TRUE CRAMER-RAO BOUND FOR ESTIMATING SYNCHRONIZATION PARAMETERS FROM A LINEARLY MODULATED BANDPASS SIGNAL WITH UNKNOWN DATA SYMBOLS

N. Noels and M. Moeneclaey

DIGCOM Research Group, TELIN Department, Ghent University, St.-Pietersnieuwstraat 41
B-9000 Gent, Belgium
nnoels,mm@telin.UGent.be

ABSTRACT

This paper considers the Cramer-Rao bound (CRB) for estimating the synchronization parameters from a linearly modulated waveform with unknown data symbols. The key idea is to reduce the computational complexity associated with the CRB by exploiting the specific structure of the observed signal, see e.g., [1, 2] and references therein. The current contribution builds on this previous work, but in this case, no distributional assumptions are made about the data symbols. A generalized closed-form expression of the CRB is presented, the resulting computational complexity is discussed, and numerical results are provided.

1. INTRODUCTION

The Cramer-Rao bound (CRB) provides a fundamental lower bound on the achievable mean square estimation error (MSEE), and as such serves as a useful benchmark for practical estimators [3]. The present paper specifically focuses on estimating synchronization parameters from a linearly modulated waveform. The presence of random data symbols makes the computation of the CRB very hard. In [4], the high signal-to-noise ratio (SNR) limit of the CRB has been evaluated analytically. However, in the presence of coding, synchronization algorithms must operate at low SNR, so that this high-SNR limit of the CRB might no longer be a relevant benchmark. Several numerical procedures, enabling the efficient evaluation of the CRB at an arbitrary SNR, have been proposed over the last decade: [1, 2] and references therein. In order to unify and generalize these contributions, we consider in the next section an arbitrarily distributed data symbol vector. Section 3 subsequently addresses the relation between various estimation modes and the associated distribution of the data symbol vector to be used when evaluating the CRB. In Section 4, a generalized closed-form expression of the CRB is presented, and the resulting computational complexity is discussed. Section 5 provides numerical results evaluated for some particular cases, pertaining to the estimation of the carrier phase from a noisy linearly modulated signal with encoded data and known pilots. The main conclusions are listed in Section 6.

This work was supported by the FWO project G.0047.06 "Advanced space-time processing techniques for communication through multi-antenna systems in realistic mobile channels.

2. PROBLEM FORMULATION

We consider the transmission of a linearly modulated signal, obtained by applying a burst of N data symbols to a square-root Nyquist transmit filter, over a bandlimited additive white Gaussian noise (AWGN) channel. The complex baseband representation of the received noisy signal is given by

$$r(t) = \sqrt{E_s}\sum_{k=1}^{N} a_k h(t - kT - \tau)\, e^{j(2\pi F t + \theta)} + w(t) \quad (1)$$

where $\mathbf{a} = (a_1 a_2 ... a_N)$ are the transmitted symbols with $E[a_k^* a_l] = \delta_{k-l}$, $h(t)$ is the transmit pulse and $w(t)$ is complex-valued zero-mean AWGN with independent real and imaginary parts each having a power spectral density equal to $N_0/2$. The quantity E_s in (1) denotes the average symbol energy, and T is the symbol period. Finally, the parameters F, θ and τ stand for the frequency offset, the phase shift and the time delay of the received signal vis-a-vis the local reference carrier.

Assuming that the channel has an ideal lowpass transfer function with bandwidth B, the signal $r(t)$ is sampled at a rate $\frac{1}{T_s} = 2B$, and scaled by a factor $\sqrt{\frac{T_s}{E_s}}$. This yields the observation vector $\mathbf{r}$, with

$$\mathbf{r} = \mathbf{a}\mathbf{S}(\mathbf{u}) + \mathbf{w} \quad (2)$$

where $\mathbf{w}$ is a complex-valued white Gaussian noise vector with zero mean and $E[w_i^* w_j] = N_0 \delta_{i-j}$. It follows from (1) that the matrix $\mathbf{S}(\mathbf{u})$ satisfies the following property: $\mathbf{S}(\mathbf{u})\mathbf{S}^H(\mathbf{u}) = \mathbf{I}_N$, where $\mathbf{I}_k$ represents a $k \times k$ identity matrix. The column vector $\mathbf{u} = (u_1 u_2 ... u_n)^T$ in (2) contains one or more unknown synchronization parameters (F, θ, τ) and is assumed to be uniformly distributed over its domain.

