The Cramer-Rao Bound for Phase Estimation From Coded Linearly Modulated Signals

Nele Noels, *Student Member, IEEE*, Heidi Steendam, *Member, IEEE*, and Marc Moeneclaey, *Fellow, IEEE*

***Abstract*—In this letter, we express the Cramer-Rao Bound (CRB) for carrier phase estimation from a noisy linearly modulated signal with *encoded* data symbols, in terms of the *marginal a posteriori* probabilities (APPs) of the coded symbols. For a wide range of classical codes (block codes, convolutional codes, and trellis-coded modulation), these marginal APPs can be computed efficiently by means of the Bahl–Cocke–Jelinke–Raviv (BCJR) algorithm, whereas for codes that involve interleaving (turbo codes and bit interleaved coded modulation), iterated application of the BCJR algorithm is required. Our numerical results show that when the BER of the coded system is less than about 10^{-3}, the resulting CRB is essentially the same as when transmitting a training sequence.**

***Index Terms*—Carrier synchronization, channel coding, Cramer-Rao lower bound, phase estimation.**

I. Introduction

THE Cramer-Rao Bound (CRB) is a lower bound on the error variance of any unbiased estimate, and as such serves as a useful benchmark for practical estimators [1]. In many cases, the statistics of the observed vector depend not only on the parameter to be estimated, but also on a number of nuisance parameters we do not want to estimate. In order to avoid the computational complexity caused by the nuisance parameters, a *modified* CRB (MCRB) has been derived in [2]. The MCRB is much simpler to evaluate than the CRB, but is in general looser than the CRB. The CRB related to carrier phase and frequency estimation has been derived assuming transmission of *uncoded* symbols from a PSK [3]–[5] or a symmetric QAM [6] constellation.

In this contribution we investigate the CRB related to the estimation of the carrier phase of a noisy linearly modulated signal *in the presence of coding*. We derive an expression for the CRB in terms of the marginal a posteriori probabilities (APPs) of the coded symbols, and discuss its numerical evaluation. In Section IV, we present numerical results for the CRB resulting from QPSK transmission with convolutional coding and turbo coding, and compare them with the CRB for uncoded QPSK transmission from [3]–[5] and with the MCRB from [2].

Manuscript received November 12, 2002. The associate editor coordinating the review of this letter and approving it for publication was Dr. E. Serpedin. This work has been supported by the Interuniversity Attraction Poles Program, Belgian Sate—Federal Office for Scientific Technical and Cultural Affairs.

The authors are with the Department of Telecommunications and Information Processing, Ghent University, Gent, B-9000 Belgium (e-mail: {nnoels@ telin.rug.ac.be; hs@ telin.rug.ac.be; mm@ telin.rug.ac.be).

Digital Object Identifier 10.1109/LCOMM.2003.812168

II. Problem Formulation

We consider a linearly modulated signal, obtained by applying an encoded data symbol sequence to a square-root Nyquist transmit filter, that is transmitted over an additive white Gaussian noise channel. The resulting noisy signal is applied to a receiver filter, matched to the transmit filter. The receiver filter output signal is sampled at the correct decision instants (assuming that accurate timing estimation has been established before), which yields the observation vector $\boldsymbol{r} = (r_{-K}, \ldots, r_K)$, with

$$r_k = \varepsilon a_k \exp(j\theta) + w_k \tag{1}$$

for $k = -K, \ldots, K$. In (1), $\{a_k\}$ is a sequence of $L = 2K + 1$ data symbols taken from an M-PSK, M-PAM or M-QAM constellation $((E[|a_k|^2] = 1)$ according to a combination of an encoding rule and a mapping rule. The sequence $\{w_k\}$ consists of independent identically distributed zero-mean complex Gaussian noise variables, with independent real and imaginary parts each having a variance of 1/2. The (unknown but deterministic) parameter θ represents the carrier phase. Finally, $\varepsilon = (E_s/N_0)^{1/2}$, with E_s and N_0 denoting the energy per coded symbol and the noise power spectral density, respectively.

