The True Cramer–Rao Bound for Carrier Frequency Estimation From a PSK Signal

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

***Abstract*—This paper considers the Cramer–Rao bound (CRB) related to estimating the carrier frequency of a noisy phase-shift keying signal. The following scenarios are discussed: 1) carrier frequency estimation irrespective of the carrier phase, based on either known or random data and 2) joint carrier phase and frequency estimation, based on either known or random data. Ideal symbol timing is assumed. We compare the results obtained from a (commonly used) simplified observation model against those resulting from the correct model. Because of the presence of nuisance parameters (random data and/or random carrier phase), the analytical computation of the corresponding CRBs is often not feasible. Here we present results that are based upon a combined analytical/numerical approach. Our results show that the choice of the observation model has essentially no effect on the CRBs at moderate and high signal-to-noise ratios. We also show that of the two scenarios considered, joint frequency and phase estimation yields the smaller CRB; the penalty resulting from frequency estimation, irrespective of the carrier phase, decreases with increasing observation interval.**

***Index Terms*—Cramer–Rao bound (CRB), frequency estimation, carrier synchronization.**

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 observation depend not only on the vector parameter to be estimated, but also on a nuisance vector parameter we do not want to estimate. The presence of this nuisance parameter makes the analytical computation of the CRB very hard, if not impossible.

In order to avoid the computational complexity caused by the nuisance parameters, a modified CRB (MCRB) has been derived in [2] and [3]. The MCRB is much simpler to evaluate than the CRB, but is, in general, looser than the CRB. In [4], the high signal-to-noise ratio (SNR) limit of the CRB has been evaluated analytically, and has been shown to coincide with the MCRB when estimating the delay, the frequency offset, or the carrier phase of a linearly modulated waveform.

The true CRB related to joint carrier phase and frequency estimation (but not frequency estimation irrespective of the carrier phase) has been derived for binary phase-shift keying (BPSK) and quaternary PSK (QPSK) in [5], and for quadrature amplitude modulation (QAM) in [6], assuming a simplified observation model. In this model, frequency error correction is applied after the matched filter, and the observation consists of the matched-filter output samples (taken at the decision instants) before frequency correction. The simplification consists of neglecting the signal reduction and intersymbol interference (ISI) that occur at the matched-filter output when the frequency offset is nonzero. In [7], the low SNR limit of the CRB for carrier phase and frequency estimation, again assuming this simplified observation model, has been obtained analytically for M-ary (M)-PSK, M-ary (M)-QAM, and M-ary pulse amplitude modulation (M-PAM).

In this paper, we investigate the true CRBs related to frequency estimation, irrespective of the carrier phase, and to joint phase and frequency estimation. The data symbols are either known (e.g., training sequence) or random, and are taken from an M-PSK constellation. The corresponding low SNR limits of these CRBs are presented as well. We consider both the correct model (where the observation consists of the noisy PSK signal at the receiver input) and the simplified observation model adopted in [5]–[7]. The transmit pulse is a square-root Nyquist pulse, and we assume the time delay to be known. Results are presented for various PSK constellations and several transmitted sequence lengths. The main conclusions are: 1) both observation models yield essentially the same results for moderate and high SNR; and 2) frequency estimation irrespective of the carrier phase exhibits a performance penalty, as compared with joint frequency and phase estimation, but this penalty decreases with increasing observation intervals.

II. Problem Formulation

Let us consider the complex baseband representation $r(t)$ of a noisy PSK signal

$$r(t) = \varepsilon \sum_{k=-K}^{K} a_k h(t - kT) \exp(j(2\pi F t + \theta)) + w(t) \quad (1)$$

where $\mathbf{a} = (a_{-K}, \ldots a_K)$ is a vector of $L = 2K + 1$ M-PSK symbols ($|a_k|^2 = 1$); $h(t)$ is a real-valued unit-energy square-root Nyquist pulse; F is the carrier frequency offset; θ is the carrier phase at $t = 0$; T is the symbol interval; $w(t)$ is complex-valued zero-mean Gaussian noise with independent real and imaginary parts, each having a normalized power spectral density of $1/2$; and $\varepsilon = (E_s/N_0)^{1/2}$, with E_s and N_0 denoting the symbol energy and the noise power spectral density, respectively. Depending on the scenario to be considered,

Paper approved by R. Regiannini, the Editor for Synchronization and Wireless Applications of the IEEE Communications Society. Manuscript received April 24, 2002; revised April 9, 2003 and November 18, 2003. This work was supported by the Interuniversity Attraction Poles Programme P5/11-Belgian Science Policy. This paper was presented in part at the IEEE Global Telecommunications Conference, Taipei, Taiwan, November 17–21, 2002.

