DERIVATION OF A BAYESIAN CRAMÉR-RAO BOUND FOR DYNAMICAL PHASE OFFSET ESTIMATION

S. Bay[(1)], *C. Herzet*[(1,2)], *J.M. Brossier*[(3)], *J.P. Barbot*[(1)], *A. Renaux*[(1)] *and B. Geller*[(1,4)]

[(1)]ENS Cachan / Laboratoire SATIE, France
[(2)]UCL / TELE, Louvain La Neuve, Belgique
[(3)]INPG / Laboratoire LIS, Saint Martin d'Hères, France
[(4)]Université Paris 12 IUTGTR Créteil, France

ABSTRACT

In this paper, we derive a closed-form expression of a Bayesian Cramér-Rao bound (BCRB) for the estimation of a dynamical phase offset. We consider the scenario of an uncoded AWGN transmission and a Wiener phase-offset model. We provide an analytical expression of this bound in both the so-called *off-line* and *on-line* contexts. Taking benefit from the derived BCRB expression, we study its asymptotic behavior.

***Index Terms*—** Synchronization, phase estimation, communication system performance

1. INTRODUCTION

In digital communication systems, phase error considerably degrades the receiver performance. Synchronization is therefore a fundamental task of the receiver. In order to assess the performance of practical phase estimators [1], a classical tool is the Cramér-Rao lower Bound (CRB) which is a fundamental limit on the variance of any unbiased estimator. Considering dynamical parameter estimation, the unknown parameters can no longer be regarded as "deterministic" and the performance analysis requires the derivation of the Bayesian CRB (BCRB). So far, BCRBs associated to dynamical carrier-phase synchronization have already been considered in some contributions, see particularly Brossier *et al.* [2] and Dauwels [3, 4]. In [2], the authors provide a sequential expression of the BCRB for the one-line dynamical phase estimation. In [3, 4], the author proposes a numerical method based on graph representation to efficiently evaluate CRBs.

In this contribution, we derive a BCRB associated to the estimation of a dynamical phase offset obeying a Wiener model. We consider an uncoded BSPK scenario. We first focus on the off-line scenario by providing a closed-form expression of the BCRB whereas Dauwels' results are numerical [3]. We then consider its asymptotic behavior at low and high SNR. As a by-product, we show that the change from the off-line scenario to the on-line scenario is straightforward, completing in this way the result of Brossier *et al* [2]. Finally, we derive the asymptote of the off-line BCRB for an infinite number of observations.

This paper is organized as follows. In Section 2, we set the system model. In Section 3, we present and derive the BCRB. In Section 4, the asymptotic cases at low and high-SNR are studied. In Section 5, an on-line interpretation of the bound is given. Finally, Section 6 gives an illustration of the bounds for different scenarios.

2. MODEL

We consider the transmission of a BPSK modulated sequence $\mathbf{a} = [a_0 \cdots a_{K-1}]^T$ over an AWGN channel affected by a time-varying phase offset. Assuming that the received signal has been ideally filtered and sampled at the optimum sampling instant, the discrete-time baseband signal is given by

$$y_k = a_k e^{j\theta_k} + n_k \quad \text{with } k = 0 \ldots K-1, \tag{1}$$

where θ_k is the phase offset affecting y_k and n_k is a zero-mean circular Gaussian noise with known variance σ_n^2, so that the SNR is related to the E_s/N_0-ratio as $\frac{E_s}{N_0} = \frac{1}{\sigma_n^2}$. We assume an uncoded scenario so that the transmitted symbols are independent and identically distributed (i.i.d.) with $p(a_k = \pm 1) = \frac{1}{2}$. We furthermore assume that the phase offset obeys a Wiener model, *i.e.*,

$$\theta_k = \theta_{k-1} + w_k, \tag{2}$$

where w_k is an i.i.d. zero-mean Gaussian noise with known variance σ_w^2. This model is commonly used [5] to describe the behavior of practical oscillators. In the sequel, we stack the carrier phase offsets in a vector $\boldsymbol{\theta} = [\theta_0 \cdots \theta_{K-1}]^T$ and the observations in a vector $\mathbf{y} = [y_0 \cdots y_{K-1}]^T$.