Suppose that one is able to produce from the observation vector $\mathbf{r}$ an unbiased estimate $\hat{\theta}$ of the deterministic parameter θ. Then the estimation error variance is lower bounded by the CRB: $E_r[(\hat{\theta} - \theta)^2] \geq \mathrm{CRB}_\theta$, where CRB_θ is given by [1]

$$\mathrm{CRB}_\theta = E_r\left[\left(\frac{\partial}{\partial\theta}\ln\left(p\left(\mathbf{r};\theta\right)\right)\right)^2\right]^{-1}. \tag{2}$$

The probability density $p(\mathbf{r};\theta)$ of $\mathbf{r}$, corresponding to a given value of θ, is called the likelihood function of θ, while $\ln(p(\mathbf{r};\theta))$ is the log-likelihood function of θ. The expectation $E_r[\cdot]$ in (2) is with respect to $p(\mathbf{r};\theta)$.

Since the observation $\mathbf{r}$ depends not only on the parameter θ to be estimated but also on the nuisance vector $\boldsymbol{a}$ of random data symbols, the likelihood function of θ is obtained by averaging the likelihood function $p(\mathbf{r}|\mathbf{a};\theta)$ of the vector $(\theta, \mathbf{a})$, over the *a priori* distribution of the nuisance parameter vector $\mathbf{a}$. The log-likelihood function $\ln(p(\mathbf{r};\theta))$ of θ is given by

$$\ln(p(\mathbf{r};\theta)) = \ln(E_a[p(\mathbf{r}|\mathbf{a};\theta)]) \tag{3}$$

where, within a factor not depending on θ

$$p(\mathbf{r}|\mathbf{a};\theta) = \prod_{k=-K}^{K} \exp\left(2\varepsilon \mathrm{Re}\left(a_k^* r_k \exp(-j\theta)\right) - \varepsilon^2 \left|a_k\right|^2\right). \tag{4}$$

Computation of the CRB requires the substitution of (3) into (2), and the evaluation of the various expectations included in (2) and (3).

As the evaluation of the expectations involved in (2) and (3) is quite tedious (in particular for coded data sequences), a simpler but looser lower bound, the MCRB, has been derived in [2], i.e., $E_r[(\hat{\theta}-\theta)^2] \geq \mathrm{CRB}_\theta \geq \mathrm{MCRB}_\theta$. The MCRB for phase estimation is given by [2].

$$\mathrm{MCRB}_\theta = \frac{1}{2L}\frac{N_0}{E_s}. \tag{5}$$

The expression (5) equals the CRB that corresponds to the transmission of a long sequence of known training symbols. In [7] it has been shown that the high signal-to-noise ratio (SNR) limit of the CRB coincides with the MCRB from (5).

III. Evaluation of the CRB

A. CRB in Terms of the APPs of the Data Symbols

We obtain for the log-likelihood function $\ln(p(\mathbf{r};\theta))$ from (3)

$$\ln(p(\mathbf{r};\theta)) = \ln\left(\sum_{i=0}^{M^L-1} \Pr[\mathbf{a}=\mathbf{c}_i]p(\mathbf{r}|\mathbf{c}_i;\theta)\right) \tag{6}$$

where i enumerates all M^L symbol sequences $\{c_{i,k}\}$ of length L. Denoting by ξ the set of legitimate coded sequences of length L, we have $\Pr[\mathbf{a}=\mathbf{c}_i] = M^{-rL}$ for $\mathbf{c}_i \in \xi$ and $\Pr[\mathbf{a}=\mathbf{c}_i] = 0$ otherwise, with r and M denoting the rate of the code and the number of constellation points, respectively. Differentiating (6) with respect to θ yields

$$\frac{d}{d\theta}\ln(p(\mathbf{r};\theta)) = \sum_{i=0}^{M^L-1} \frac{\Pr[\mathbf{a}=\mathbf{c}_i]p(\mathbf{r}|\mathbf{c}_i;\theta)}{p(\mathbf{r};\theta)}\frac{d}{d\theta}\ln(p(\mathbf{r}|\mathbf{c}_i;\theta)) \tag{7}$$

Made use of Bayes' rule, i.e.,

$$\frac{\Pr[\mathbf{a}=\mathbf{c}_i]p(\mathbf{r}|\mathbf{c}_i;\theta)}{p(\mathbf{r};\theta)} = \Pr[\mathbf{a}=\mathbf{c}_i|\mathbf{r};\theta] \tag{8}$$

and of (4), (7) is transformed into

$$\begin{aligned}\frac{d}{d\theta}\ln(p(\mathbf{r};\theta)) &= 2\varepsilon E_{\mathbf{a}|\mathbf{r}}\left[\sum_{k=-K}^{K}\mathrm{Im}\left(a_k^* r_k e^{-j\theta}\right)\right] \\ &= 2\varepsilon \sum_{k=-K}^{K} E_{a_k|\mathbf{r}}\left[\mathrm{Im}\left(a_k^* r_k e^{-j\theta}\right)\right] \\ &= 2\varepsilon \sum_{k=-K}^{K}\sum_{m=0}^{M-1}\Pr[a_k=\alpha_m|\mathbf{r};\theta]\mathrm{Im}\left(\alpha_m^* r_k e^{-j\theta}\right)\end{aligned} \tag{9}$$