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

Digital Object Identifier 10.1109/TCOMM.2004.826244

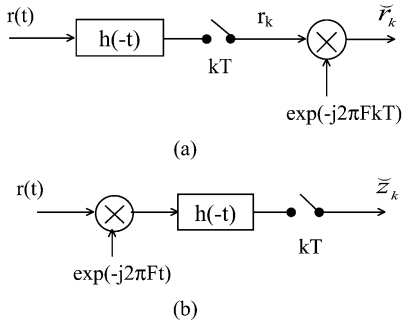

Fig. 1. Computation of (a) $\breve{r}_k$ and (b) $\breve{z}_k$.

the M-PSK symbols are either *a priori* known to the receiver (training sequence), or they are statistically independent and uniformly distributed over the M-PSK constellation (random data).

In [5]–[7], CRBs related to frequency estimation have been derived, assuming the following observation model $(k = -K, \ldots, K)$:

$$r_k = \varepsilon a_k \exp(j(2\pi F kT + \theta)) + w_k \tag{2}$$

where $\{w_k\}$ is a sequence of independent zero-mean complex-valued Gaussian random variables, with independent real and imaginary parts that each have variance equal to $1/2$. In (2), r_k stands for the matched-filter output sample taken at the correct decision instant kT, when $r(t)$ from (1) is applied to the matched filter, and the frequency offset is assumed to be small (i.e., $|FT| \ll 1$) [see Fig. 1(a)]. It is important to realize that the observations $r(t)$ from (1) and $\{r_k\}$ from (2) are not equivalent, as will be pointed out in the following.

Suppose that one is able to produce from an observation vector $\mathbf{r}$ an *unbiased* estimate $\hat{\mathbf{u}}$ of a deterministic vector parameter $\mathbf{u}$. Then the estimation error variance is lower bounded by the CRB [1]: $E_{\mathbf{r}}[(\hat{u}_i - u_i)^2] \geq \mathrm{CRB}_i(\mathbf{u})$, where $\mathrm{CRB}_i(\mathbf{u})$ is the ith diagonal element of the inverse of the *Fisher information matrix* $\mathbf{J}(\mathbf{u})$. The (i,j)th element of $\mathbf{J}(\mathbf{u})$ is given by

$$\begin{aligned}\mathbf{J}_{ij}(\mathbf{u}) &= E_{\mathbf{r}}\left[-\frac{\partial^2}{\partial u_i \partial u_j}\ln(p(\mathbf{r};\mathbf{u}))\right] \\ &= E_{\mathbf{r}}\left[\frac{\partial}{\partial u_i}\ln(p(\mathbf{r};\mathbf{u}))\frac{\partial}{\partial u_j}\ln(p(\mathbf{r};\mathbf{u}))\right].\end{aligned} \tag{3}$$

Note that $\mathbf{J}(\mathbf{u})$ is a symmetrical matrix. The probability density $p(\mathbf{r};\mathbf{u})$ of $\mathbf{r}$, corresponding to a given value of $\mathbf{u}$, is called the *likelihood function* of $\mathbf{u}$, while $\ln(p(\mathbf{r};\mathbf{u}))$ is the *log-likelihood function (LLF)* of $\mathbf{u}$. The expectation $E_r[\cdot]$ in (3) is with respect to $p(\mathbf{r};\mathbf{u})$.

We investigate two scenarios. For each scenario, the data symbols are either known (e.g., training sequence) or random to the receiver. Also, the effect of the observation model (correct/simplified) will be considered.

- **Scenario 1)**: Estimation of F, irrespective of θ.