3. EXPRESSION OF THE OFF-LINE BCRB

Van Trees [6] shows that any estimator $\hat{\boldsymbol{\theta}}(\mathbf{y})$ is bounded by the inverse of the Bayesian Information Matrix (BIM), say $\mathbf{B}$, as follows

$$E_{\mathbf{y},\boldsymbol{\theta}} \; (\hat{\boldsymbol{\theta}}(\mathbf{y}) - \boldsymbol{\theta})(\hat{\boldsymbol{\theta}}(\mathbf{y}) - \boldsymbol{\theta})^T \; \geq \mathbf{B}^{-1}. \tag{3}$$

The BIM can be written as a function of the Fisher Information Matrix $\mathbf{F}(\boldsymbol{\theta})$ as follows

$$\mathbf{B} = E_{\boldsymbol{\theta}} \; \mathbf{F}(\boldsymbol{\theta}) \; + E_{\boldsymbol{\theta}} \; -\Delta_{\boldsymbol{\theta}}^{\boldsymbol{\theta}} \log p(\boldsymbol{\theta}) \; , \tag{4}$$

$$\mathbf{F}(\boldsymbol{\theta}) = E_{\mathbf{y}|\boldsymbol{\theta}} \; -\Delta_{\boldsymbol{\theta}}^{\boldsymbol{\theta}} \log p(\mathbf{y}|\boldsymbol{\theta}) \; , \tag{5}$$

where $\Delta_{\boldsymbol{\psi}}^{\boldsymbol{\mu}}$ represents the second-order partial derivative operator *i.e.*, $\left[\Delta_{\boldsymbol{\psi}}^{\boldsymbol{\mu}}\right]_{k,l} = \frac{\partial^2}{\partial \psi_k \partial \mu_l}$.

The first term of (4) can be interpreted as the average information with respect to $\boldsymbol{\theta}$ brought by the observations $\mathbf{y}$; on the other hand, the last term can be regarded as the information available from the prior knowledge of $\boldsymbol{\theta}$, *i.e.*, $p(\boldsymbol{\theta})$. This term actually accounts for the time dependence between phase offsets at different instants.

This work is partially supported by the European Community contract NO. 507325, NEWCOM.

3.1. COMPUTATION OF $E_{\boldsymbol{\theta}}[\mathbf{F}(\boldsymbol{\theta})]$

The evaluation of $E_{\boldsymbol{\theta}}[\mathbf{F}(\boldsymbol{\theta})]$ requires the computation of $\mathbf{F}(\boldsymbol{\theta})$. Using the observation model defined in Section 2, the log-likelihood function can be expanded as

$$\log\, p(\mathbf{y}|\boldsymbol{\theta}) = \log \sum_{\mathbf{a}} p(\mathbf{y}|\mathbf{a},\boldsymbol{\theta})\, p(\mathbf{a}). \tag{6}$$

Using the whiteness of the noise and the independence of the transmitted symbols, one then obtains that

$$\Delta_{\boldsymbol{\theta}}^{\boldsymbol{\theta}} \log p\left(\mathbf{y}|\boldsymbol{\theta}\right) = \sum_{k=0}^{K-1} \Delta_{\boldsymbol{\theta}}^{\boldsymbol{\theta}} \log p\left(y_k|\theta_k\right). \tag{7}$$

It is important to note that each term of the sum (7) is a matrix with only one non-zero element, namely,

$$\left[\Delta_{\boldsymbol{\theta}}^{\boldsymbol{\theta}} \log p\left(y_k|\theta_k\right)\right]_{k,k} = \frac{\partial^2}{\partial\theta_k^2} \log p\left(y_k|\theta_k\right). \tag{8}$$

As a direct consequence, $\Delta_{\boldsymbol{\theta}}^{\boldsymbol{\theta}} \log p\left(\mathbf{y}|\boldsymbol{\theta}\right)$ is a diagonal matrix with the k^{th} diagonal element given by equation (8). Moreover, because of the phase model and the Gaussian nature of the noise, $p\left(y_k|\theta_k\right)$ is invariant with respect to the time index k. Then, one has that

$$E_{\boldsymbol{\theta}}\left[\mathbf{F}_K\left(\boldsymbol{\theta}\right)\right] = J_D \mathbf{I}_K, \tag{9}$$

where $\mathbf{I}_K$ is the $K \times K$ identity matrix and where J_D is defined as follows

$$J_D \triangleq E_{\mathbf{y},\boldsymbol{\theta}} \quad -\frac{\partial^2 \log p\left(y_k|\theta_k\right)}{\partial\theta_k^2} \quad . \tag{10}$$