where $E_{\mathbf{a}|\mathbf{r}}[\cdot]$ and $E_{a_k|\mathbf{r}}[\cdot]$ refer to averaging over $\Pr[\mathbf{a}=\mathbf{c}_i|\mathbf{r};\theta]$ and $\Pr[a_k=\alpha_m|\mathbf{r};\theta]$, respectively, and $(\alpha_0,\alpha_1,\ldots,\alpha_{M-1})$ denotes the set of constellation points. Note that no approximation is involved in obtaining (9). Substitution of (9) into (2) yields an exact expression of the CRB in terms of the *marginal* APPs $\Pr[a_k=\alpha_m|\mathbf{r};\theta]$ of the coded data symbols.

B. Evaluation of the Marginal APPs of the Data Symbols

In principle, any marginal APP $\Pr[a_k=\alpha_m|\mathbf{r};\theta]$ can be obtained as a summation of joint APPs $\Pr[\mathbf{a}=\mathbf{c}_i|\mathbf{r};\theta]$, which in turn can be computed from (8). However, the computational complexity of this procedure increases exponentially with the sequence length L.

For codes that are described by means of a trellis, the marginal APPs can be determined directly by means of the BCJR algorithm [8]. As its computational complexity grows only linearly with the number of states and with the sequence length L, the BCJR algorithm is the appropriate tool for marginal APP computation in case of linear block codes, convolutional codes and trellis codes, provided that the number of states is manageable.

When the coded symbol sequence results from the (serial or parallel) concatenation of two encoders that are separated by an interleaver (such as turbo codes [9]), the underlying overall trellis has a number of states that grows exponentially with the interleaver size. However, when the encoders themselves are described by a small trellis, the marginal APPs are computed by means of iterated application of the BCJR algorithm to the individual trellises, with exchange of extrinsic information between the BCJR algorithms at each iteration (the same computation is carried out when performing iterated turbo decoding instead of the too complex MAP symbol decoding). When the coded bits (conditioned on $\mathbf{r}$ and θ) can be considered as independent (which is a reasonable assumption when the interleaver size is large), this iterative procedure yields the correct marginal APPs when reaching the steady state [10]. This approach is easily extended to other systems that use iterative decoding, such as bit interleaved coded modulation [11].

IV. Results and Discussion

Assuming the transmission of $L = 1001$ QPSK symbols with Gray mapping, we have numerically evaluated the ratio CRB/MCRB for phase estimation, by substituting (9) into (2), and approximating the statistical expectation in (2) by an arithmetical average. We have considered the following scenarios: 1) *uncoded* transmission; 2) a nonrecursive *convolutional code* (NRCC) with rate $r = 1/2$, $n = 16$ states and generator matrix $G = ((23)_8(35)_8)$; and 3) a *turbo code* (TC) consisting of the parallel concatenation of two identical recursive systematic convolutional codes (RSCC) with $r = 1/2$, $n = 16$ and $G = ((37)_8(21)_8)$, through a pseudorandom even/odd interleaver of length L; the output of the turbo encoder is punctured to obtain an overall rate of 1/2. Fig. 1 displays the ratio CRB/MCRB as a function of E_s/N_0 for the considered scenarios.

Fig. 1 shows that for very large E_s/N_0 the CRB converges to the MCRB. When E_s/N_0 decreases, a critical value $(E_s/N_0)_{\mathrm{crit}}$ is reached where the CRB starts to diverge from the MCRB; for turbo coding, convolutional coding and uncoded transmission, $(E_s/N_0)_{\mathrm{crit}}$ is about 1.5 dB, 3.5 dB and 10.5 dB, respectively. From simulations (not reported here), we have verified that this critical value corresponds to a BER of about 10^{-3} for the considered scenarios (and for other scenarios as well). Hence, $(E_s/N_0)_{\mathrm{crit}}$ is determined by the coding gain near $\mathrm{BER} = 10^{-3}$, and decreases with increasing coding