  The useful parameter is given by $\mathbf{u} = F$. The nuisance parameter is given by $\mathbf{v} = \theta$ when the transmitted data symbols are known, or by $\mathbf{v} = (\theta, \mathbf{a})$ when the data symbols are random. In this scenario, θ is considered as uniformly distributed in $(-\pi, \pi)$.

- **Scenario 2)**: Joint estimation of (F, θ).

  The useful parameter is given by $\mathbf{u} = (F, \theta)$. There is no nuisance parameter when the transmitted data symbols are known. In the case of random data, the nuisance parameter is given by $\mathbf{v} = \mathbf{a}$.

For both scenarios, the joint likelihood function $p(\mathbf{r} \,|\, \mathbf{v};\mathbf{u})$ is, within a factor not depending on $(\mathbf{u}, \mathbf{v})$, given by

$$p(\mathbf{r} \,|\, \mathbf{v};\mathbf{u}) = \prod_{k=-K}^{K} F(a_k, x_k(F)e^{-j\theta}) \tag{4}$$

where

$$F(a_k, x_k(F)e^{-j\theta}) = \exp(\varepsilon a_k x_k^*(F)e^{j\theta} + \varepsilon a_k^* x_k(F)e^{-j\theta}). \tag{5}$$

When using the simplified observation model (2), the vector $\mathbf{r}$ is given by $(r_{-K}, \ldots, r_K)$, and $x_k(F)$ equals $\breve{r}_k(F)$ given by

$$\breve{r}_k(F) = r_k \exp(-j2\pi F kT). \tag{6}$$

As indicated in Fig. 1(a), the quantity r_k is obtained by feeding $r(t)$ to a filter matched to the transmit pulse $h(t)$, and sampling the matched-filter output at instant kT. The quantity $\breve{r}_k(F)$ is obtained by applying to r_k a rotation of $-2\pi FkT$ rad. When using the correct observation model (1), $\mathbf{r}$ is a vector representation of the signal $r(t)$ from (1), and $x_k(F)$ equals $\breve{z}_k(F)$ given by

$$\breve{z}_k(F) = \int r(t)\exp(-j2\pi Ft)h(t - kT)\,dt. \tag{7}$$

As indicated in Fig. 1(b), the quantity $\breve{z}_k(F)$ is obtained by first applying to $r(t)$ a constant-speed rotation of $-2\pi F$ rad/s, feeding the result to a filter matched to the transmit pulse $h(t)$, and sampling the matched-filter output at instant kT. Note the similarity between (6) and (7). However, unless $F = 0$, it follows from Fig. 1 that $\breve{r}_k(F) \neq \breve{z}_k(F)$. Actually, $\{\breve{z}_k(F)\}$ cannot be computed from $\{\breve{r}_k(F)\}$, and therefore, estimating F from $\{r_k\}$ instead of $r(t)$ is suboptimum.

The LLF $\ln(p(\mathbf{r};\mathbf{u}))$ resulting from (4) is given by

$$\ln p(\mathbf{r};\mathbf{u}) = \ln\left(E_{\mathbf{v}}\left[\prod_{k=-K}^{K} F(a_k, x_k(F)e^{-j\theta})\right]\right). \tag{8}$$

Computation of the CRB requires the substitution of (8) into (3), and the evaluation of the various expectations included in (8) and (3).

As the evaluation of the expectations involved in $\mathbf{J}(\mathbf{u})$ and $p(\mathbf{r};\mathbf{u})$ is quite tedious, a simpler lower bound, called the MCRB, has been derived in [2] and [3], i.e., $E_{\mathbf{r}}[(\hat{u}_i - u_i)^2] \geq \mathrm{CRB}_i(\mathbf{u}) \geq \mathrm{MCRB}_i(\mathbf{u})$. The MCRB for frequency estimation, is given by [2], [3]

$$\mathrm{MCRB}_F \cong \frac{3}{2\pi^2 \frac{E_s}{N_0} L(L^2 - 1)} \cdot \frac{1}{T^2} \tag{9}$$

where $L = 2K + 1$ denotes the number of symbols transmitted within the observation interval. When using the simplified observation model (2), the MCRB (9) is valid for both scenarios 1) and 2), irrespective of the data symbols being known or random. When using the correct observation model (1), the resulting MCRB converges to (9) for large L, for both scenarios and both cases of known or random data symbols. In [4], it has

been shown that for *high* SNR (i.e., $E_s/N_0 \to \infty$), the CRB for frequency estimation converges to the MCRB corresponding to the considered scenario.