3.2. COMPUTATION OF $E_{\boldsymbol{\theta}}[\Delta_{\boldsymbol{\theta}}^{\boldsymbol{\theta}} \log(p(\boldsymbol{\theta}))]$

Due to the Wiener structure of the phase model (2), $\Delta_{\boldsymbol{\theta}}^{\boldsymbol{\theta}} \log p\left(\boldsymbol{\theta}\right)$ can be expanded as

$$\Delta_{\boldsymbol{\theta}}^{\boldsymbol{\theta}} \log p\left(\boldsymbol{\theta}\right) = \Delta_{\boldsymbol{\theta}}^{\boldsymbol{\theta}} \log p\left(\theta_0\right) + \sum_{k=1}^{K-1} \Delta_{\boldsymbol{\theta}}^{\boldsymbol{\theta}} \log p\left(\theta_k|\theta_{k-1}\right). \tag{11}$$

The first term is a matrix with only one non-zero element, namely, the element (0×0) which is equal to $\left[\Delta_{\boldsymbol{\theta}}^{\boldsymbol{\theta}} \log p\left(\theta_0\right)\right]_{0,0} = \frac{\partial^2 \log p(\theta_0)}{\partial\theta_0^2}$. The other terms in (11) are matrices with only four non-zero elements, namely, $(k-1,k-1)$, $(k-1,k)$, $(k,k-1)$ and (k,k). Due to the Gaussian nature of the noise, one finds that:

$$\left[\Delta_{\boldsymbol{\theta}}^{\boldsymbol{\theta}} \log p\left(\theta_k|\theta_{k-1}\right)\right]_{k,k} = \left[\Delta_{\boldsymbol{\theta}}^{\boldsymbol{\theta}} \log p\left(\theta_k|\theta_{k-1}\right)\right]_{k-1,k-1} = \frac{-1}{\sigma_w^2}$$

$$\left[\Delta_{\boldsymbol{\theta}}^{\boldsymbol{\theta}} \log p\left(\theta_k|\theta_{k-1}\right)\right]_{k,k-1} = \left[\Delta_{\boldsymbol{\theta}}^{\boldsymbol{\theta}} \log p\left(\theta_k|\theta_{k-1}\right)\right]_{k-1,k} = \frac{1}{\sigma_w^2}$$

With these previous expressions, one finally obtains

$$-E_{\boldsymbol{\theta}}[\Delta_{\boldsymbol{\theta}}^{\boldsymbol{\theta}} \log p(\boldsymbol{\theta})] =$$

$$\begin{pmatrix} \frac{1}{\sigma_w^2} - E_{\theta_0}\left[\frac{\partial^2}{\partial\theta_0^2}\log p(\theta_0)\right] & \frac{-1}{\sigma_w^2} & 0 & \dots & 0 \\ \frac{-1}{\sigma_w^2} & \frac{2}{\sigma_w^2} & \frac{-1}{\sigma_w^2} & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & \frac{-1}{\sigma_w^2} & \frac{2}{\sigma_w^2} & \frac{-1}{\sigma_w^2} \\ 0 & \dots & 0 & \frac{-1}{\sigma_w^2} & \frac{1}{\sigma_w^2} \end{pmatrix} \tag{12}$$

In the sequel, for the sake of conciseness, we set $E_{\theta_0}\left[\frac{\partial^2 \log p(\theta_0)}{\partial\theta_0^2}\right] = 0$. This corresponds to the case of a non-informative prior about θ_0.

3.3. ANALYTICAL EXPRESSION OF THE OFF-LINE BCRB

In this subsection, an analytical expression of the diagonal elements of the inverse of the BIM is displayed. These elements lower-bound the MSE achievable by any off-line estimator of the time-varying phase offset θ_k's.