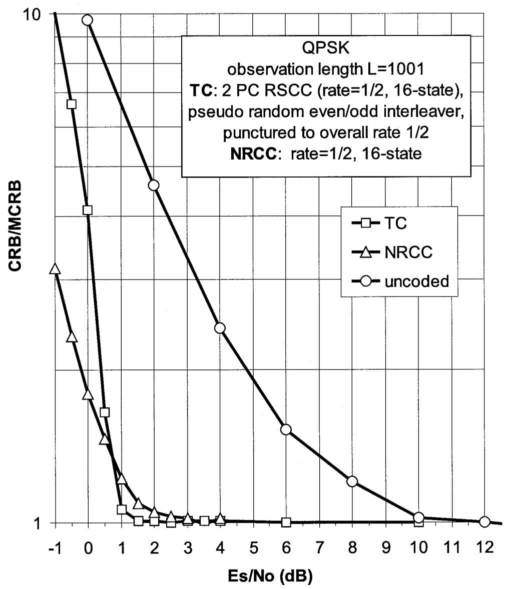

Fig. 1. The ratio CRB/MCRB for QPSK symbols and an observation length $L = 1001$. (NRCC: nonrecursive convolutional code, TC: turbo code, PC RSCC: parallel concatenated recursive systematic convolutional codes).

gain. This indicates that, as far as the CRB for carrier phase estimation is concerned, transmission at a BER less than 10^{-3} is nearly equivalent to transmitting a training sequence.

V. Conclusions

In this contribution we have expressed the CRB for carrier phase estimation in terms of the marginal APPs of the data symbols. In the case of block codes, convolutional codes or trellis codes that are described by means of a trellis of reasonable size, these marginal APPs are obtained efficiently from the BCJR algorithm. When the encoding involves two subsystems that are separated by an interleaver (turbo codes, bit interleaved coded modulation), the APPs result from an iterated application of the BCJR algorithm. We have presented numerical results indicating that the CRB essentially coincides with the MCRB, provided that the SNR exceeds some critical value. This critical SNR corresponds to a BER of about 10^{-3}, and hence decreases with the coding gain of the considered system. When operating below this critical SNR, the CRB diverges from the MCRB.

In this contribution we have restricted our attention to carrier phase estimation from coded signals. The investigation of the CRB for the estimation of multiple parameters (such as carrier phase, carrier frequency, symbol timing) from coded signals is a topic for further research.

References

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

[2] A. N. D'Andrea, U. Mengali, and R. Reggiannini, "The modified Cramer-Rao bound and its applications to synchronization problems," *IEEE Trans. Commun.*, vol. 42, pp. 1391–1399, Feb.–Apr. 1994.

[3] N. Noels, H. Steendam, and M. Moeneclaey, "The true Cramer-Rao bound for phase-independent carrier frequency estimation from a PSK signal," presented at the IEEE Globecom, Taipei, Taiwan, R.O.C., Nov. 17–21, 2002, Paper CTS-04-2, pp. 1137–1141.

[4] ——, "The impact of the observation model on the Cramer-Rao bound for carrier and frequency synchronization," presented at the IEEE Int. Conf. Communications 2003, Anchorage, AK, May 11–15, 2003.

[5] W. G. Cowley, "Phase and frequency estimation for PSK packets: Bounds and algorithms," *IEEE Trans. Commun.*, vol. 44, pp. 26–28, Jan. 1996.

[6] F. Rice, B. Cowley, B. Moran, and M. Rice, "Cramer–Rao lower bounds for QAM phase and frequency estimation," *IEEE Trans. Commun.*, vol. 49, pp. 1582–1591, Sept. 2001.

[7] 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. Commun.*, vol. 46, pp. 1536–1544, Nov. 1998.

[8] L. R. Bahl, J. Cocke, F. Jelinek, and J. Raviv, "Optimal decoding of linear codes for minimizing symbol error rate," *IEEE Trans. Inform. Theory*, pp. 284–287, Mar. 1974.

[9] C. Berrou and A. Glavieux, "Near optimum error correcting coding and decoding: Turbo codes," *IEEE Trans. Commun.*, vol. 44, pp. 1261–1271, Oct. 1996.

[10] T. Richardson, "The geometry of turbo-decoding dynamics," *IEEE Trans. Inform. Theory*, vol. 46, pp. 9–23, Jan 2000.

[11] S. T. Brink, J. Speidel, and R. H. Yan, "Iterative demapping and decoding for multi-level modulation," in *Proc. IEEE GLOBECOM 1998*, Sydney, Australia, pp. 579–584.