From the observation vector $\mathbf{r}$, we want to recover the value of $\mathbf{u}$. A popular criterion to derive an estimate $\hat{\mathbf{u}}$ of $\mathbf{u}$ is the a posteriori probability maximization [3]. When $\mathbf{u}$ is uniformly distributed over its domain, estimates satisfying this criterion may be computed via the simpler maximum likelihood criterion, i.e.,

$$\hat{\mathbf{u}} = \arg\max_{\tilde{\mathbf{u}}} \ln p\left(\mathbf{r}\left|\tilde{\mathbf{u}}\right.\right) \tag{3}$$

A common approach to evaluate the quality of an estimator of $\mathbf{u}$ consists in comparing its resulting MSEE with the CRB. The latter imposes a fundamental lower bound on the MSEE of *any* estimator, and results from the inequality $\mathbf{R_u} - \mathbf{J}^{-1} \geq \mathbf{0}$ [3]. Here, $\mathbf{R_u}$ is the error correlation matrix related to the estimation of $\mathbf{u}$, the notation $\mathbf{A} \geq \mathbf{0}$ indicates that $\mathbf{A}$ is a positive semi-definite matrix, and $\mathbf{J}^{-1}$ denotes the inverse of the Fisher Information Matrix (FIM) $\mathbf{J}$. Taking into account that $\mathbf{u}$ has a uniform a priori distribution, the (i,j)-th element of $\mathbf{J}$ is given by

$$J_{i,j} = E\left[\ell_i\left(\mathbf{u};\mathbf{r}\right)\ell_j\left(\mathbf{u};\mathbf{r}\right)\right] \tag{4}$$

where $E[.]$ denotes averaging with respect to $p\left(\mathbf{r},\mathbf{u}\right)$, and

$$\ell_i\left(\mathbf{u};\mathbf{r}\right) = \frac{\partial \ln p\left(\mathbf{r}\left|\mathbf{u}\right.\right)}{\partial u_i} \tag{5}$$

is a short-hand notation for the derivative of $\ln p\left(\mathbf{r}\left|\mathbf{u}\right.\right)$ with respect to the i-th parameter u_i. It easily follows from $\mathbf{R_u} - \mathbf{J}^{-1} \geq \mathbf{0}$ that

$$E\left[\left(u_i - \hat{u}_i\right)^2\right] \geq CRB_i \tag{6}$$

where CRB_i is the i-th diagonal element of the inverse of the FIM $\mathbf{J}$. The right-hand side of (6) is referred to as the CRB.

3. ESTIMATION MODES AND THE CRB

The quantity $p\left(\mathbf{r}\left|\mathbf{u}\right.\right)$, in (3) and (4)-(5), can be rewritten as

$$p\left(\mathbf{r}\left|\mathbf{u}\right.\right) = \sum_{\tilde{\mathbf{a}}} p\left(\mathbf{r}\left|\mathbf{a}=\tilde{\mathbf{a}},\mathbf{u}\right.\right)\Pr\left[\mathbf{a}=\tilde{\mathbf{a}}\right] \tag{7}$$

where $\Pr\left[\mathbf{a}=\tilde{\mathbf{a}}\right]$ denotes the probability that the transmitted data symbol vector equals $\tilde{\mathbf{a}}$. In principle, estimators should take into account as much statistical information about $\mathbf{a}$ as possible. Unfortunately, the complexity associated with using the correct data symbol a priori probability may often turn out to be too large. For example, the evaluation of (7) in the presence of a non-trivial channel code gives rise to a computational burden that is exponential in the size N of the data burst. A solution to simplify the estimator implementation consists in dropping/approximating some parts of the available statistical information. It should be noted that the MSEE resulting from estimation algorithms that rely on such an approximation of $\Pr\left[\mathbf{a}=\tilde{\mathbf{a}}\right]$ is lower bounded by the CRB that is computed according to (4) but with the same approximate hypothesis about $\Pr\left[\mathbf{a}=\tilde{\mathbf{a}}\right]$.