Also, a closed-form expression can be derived for the *low*-SNR limit (i.e., $E_s/N_0 \to 0$) of the CRB, which we call the asymptotic CRB (ACRB). In [7], this has been accomplished for scenario 2) with random data and using the simplified observation model.

In this paper, we compute the CRBs resulting from the scenarios mentioned above, and present the expressions for the corresponding ACRBs. It should be noted that these ACRBs do not necessarily provide a lower bound on the actual frequency-error variance for moderate and large SNR.

III. Evaluation of the True CRB

A. Estimation of F Irrespective of θ; Random Data Symbols

Taking in (8) $\mathbf{v} = (\mathbf{a}, \theta)$, the LLF $\ln(p(\mathbf{r} \mid F))$ is (within an arbitrary constant) given by

$$\ln(p(\mathbf{r} \mid F)) = \ln\left(\int_{-\pi}^{\pi} K(\theta, \mathbf{x}(F))\, d\theta\right) \tag{10}$$

where $\mathbf{x}(F) = (x_{-K}(F), \ldots, x_K(F))$

$$K(\theta, \mathbf{x}(F)) = \prod_{k=-K}^{K} I(x_k(F) e^{-j\theta}) \tag{11}$$

$$I(x_k(F) e^{-j\theta}) = \sum_{i=0}^{M-1} F(\alpha_i, x_k(F) e^{-j\theta}) \tag{12}$$

and $\{\alpha_0, \alpha_1, \ldots, \alpha_{M-1}\}$ is the set of PSK constellation points. Differentiation of (10) yields

$$\frac{d}{dF} \ln p(\mathbf{r} \mid F) = \varepsilon \int_{-\pi}^{\pi} \sum_{k=-K}^{K} \sum_{i=0}^{M-1} M(\theta, \mathbf{x}) H(\alpha_i, x_k e^{-j\theta}, x_{Fk} e^{-j\theta})\, d\theta \tag{13}$$

where $\mathbf{x}, x_k$ and x_{Fk} are shorthand notations for $\mathbf{x}(F), x_k(F)$ and $x_{kF}(F)$, respectively, and

$$M(\theta, \mathbf{x}) = \frac{K(\theta, \mathbf{x})}{\int_{-\pi}^{\pi} K(\theta, \mathbf{x})\, d\theta} \tag{14}$$

$$H(\alpha_i, x_k e^{-j\theta}, x_{Fk} e^{-j\theta}) = G(\alpha_i, x_k e^{-j\theta}) 2\Re(\alpha_i^* x_{Fk} e^{-j\theta}) \tag{15}$$

$$G(\alpha_i, x_k e^{-j\theta}) = \frac{F(\alpha_i, x_k e^{-j\theta})}{I(x_k e^{-j\theta})} \tag{16}$$

and the subscript F denotes differentiation with respect to F. In these expressions, $x_k(F) = \breve{r}_k(F)$ or $x_k(F) = \breve{z}_k(F)$, and $x_{Fk}(F) = -j2\pi kT\, \breve{r}_k(F)$ or $x_{Fk}(F) = \breve{z}_{Fk}(F)$, according to the observation model that is being used. As the variable F in (6) and (7) corresponds to the actual frequency offset, the quantities $\breve{r}_k(F)$, $\breve{z}_k(F)$, and $\breve{z}_{Fk}(F)$ can be decomposed as

$$\breve{r}_k(F) = \breve{z}_k(F) = \varepsilon a_k + N(k) \tag{17}$$

$$\breve{z}_{Fk}(F) = -j2\pi kT\varepsilon a_k + N_F(k) \tag{18}$$

where $N(k)$ and $N_F(k)$ are zero-mean complex Gaussian random variables, with

$$\begin{aligned} E[N(m)N^*(n)] &= \delta_{m-n} \\ E[N_F(m)N^*(n)] &= -j2\pi mT\delta_{m-n} \\ E[N_F(m)N_F^*(n)] &= 4\pi^2 m^2 T^2 \delta_{m-n} \\ &\quad + 4\pi^2 \int_{-\infty}^{+\infty} t^2 h(t) h(t-(m-n)T)\, dt. \end{aligned} \tag{19}$$