From (9) and (12), the BIM can be written as

$$\mathbf{B}_K = b \begin{pmatrix} A+1 & 1 & & & \\ 1 & A & 1 & & \\ & \ddots & \ddots & \ddots & \\ & & 1 & A & 1 \\ & & & 1 & A+1 \end{pmatrix}, \tag{13}$$

where A and b are defined by $A \triangleq -\sigma_w^2 J_D - 2$ and $b \triangleq -1/\sigma_w^2$. In particular, $\mathbf{B}_K$ is a symmetric sparse matrix. Based on its structure, one finds that the k^{th} diagonal element of $\mathbf{B}_K^{-1}$ can be expressed as (see Appendix A)

$$\begin{aligned} \mathbf{B}_K^{-1}\;{}_{k,k} = &\frac{1}{|\mathbf{B}_K|}\Big[\rho_1^2\left(b+r_1\right)^2 r_1^{K-3} + \rho_2^2(b+r_2)^2 r_2^{K-3} \\ &- \frac{b^2}{A-2}(r_1^{k-1} r_2^{K-k-2} + r_1^{K-k-2} r_2^{k-1})\Big], \end{aligned} \tag{14}$$

where we use the following definitions:

$$r_1 = \frac{b}{2} \quad A + \sqrt{A^2-4} \quad , \qquad r_2 = \frac{b}{2} \quad A - \sqrt{A^2-4} \quad , \tag{15}$$

$$\rho_1 = \frac{\sqrt{1-\frac{4}{A^2}}+1}{2\sqrt{1-\frac{4}{A^2}}}, \qquad \rho_2 = \frac{\sqrt{1-\frac{4}{A^2}}-1}{2\sqrt{1-\frac{4}{A^2}}}. \tag{16}$$

4. J_D EVALUATION AND ASYMPTOTIC CASES

In this section, we briefly discuss the practical evaluation of the BCRB. Then, we derive high and low-SNR approximation of the BCRB which give an alternative to the evaluation of (14).

4.1. BCRB EVALUATION

It is clear from (14)-(16) that the computation of the BCRB requires to evaluate J_D (10). Notice that, using the Gaussian nature of the noise and the equiprobability of the data symbols, the term in the expectation (10) may be written as

$$\begin{aligned} \frac{\partial^2 \log p(y_k|\theta_k)}{\partial\theta_k^2} = &-\frac{2}{\sigma_n^2}\mathrm{Re}(x_k)\tanh \quad \frac{2}{\sigma_n^2}\mathrm{Re}(x_k) \\ &+ \frac{4}{\sigma_n^4}\mathrm{Im}^2(x_k) \quad 1-\tanh^2 \quad \frac{2}{\sigma_n^2}\mathrm{Re}(x_k) \quad , \end{aligned} \tag{17}$$

where $x_k \triangleq y_k\, e^{-j\theta_k}$. Unfortunately, the expectation of (17) with respect to $\mathbf{y}$ and $\boldsymbol{\theta}$ (see (10)) does not have any analytical solution. J_D can however be straightforwardly evaluated via numerical integration methods.

4.2. High-SNR Asymptote

From the definition of the BIM (4), it is clear that only $E_{\boldsymbol{\theta}}[\mathbf{F}(\boldsymbol{\theta})]$ depends on the SNR and from (9) that $E_{\boldsymbol{\theta}}[\mathbf{F}(\boldsymbol{\theta})]$ is fully characterized by J_D (10). Consequently, we can exclusively focus on the high-SNR behavior of J_D to provide a high-SNR asymptote of the BCRB.

At high SNR, $\tanh$ in (17) can be approximated as $\tanh\left[\frac{2}{\sigma_n^2}\mathrm{Re}(x_k)\right] \approx \text{sign}\ \mathrm{Re}(x_k)$. As a consequence one has (see *e.g.*,[2])

$$J_D \approx \frac{2}{\sigma_n\sqrt{\pi}} e^{-\frac{1}{\sigma_n^2}} + \frac{2}{\sigma_n^2}\mathrm{erf}\ \ \frac{1}{\sigma_n}\ \ , \tag{18}$$

where $\mathrm{erf}\,(x) = \frac{2}{\sqrt{\pi}}\int_0^x e^{-t^2}dt$ is the error function. So that $\lim_{\sigma_n^2\to 0} J_D = 2/\sigma_n^2$. Replacing J_D by its limit in (14), we obtain the high-SNR asymptote.