4. DERIVING THE CRB

A brute force numerical evaluation of the FIM involves replacing in (4) the statistical average $E\left[.\right]$ by an arithmetical average over a large number of realizations of $\left(\mathbf{r},\mathbf{u}\right)$, that are computer-generated according to the joint distribution of $\mathbf{r}$ and $\mathbf{u}$. The numerical evaluation of the FIM further requires the computation of the derivatives $\ell_i\left(\mathbf{u};\mathbf{r}\right)$, $i=1,2$ that correspond to the realizations of $\left(\mathbf{r},\mathbf{u}\right)$. Making use of (7), these derivatives can be put into the following form

$$\ell_i\left(\mathbf{u};\mathbf{r}\right) = \sum_{\tilde{\mathbf{a}}} \frac{\partial \ln p\left(\mathbf{r}\left|\mathbf{a}=\tilde{\mathbf{a}},\mathbf{u}\right.\right)}{\partial u_i}\Pr\left[\mathbf{a}=\tilde{\mathbf{a}}\left|\mathbf{r},\mathbf{u}\right.\right] \tag{8}$$

As $p\left(\mathbf{r}\left|\mathbf{a},\mathbf{u}\right.\right)$ is Gaussian, the logarithm $\ln p\left(\mathbf{r}\left|\mathbf{a},\mathbf{u}\right.\right)$ is readily available in closed-form:

$$\ln p\left(\mathbf{r}\left|\mathbf{a},\mathbf{u}\right.\right) = -\frac{E_s}{N_0}\left|\mathbf{r}-\mathbf{a}\mathbf{S}\left(\mathbf{u}\right)\right|^2 \tag{9}$$

Differentiating with respect to u_i yields

$$\begin{aligned}&\frac{\partial \ln p\left(\mathbf{r}\left|\mathbf{a},\mathbf{u}\right.\right)}{\partial u_i} \\ &= -\frac{2E_s}{N_0}\Re\left\{\left(\mathbf{r}-\mathbf{a}\mathbf{S}\left(\mathbf{u}\right)\right)^H\left(\mathbf{a}\frac{\partial \mathbf{S}\left(\mathbf{u}\right)}{\partial u_i}\right)\right\}\end{aligned} \tag{10}$$

The joint symbol a posteriori probabilities (APP) $\Pr\left[\mathbf{a}\left|\mathbf{r},\mathbf{u}\right.\right]$ in (8) can be computed from $p\left(\mathbf{r}\left|\mathbf{a},\mathbf{u}\right.\right)$ and $\Pr\left[\mathbf{a}\right]$, according to

$$\Pr\left[\mathbf{a}\left|\mathbf{r},\mathbf{u}\right.\right] = \frac{\Pr\left[\mathbf{a}\right]p\left(\mathbf{r}\left|\mathbf{a},\mathbf{u}\right.\right)}{\sum_{\tilde{\mathbf{a}}} p\left(\mathbf{r}\left|\mathbf{a}=\tilde{\mathbf{a}},\mathbf{u}\right.\right)\Pr\left[\mathbf{a}=\tilde{\mathbf{a}}\right]} \tag{11}$$

Although this procedure yields the exact derivatives $\ell_i\left(\mathbf{u};\mathbf{r}\right)$, $i=1,2$, the summations in (8) and (11) gives rise to a computational complexity that is exponential in the burst size N. The computational complexity associated with the evaluation of these derivatives can be drastically reduced by taking into account the specific structure of the useful signal in (2). Because $\mathbf{S}\left(\mathbf{u}\right)\mathbf{S}^H\left(\mathbf{u}\right) = \mathbf{I}_N$ does not depend on $\mathbf{u}$, we obtain

$$\frac{\partial \ln p\left(\mathbf{r}\left|\mathbf{a},\mathbf{u}\right.\right)}{\partial u_i} = \frac{2E_s}{N_0}\Re\left\{\mathbf{r}\left(\frac{\partial \mathbf{S}\left(\mathbf{u}\right)}{\partial u_i}\right)^H\mathbf{a}^H\right\} \tag{12}$$