As $M(\theta, \mathbf{x}), H(\alpha_i, x_k e^{-j\theta}, x_{Fk} e^{-j\theta})$ and $K(\theta, \mathbf{x})$ are periodic in θ, with period equal to 2π, there is no need to include a carrier phase in the first term of (17) and (18). Note that the statistics of $\breve{r}_k(F)$, $\breve{z}_k(F)$, and $\breve{z}_{Fk}(F)$ in (17) and (18) do not depend on the value of the actual frequency offset F.

Taking (13) into account, we derive from (3)

$$\mathrm{CRB}^{-1} = \frac{E_s}{N_0} \int_{\theta_1,\theta_2} \sum_{k_1,k_2=-K}^{K} \sum_{i_1,i_2=0}^{M-1} Q_{i_1,i_2,k_1,k_2}(\theta_1,\theta_2) d\theta_1 d\theta_2 \tag{20}$$

where

$$\begin{aligned} &Q_{i_1,i_2,k_1,k_2}(\theta_1,\theta_2) \\ &\quad = E\left[\begin{matrix} M(\theta_1,\mathbf{x}) H\left(\alpha_{i_1}, x_{k_1} e^{-j\theta_1}, x_{Fk_1} e^{-j\theta_1}\right) \\ \cdot M(\theta_2,\mathbf{x}) H\left(\alpha_{i_2}, x_{k_2} e^{-j\theta_2}, x_{Fk_2} e^{-j\theta_2}\right) \end{matrix}\right] \end{aligned} \tag{21}$$

and $E[\cdot]$ denotes averaging over the data symbols and the noise. Using the method outlined in [8], (21) can be further simplified by making use of the statistical properties of x_{Fk_1} and x_{Fk_2}, conditioned on $\mathbf{x}$

$$\begin{aligned} &E[x_{Fk} \mid \mathbf{x}] = -j2\pi kT x_k \\ &E\left[x_{Fk_1}^* x_{Fk_2} \mid \mathbf{x}\right] \\ &\quad = \begin{cases} 4\pi^2 k_1 k_2 T^2 x_{k_1}^* x_{k_2} & \text{(si)} \\ 4\pi^2 k_1 k_2 T^2 x_{k_1}^* x_{k_2} & \\ \quad + 4\pi^2 \int_{-\infty}^{\infty} t^2 h(t) h(t-(k_1-k_2)T)\, dt & \text{(co)} \end{cases} \\ &E\left[x_{Fk_1} x_{Fk_2} \mid \mathbf{x}\right] = -4\pi^2 k_1 k_2 T^2 x_{k_1} x_{k_2}. \end{aligned} \tag{22}$$

In (22), the captions si and co refer to the observation model (si: simplified, co: correct). The average in (21) is computed by first taking the expectation conditioned on $\mathbf{x}$, and then averaging over $\mathbf{x}$. We obtain

$$\frac{1}{\mathrm{CRB}^{(i),\mathrm{ra,si}}} = \frac{E_s}{N_0} \frac{4\pi^2 L(L^2-1)T^2}{3} \cdot A\left(\frac{E_s}{N_0}, M, L\right) \tag{23a}$$

$$\frac{1}{\mathrm{CRB}^{(i),\mathrm{ra,co}}} = \frac{1}{\mathrm{CRB}^{(i),\mathrm{ra,si}}} + 8\pi^2 \frac{E_s}{N_0} L \times \int_{-\infty}^{+\infty} t^2 h^2(t)\, dt \cdot B\left(\frac{E_s}{N_0}, M, L\right) \tag{23b}$$

where

$$A\left(\frac{E_s}{N_0}, M, L\right) = E\left[\begin{matrix} \Im^2(N(k,\mathbf{x})x_k) \\ -\Im(N(k,\mathbf{x})x_k)\Im(N(m,\mathbf{x})x_m) \end{matrix}\right] \tag{24}$$

$$B\left(\frac{E_s}{N_0}, M, L\right) = E[|N(k,\mathbf{x})|^2] \tag{25}$$

$$N(k,\mathbf{x}) = \int_{-\pi}^{\pi} \sum_{i=0}^{M-1} M(\theta,\mathbf{x}) G(\alpha_i, x_k e^{-j\theta}) \alpha_i^* e^{-j\theta}\, d\theta. \tag{26}$$