4.3. Low-SNR Asymptote

By the same reasoning as previously, we only focus on J_D. With a first order Taylor expansion of $\tanh(z)$ around $z = 0$, equation (17) becomes

$$\frac{\partial^2 \log p\,(y_k|\theta_k)}{\partial\theta_k^2} \approx - \left(\frac{2}{\sigma_n^2}\mathrm{Re}(x_k)\right)^2 + \left(\frac{2}{\sigma_n^2}\mathrm{Im}(x_k)\right)^2, \tag{19}$$

and thus from (10) one obtains an asymptote simply depending on the observation noise:

$$J_D \approx \frac{4}{\sigma_n^4}. \tag{20}$$

Plugging (20) into (10), we get a low-SNR asymptote of the BCRB.

5. ON-LINE BCRB

Up to this section, we have focused on the off-line scenario, *i.e.*, the receiver waits until it has received all the observation (*i.e.*, y_k with $k = 0, \cdots, K-1$) before estimating the phase offsets. We now show how the previous results can be used in the case of an on-line synchronization mode. In the on-line mode, the receiver estimates θ_k upon the arrival of y_k, *i.e.*, it bases its estimation on the y_l's with $l = 0, ..., k-1$. In order to deal with such scenario, a Posterior (on-line) Cramér-Rao Bound has been derived in [7]. More particularly, the authors provide a method for updating the BIM from the time index k to the time index $k+1$. This method has already been applied to the same scenario as the one considered in this paper in [2].

Using our previous derivations, we can provide an alternative expression of the on-line BCRB. Indeed, the on-line BCRB associated to the observation vector $\mathbf{y}_0^{k-1} \triangleq [y_0 \cdots y_{k-1}]^T$ is clearly equal to element $(k-1, k-1)$ of the inverse of the BIM, *i.e.*, $\left[\mathbf{B}_k^{-1}\right]_{k-1,k-1}$. Using expression (14), we therefore have

$$\left[\mathbf{B}_k^{-1}\right]_{k-1,k-1} = \frac{1}{f_k}\Big[\rho_1^2\,(b+r_1)^2\,r_1^{k-3} + \rho_2^2(b+r_2)^2 r_2^{k-3} - \frac{b^2}{A-2}(r_1^{k-2}r_2^{-1} + r_1^{-1}r_2^{k-2})\Big]. \tag{21}$$

From (21), we can derive (see Appendix B) the asymptote of the on-line BCRB when the number of observations tends to infinity:

$$\lim_{k\to\infty}\left[\mathbf{B}_k^{-1}\right]_{k-1,k-1} = \frac{-\sigma_w^2 + \sqrt{\sigma_w^4 + 4\frac{\sigma_w^2}{J_D}}}{2}. \tag{22}$$

6. DISCUSSION

In this section, we analyse and illustrate the behavior of the different bounds according to both the SNR and the observation number.

We first consider a transmission disturbed by an AWGN with variance $\sigma_n^2 = 0.25$ and phase noise with variance $\sigma_w^2 = 0.04\,\mathrm{rad}^2$. Figure 1 superimposes the on-line BCRB and its asymptote, and the off-line BCRBs for different block-observation lengths versus the time index. Then we obtain the lower bound for each phase offset θ_k according to the considered scenario. In the off-line context, we can see that the best phase estimate is achieved at mid-block, whereas the estimates are likely to be poorer at the block boarder (this result can be proved by analyzing (14)). In this case, all the observations are used to estimate the k^{th} phase offset. Then at time index $\lfloor\frac{K-1}{2}$, the estimate takes equally advantage of the previous and the next half-block of observations. In other words, one better benefits from all the *a priori* information. On the contrary, the first (and respectively the last) estimate can only use the following (respectively previous) observations. The performance is necessarily degraded.

Concerning the on-line bound, at the beginning when the observation number increases, the estimator takes into account more information and the estimation can be improved, so that the bound decreases and converges to its asymptote: the estimation performance is limited by the phase noise and the observation noise independently of the observation number.