Substituting (12) into (8), then yields

$$\ell_i\left(\mathbf{u};\mathbf{r}\right) = \frac{2E_s}{N_0}\Re\left\{\mathbf{r}\left(\frac{\partial \mathbf{S}\left(\mathbf{u}\right)}{\partial u_i}\right)^H\boldsymbol{\mu}^H\left(\mathbf{r},\mathbf{u}\right)\right\} \tag{13}$$

where $\boldsymbol{\mu}\left(\mathbf{r},\mathbf{u}\right)$ is a short-hand notation for the a posteriori average of the symbol vector $\mathbf{a}$, with N components

$$\mu\left(k;\mathbf{r},\mathbf{u}\right) = E_{\mathbf{a}}\left[a_k\left|\mathbf{r},\mathbf{u}\right.\right] \tag{14}$$

$$= \sum_{\omega\in\Omega}\Pr\left[a_k=\omega\left|\mathbf{r},\mathbf{u}\right.\right]\omega \tag{15}$$

where Ω denotes the set of constellation points and the averaging $E_{\mathbf{a}}[.|\mathbf{r},\mathbf{u}]$ is with respect to $\Pr[\mathbf{a}|\mathbf{r},\mathbf{u}]$.

We emphasize that no approximation is involved in obtaining (15); the right hand side simply expresses the a posteriori average of the k-th data symbol a_k in terms of the *marginal* APP of a_k, rather than the *joint* APP of all components of $\mathbf{a}$. Computing the marginal APPs from the the joint APP still requires a complexity that increases exponentially with N. However, in most practical scenarios, the required marginal symbol APPs can be directly computed in an efficient way, by applying the sum-product algorithm to a factor graph representing a suitable factorization of the joint symbol APP [5]. The application of the sum-product algorithm on a graph that corresponds to a tree is straightforward and yields the exact marginals. When the graph contains cycles, the sum-product algorithm becomes an iterative procedure that, after convergence, yields only an approximation of the marginals. However, when the cycles in the graph are large, the resulting marginals turn out to be quite accurate. Computing the derivatives $\ell_i(\mathbf{u};\mathbf{r})$, $i=1,2$ according to (12), for a given realization of $(\mathbf{r},\mathbf{u})$, yields a complexity that is linear (and not exponential) in the number of data symbols N.

5. NUMERICAL EXAMPLE

As a case study, we consider a turbo-coded 4-PSK transmission scheme. The rate 1/2 turbo encoder encompasses parallel concatenation of two identical binary 16-state rate-1/2 recursive systematic convolutional encoders with generator polynomials $(21)_8$ and $(37)_8$ in octal notation, via a pseudo random interleaver with block length N_d information bits, and an appropriate puncturing pattern so that the block at the turbo encoder output comprises $2N_d$ coded bits. This binary turbo code is followed by conventional Gray-mapped 4-PSK modulation, giving rise to a block of N_d random data symbols. Finally, a preamble of N_p known pilot symbols with value 1 is added at the beginning of each block; the transmitted symbol vector has the following a priori distribution

$$\Pr[\mathbf{a}=\tilde{\mathbf{a}}]=\begin{cases} 2^{-N_d} & ,\tilde{\mathbf{a}}\in\mathcal{S}_0 \\ 0 & ,\tilde{\mathbf{a}}\in\mathcal{S}\backslash\mathcal{S}_0\end{cases} \tag{16}$$

where $\mathcal{S}$ denotes the set of all possible vectors of (N_d+N_p) 4-PSK symbols, and $\mathcal{S}_0\subset\mathcal{S}$ denotes the subset of these vectors that result from applying a sequence of N_d bits to the above turbo encoder and mapper. Synchronization parameter estimation algorithms that take (16) into account are said to operate in a *pilot-aided code-aided* (PA CA) mode. Three other approximate operation modes are commonly considered in the framework of parameter estimation: the *data-aided* (DA), the *pilot-aided non-code-aided* (PA NCA) and the *non-pilot-aided non-code-aided* (NPA NCA) modes. The DA approach consists in computing the estimate $\hat{\mathbf{u}}$ as if only the pilot symbols had been transmitted, i.e., by assuming that

$$\Pr[\mathbf{a}=\tilde{\mathbf{a}}]=\prod_{k=1}^{N_p}I[\tilde{a}_k=1]\prod_{k=N_{p+1}}^{N_d+N_p}I[\tilde{a}_k=0],\ \forall\tilde{\mathbf{a}}\in\mathcal{S} \tag{17}$$