In (24), x_k and x_m represent two statistically independent quantities ($k \neq m$). The three superscripts in (23) and in subsequent equations related to the CRBs refer to scenario 1) or 2), to the receiver's knowledge about the data symbols (ra: random, kn: known), and to the observation model (si: simplified, co: correct). We observe from (23) that $1/\mathrm{CRB}^{(i),\mathrm{ra,co}}$ equals the sum of $1/\mathrm{CRB}^{(i),\mathrm{ra,si}}$ and an additional positive term. Hence, $\mathrm{CRB}^{(i),\mathrm{ra,co}} < \mathrm{CRB}^{(i),\mathrm{ra,si}}$, which is in keeping with the fact that frequency estimation from $\{r_k\}$ in (2) is suboptimum. As outlined in the Appendix, $1/\mathrm{CRB}^{(i),\mathrm{ra,si}}$ dominates at high E_s/N_0 (for large L), whereas at low SNR values, the additional term becomes the largest. Consequently, the simplified observation model will provide a CRB that approaches the CRB from the correct observation model at sufficiently high SNR. Note that $A(E_s/N_0, M, L)$ and $B(E_s/N_0, M, L)$ depend on E_s/N_0, on the size of the constellation, and on the transmitted sequence length, but not on the shape of the square-root Nyquist transmit pulse. The shape of the transmit pulse affects only the coefficient of $B(E_s/N_0, M, L)$ in (23b).

For low SNR, (23a) and (23b) converge to the corresponding ACRBs. The computation of these ACRBs is outlined in the Appendix; we obtain, for the simplified observation model

$$\mathrm{ACRB}_F^{(i),\mathrm{ra,si}} = \frac{3((M-1)!)^2}{2\pi^2 \left(\frac{E_s}{N_0}\right)^{2M} L^2(L^2-1)} \frac{1}{T^2} \tag{27}$$

where $L = 2K+1$ and M represents the number of constellation points. The ACRB (27) is proportional to $(E_s/N_0)^{-2M}$ and to L^{-4}. With the correct model, we get the following expression for the ACRB:

$$\mathrm{ACRB}_F^{(i),\mathrm{ra,co}} = \frac{1}{8\pi^2 \left(\frac{E_s}{N_0}\right)^2 L \int_{-\infty}^{+\infty} t^2 h^2(t)\, dt} \tag{28}$$

which is only proportional to $(E_s/N_0)^{-2}$ and to L^{-1}. Note that (28) is independent of the constellation size M, but is affected by the transmit pulse $h(t)$.

B. Estimation of F Irrespective of θ; Known Data Symbols

When the transmitted data symbols are known at the receiver, the nuisance parameter is given by $\mathbf{v} = \theta$, and no averaging over the data is required. Equations (20)–(22) remain valid, provided we remove from (12) the summations over the constellation points, and replace α_i by the actual symbol a_k. The average in (21) is computed by first taking the expectation conditioned on $\mathbf{N} = (N_{-K}, \ldots, N_K)$, and then averaging over $\mathbf{N}$. We obtain

$$\frac{1}{\mathrm{CRB}^{(i),\mathrm{kn,si}}} = \frac{2\pi^2 L(L^2-1)T^2}{3} E\left[\left|\int_{-\pi}^{\pi} M(\theta,\mathbf{x}) e^{-j\theta}\, d\theta\right|^2\right] \tag{29}$$

$$\frac{1}{\mathrm{CRB}^{(i),\mathrm{kn,co}}} = \frac{1}{\mathrm{CRB}^{(i),\mathrm{kn,si}}} \left(1 + \sum_{k_1,k_2=-K}^{K} a_{k_1}^* a_{k_2} U(k_1-k_2)\right) \tag{30a}$$

with

$$U(k) = \frac{12}{L(L^2-1)T^2} \int_{-\infty}^{+\infty} t^2 h(t) h(t-kT)\, dt. \tag{30b}$$