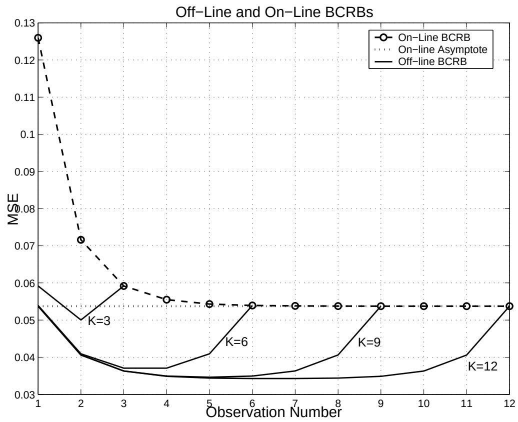

Fig. 1. On-line and Off-line scenarios with different observation block length K with $\sigma_n^2 = 0.25$, $\sigma_w^2 = 0.04\,\mathrm{rad}^2$ and 10^5 Monte-Carlo Trials

We now consider the transmission of a block of $K = 50$ BPSK symbols disturbed by a phase noise with variance $\sigma_w^2 = 0.04\,\mathrm{rad}^2$. Figure 2 superimposes the on-line BCRB and the off-line BCRB evaluated over 10^5 Monte-Carlo integrations. The on-line BCRB is plotted at time index $K = 50$ whereas the off-line BCRB is calculated at mid-block ($k = 25$). This bound can also be compared to its high-SNR and low-SNR asymptotes (which are evaluated at mid-block). We can see that these asymptotes which are simpler to evaluate than the BCRB provide an accurate approximation of the BCRB in the adequate SNR-range. We also note that the off-line BCRB is generally lower than the on-line bound. At high SNR, the bounds are close: the information brought by the current observation is accurate and becomes preponderant over the *a priori* knowledge of $\boldsymbol{\theta}$. Consequently the different observations do not seem to be linked and then we estimate θ_k with the help of y_k only. Then, the on-line and off-line bounds become equivalent.

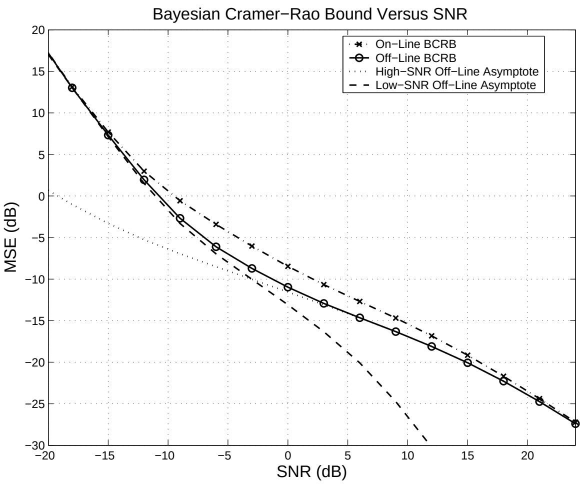

Fig. 2. BCRB versus SNR for $\sigma_w = 0.04\,\text{rad}^2$, $K = 50$ and J_D evaluated over 10^5 Monte-Carlo trials

7. CONCLUSION

In this contribution, we derived a closed-form expression of the BCRB considering an uncoded BPSK transmission and a Wiener phase model in both off-line and on-line context. Taking benefit from our derivations, we then provide asymptotes of the BCRB. In particular, we derived the low and high-SNR asymptotes of the off-line BCRB and the asymptote in terms of number of observations for the on-line scenario. Finally, we compare the BCRB behavior in the on-line and off-line scenarios.

A. APPENDIX - DIAGONAL ELEMENTS OF THE BCRB

In this appendix, we detail the calculus of the diagonal elements (14) of the inverse of the BIM (13). We use the classical matrix-inversion formula

$$[\mathbf{B}_K^{-1}]_{k,k} = C_{k,k}\,|\mathbf{B}_K|^{-1}\,, \tag{23}$$

where $C_{k,k}$ is the cofactor of the element $[\mathbf{B}_K]_{k,k}$ and $|\mathbf{B}_K|$ is the determinant of $\mathbf{B}_K$. We need to compute this cofactor and $|\mathbf{B}_K|$.