where $I[.]$ is the indicator function, i.e., $I[P]=1$ when P is true and $I[P]=0$ otherwise, where P is a proposition. The PA NCA and the NPA NCA estimators also exploit the random data symbols in the estimation process, but neglect the structure that the channel code enforces upon them. The PA NCA estimators make use of the knowledge about the pilot symbols, and assume that the N_d random data symbols are statistically independent and equiprobable; this yields

$$\Pr[\mathbf{a}=\tilde{\mathbf{a}}]=2^{-2N_d}\prod_{k=1}^{N_p}I[\tilde{a}_k=1],\ \forall\tilde{\mathbf{a}}\in\mathcal{S} \tag{18}$$

The NPA NCA estimators also ignore the fact that the receiver knows the value of the pilots, and assume that all (N_d+N_p) data symbols are statistically independent and equiprobable, i.e.,

$$\Pr[\mathbf{a}=\tilde{\mathbf{a}}]=2^{-2(N_d+N_p)},\ \forall\tilde{\mathbf{a}}\in\mathcal{S} \tag{19}$$

We assume that perfect clock and carrier frequency information is available at the receiver, so that the phase shift θ is the only synchronization parameter left to estimate. We compute the corresponding CRB from (4) for the PA CA, the DA, the PA NCA and the NPA NCA estimation modes: $E\left[\left(\theta-\hat{\theta}\right)^2\right]\geq\beta=\frac{1}{J_{1,1}}$. In (2) and subsequent equations: $\mathbf{u}=\theta$ and $\mathbf{S}(\mathbf{u})=e^{j\theta}$. The DA assumption enables the derivation of a closed-form expression of the CRB, i.e.,

$$E\left[\left(\theta-\hat{\theta}\right)^2\right]\geq\beta_{DA}=\frac{1}{2N_p}\frac{N_0}{E_s} \tag{20}$$

In the PA CA, the PA NCA and the NPA NCA scenario, the evaluation of the CRB is considerably more difficult than in the DA scenario, because of the summation over $\tilde{\mathbf{a}}$ in (7). In all three scenarios, we obtain at high SNR [4]:

$$E\left[\left(\theta-\hat{\theta}\right)^2\right]\geq\beta_{\infty}=\frac{1}{2(N_d+N_p)}\frac{N_0}{E_s},\ \frac{E_s}{N_0}\to\infty \tag{21}$$

Comparing (20) with (21), it follows this high-SNR limit equals the DA CRB that would be obtained if all (N_d+N_p) data symbols were known to the receiver. To compute the PA CA, the PA NCA and the NPA NCA CRB at an arbitrary SNR, we resort to a numerical evaluation. For N_d=1000 and N_p=100, Fig. 1 plots the resulting ratio β/β_{∞}(the left ordinate) along with the bit error rate (BER) corresponding to the coherent detection of the transmitted information bits (right ordinate) as a function of E_s/N_0. The low-SNR NPA NCA assymptotical CRB (ACRB) [6] (right ordinate, relative to β_{∞}) and the BER corresponding to the coherent detection of uncoded bits (right ordinate) are also displayed .

It is clear from Fig. 1 that the DA CRB is larger than the PA CA, the PA NCA or the NPA NCA CRB, for all practical values of the SNR (say, BER $< 10^{-2}$). This is consitent with the fact that the DA estimation mode only allows to take benefit from the received signal energy associated with the pilots. From (20), we may notice that lowering the DA CRB implies to increase the number of pilots, which leads to losses in terms of power and spectral efficiency.

Comparing the PA CA, the PA NCA and the NPA NCA modes, we see that the corresponding CRBs are equal at sufficiently high SNRs (and given by the high-SNR limit β_∞ from (21)). At lower SNRs, however, there is a gap between the PA CA CRB on one hand, and the PA NCA and the NPA NCA CRBs on the other hand. When E_s/N_0 decreases, a point $(E_s/N_0)_{thr}$ is reached where the CRBs for the PA CA, the PA NCA and the NPA NCA mode start to diverge from their high-SNR limit. For the PA CA mode, $(E_s/N_0)_{thr}$ corresponds to a coded BER of about 10^{-3}. For the PA NCA and the NPA NCA mode, $(E_s/N_0)_{thr}$ corresponds to an uncoded BER of about 10^{-3}, and consequently exceeds $(E_s/N_0)_{thr}$ for PA CA operation by an amount equal to the coding gain. When E_s/N_0 further decreases, the PA CA and the PA NCA CRBs eventually converge to the DA CRB, whereas the NPA NCA CRB converges to the NPA NCA ACRB. The latter assymptot has been analytically computed in [6].

