ETSI TR 102 376, *Digital Video Broadcasting (DVB) User guidelines for the second generation system for Broadcasting, Interactive Services, News Gathering and other broadband satellite application (DVB-S2)*, v1.1.1, Feb, 2005.

[18] A. Barbieri, D. Bolletta, and G. Colavolpe, "On the Cramer-Rao bound for carrier frequency estimation in the presence of phase noise," *IEEE GLOBECOM*, 28 Nov.-2 Dec., 2005.

[19] S. Bay, C. Herzet, J. Barbot, J. M. Brossier and B. Geller,"Analytic and Asymptotic Analysis of Bayesian Cramér–Rao Bound for Dynamical Phase Offset Estimation," *IEEE Transactions on Signal Processing*, vol. 56, pp. 61-70 Jan 2008.

[20] S. Bay, B. Geller, A. Renaux, J.P. Barbot and J.M. Brossier, "On the general form of the Hybrid CRB and its application to the dynamical phase estimation," *IEEE Signal Processing Letters*, vol. 15, pp. 453-456, 2008.

[21] J. Yang, B. Geller, and A. Wei, "Bayesian and Hybrid Cramer-Rao Bounds for QAM Dynamical Phase Estimation," *in Proc. IEEE* ICASSP'09, Taipei, 19-24 April 2009.

[22] J. Yang, B. Geller, C. Herzet, and J.M. Brossier, "Smoothing PLLs for QAM Dynamical Phase Estimation," *in Proc. IEEE Inter. Conf. on Commun. 2009, ICC'09*, Dresden, 14-18 June 2009.

[23] J. Yang, and B. Geller, "Near-optimum Low-Complexity Smoothing Loops for Dynamical Phase Estimation," *accepted by IEEE Trans. on Signal Processing.*

[24] ETSI EN 300 744, Digital *Video Broadcasting (DVB); Frame structure, channel coding and modulation for digital terrestrial television*, v1.5.1, June, 2004.