- Preliminary calculus. We define the determinant d_k of the following $k \times k$ matrix

$$D_k = b\begin{pmatrix} A & 1 & & & \\ 1 & A & 1 & & \\ & \ddots & \ddots & \ddots & \\ & & 1 & A & 1 \\ & & & 1 & A \end{pmatrix}$$

  Expanding d_k along the first and the second column, one obtains $d_k = Abd_{k-1} - b^2 d_{k-2}$ with $d_0 = 1$ and $d_1 = bA$. $\{d_k\}$ is thus a linear recurrent sequence. Using the initial terms, an analytical expression of d_k is given by

$$d_k = \rho_1\,(r_1)^k + \rho_2\,(r_2)^k \qquad \text{for } k \in \mathbb{N}, \tag{24}$$

  where ρ_1, ρ_2, r_1 and r_2 are defined by (15) and (16).

- Determinant $|\mathbf{B}_K|$.
  In a first step, by expanding $|\mathbf{B}_K|$ along the first column, we obtain a sum of two cofactors which are then expanded along the last column. Doing so, one has that $|\mathbf{B}_K| = (A+2)\,bd_{K-1}$.

- Cofactor $C_{k,k}$. In order to calculate the cofactor (k,l), one has to delete the row k and the column l of the matrix $\mathbf{B}_K$ and one obtains a two-block-diagonal matrix. The upper (respectively lower) block is noted U_{B_K} (respectively L_{B_K}). The cofactor is thus the product of two determinants:

$$\begin{aligned} C_{k,k} &= |U_{B_K}|\,|L_{B_K}| \\ &= (bd_{k-1} + d_k)\,(bd_{K-k-2} + d_{K-k-1})\,. \end{aligned} \tag{25}$$

- Diagonal elements. Rewriting (23) and using (24) and (25), one has

$$\begin{aligned} \left[\mathbf{B}_K^{-1}\right]_{k,k} = \frac{1}{|\mathbf{B}_K|}\Big[&\rho_1^2\,(b+r_1)^2\,r_1^{K-3} + \rho_2^2(b+r_2)^2 r_2^{K-3} \\ &- \frac{b^2}{A-2}(r_1^{k-1}r_2^{K-k-2} + r_1^{K-k-2}r_2^{k-1})\Big]. \end{aligned}$$

B. APPENDIX - BEHAVIOR OF THE ON-LINE BCRB

Using (23) and (25), $\left[\mathbf{B}_K^{-1}\right]_{K-1,K-1}$ can be written as

$$\left[\mathbf{B}_K^{-1}\right]_{K-1,K-1} = \frac{A+1}{b\,(A+2)} - \frac{1}{b\,(A+2)}u_K, \tag{26}$$

where $u_n \triangleq \frac{d_n}{bd_{n-1}} = \frac{Au_{n-1}-1}{u_{n-1}}$ with $u_1 = A$. Clearly, this sequence is strictly increasing and converges to $u_\infty = \frac{1}{2}\left(A - \sqrt{A^2-4}\right)$. Combining this result with (26), we have that $\left[\mathbf{B}_K^{-1}\right]_{K-1,K-1}$ is a strictly deacreasing sequence with the following limit:

$$\lim_{K\to\infty}\left[\mathbf{B}_K^{-1}\right]_{K-1,K-1} = \frac{-\sigma_w^2 + \sqrt{\sigma_w^4 + 4\frac{\sigma_w^2}{J_D}}}{2}. \tag{27}$$

C. REFERENCES

[1] Heinrich Meyr, Marc Moeneclaey, and Stefan Fechtel, *Digital Communication Receivers: Synchronization, Channel Estimation, and Signal Processing*, John Wiley & Sons, Inc., New York, NY, USA, 1997.

[2] 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.

[3] J. Dauwels, *On Graphical Models for Communications and Machine Learning: Algorithms, Bounds, and Analog Implementation*, Ph.D. thesis, Swiss Federal Institute of Technology, 2005.

[4] J. Dauwels and S. Korl, "A numerical method to compute Cramér-Rao-type bounds for challenging estimation problems," *Proc. IEEE Int. Conf. Acoust., Speech, Signal Processing*, pp. 717–720, May 2006.

[5] 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.

[6] H. L. Van Trees, *Detection, Estimation and Modulation Theory*, vol. 1, Wiley, New York, 1968.

[7] P. Tichavsky, C. H. Muravchik, and A. Nehorai, "Posterior Cramér-Rao bounds for discrete-time nonlinear filtering," *IEEE Trans. Signal Processing*, vol. 46, pp. 1386–1396, May 1998.