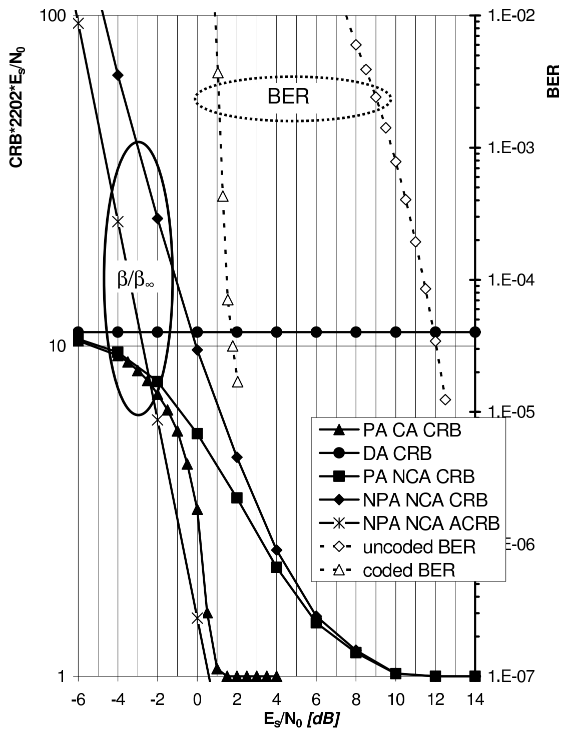

Fig. 1. The CRBs related to carrier phase estimation from N_d=1000 coded 4-PSK and N_p=100 known pilots.

6. CONCLUSIONS AND REMARKS

The paper has discussed the computation of the CRB related to synchronization parameter estimation from a linearly modulated bandpass signal. The most general case of an arbitrarily distributed data symbol vector has been considered. It has been emphasized that the different estimation modes that are commonly considered correspond to distinct hypotheses about the a priori distribution of the transmitted data symbols. It has also been demonstrated how the computational complexity related to the CRB can be significantly reduced by exploiting the structure of the received signal vector. The major conclusion that can be drawn from the reported numerical results is that the CRB that takes into account the whole available a priori information about the data symbols is very well approximated by its high-SNR limit, at the normal operating SNR of the communication system. This high-SNR limit is easily evaluated analytically. Given the estimation mode, the MSEE performance of the maximum-likelihood estimator assymptotically tends to the associated CRB. The results therefore also indicate that synchronizers that do not exploit the code properties might fail to provide accurate estimates when using powerful codes with large coding gains (and low operating SNR).

7. REFERENCES

[1] N. Noels, H. Steendam, M. Moeneclaey and H. Bruneel. ”Carrier Phase and Frequency Estimation for Pilot-Symbol Assisted Transmission: Bounds and Algorithms”. *IEEE Transactions on Signal Processing*, 53(12):4578–4587, December 2005.

[2] N. Noels, H. Steendam and M. Moeneclaey. ”Carrier and Clock recovery in (turbo) coded systems: Cramer-Rao Bound and Synchronizer Performance”. *EURASIP Journ. on Appl. Sig. Proc.*, 2005(6):972–980, May 2005.

[3] H.L. Van Trees. *Detection, Estimation, and Modulation Theory, Part I.* Wiley and Sons, October 2001.

[4] M. Moeneclaey. ”On the true and the modified Cramer-Rao bounds for the estimation of a scalar parameter in the presence of nuisance parameters”. *IEEE Trans. on Comm.*, 46:1536–1544, November 1998.

[5] H.-A. Loeliger. ”An introduction to factor graphs”. *IEEE Signal Processing Magazine*, 21(1):28–41, January 2004.

[6] S. Steendam and M. Moeneclaey. ”Low-SNR limit of the Cramer-Rao bound for estimating the carrier phase and frequency of a PAM, PSK or QAM waveform”. *IEEE Comm. Lett.*, 5:215–217, May 2001.