Note that for symbols taken from a PSK constellation, the statistics of $M(\theta,\mathbf{x})$ do not depend on the specific data sequence transmitted; hence, $\mathrm{CRB}^{(i),\mathrm{kn,si}}$ from (29) does not depend on the particular training sequence. The second term between parentheses in (30a) indicates that the CRB corresponding to the correct observation model does depend on the transmitted data sequence. However, for large L, this dependence is very weak. Indeed, for long training sequences with symbols drawn independently from a PSK constellation, the statistical fluctuation of this second term in (30a) can be ignored, as compared with its mean. Therefore, we can approximate this term by its mean over all possible training sequences, which corresponds to keeping in (30a) only the terms with $k_1 = k_2$. From the resulting expression, it follows that the ratio of the first to the second term of (30) is proportional to L^2, so that the second term can be safely ignored for large L. This indicates that $\mathrm{CRB}^{(i),\mathrm{kn,si}}$, and $\mathrm{CRB}^{(i),\mathrm{kn,co}}$ are essentially the same for long training sequences.

The corresponding low SNR limit can be derived as in [7]. For the simplified model, we obtain (see the Appendix)

$$\mathrm{ACRB}_F^{(i),\mathrm{kn,si}} = \frac{3}{2\pi^2 \left(\frac{E_s}{N_0}\right)^2 L^2(L^2-1)} \frac{1}{T^2} \tag{31}$$

which is proportional to $(E_s/N_0)^{-2}$ and to L^{-4}. For the correct observation model, the resulting $\mathrm{ACRB}_F^{(i),\mathrm{kn,co}}$ converges for large L to $\mathrm{ACRB}_F^{(i),\mathrm{kn,si}}$ from (31).

C. Joint Estimation of F and θ; Random Data Symbols

Now we take in (8) $\mathbf{v} = \mathbf{a}$. We obtain

$$\ln(p(\mathbf{r} \mid F,\theta)) = \sum_{k=-K}^{K} \ln(I(x_k(F) e^{-j\theta})) \tag{32}$$

where $\{\alpha_0, \alpha_1, \ldots, \alpha_{M-1}\}$ is the set of PSK constellation points. Differentiation of (32) yields

$$\frac{\partial}{\partial F} \ln p(\mathbf{r} \mid F,\theta) = \varepsilon \sum_{k=-K}^{K} \sum_{i=0}^{M-1} H(\alpha_i, x_k e^{-j\theta}, x_{Fk} e^{-j\theta}) \tag{33}$$

$$\frac{\partial}{\partial \theta} \ln p(\mathbf{r} \mid F,\theta) = \varepsilon \sum_{k=-K}^{K} \sum_{i=0}^{M-1} H(\alpha_i, x_k e^{-j\theta}, -j x_k e^{-j\theta}). \tag{34}$$

Taking (33) and (34) into account, J_{FF} and $J_{F\theta}$ from (3) can be represented as

$$J_{FF}=\frac{E_s}{N_0}\sum_{k_1,k_2=-K}^{K}\sum_{i_1,i_2=0}^{M-1}E\left[\begin{array}{c}H\left(\alpha_{i_1},x_{k_1}e^{-j\theta},x_{Fk_1}e^{-j\theta}\right)\\ \cdot H\left(\alpha_{i_2},x_{k_2}e^{-j\theta},x_{Fk_2}e^{-j\theta}\right)\end{array}\right] \tag{35}$$

$$J_{F\theta}=\frac{E_s}{N_0}\sum_{k_1,k_2=-K}^{K}\sum_{i_1,i_2=0}^{M-1}E\left[\begin{array}{c}H\left(\alpha_{i_1},x_{k_1}e^{-j\theta},x_{Fk_1}e^{-j\theta}\right)\\ \cdot H\left(\alpha_{i_2},x_{k_2}e^{-j\theta},-jx_{k_2}e^{-j\theta}\right)\end{array}\right] \tag{36}$$

where $E[\cdot]$ denotes averaging over the data symbols and the noise. As in Section III-A, $x_k(F) = \breve{r}_k(F)$ or $x_k(F) = \breve{z}_k(F)$; and $x_{Fk}(F) = -j2\pi kT\,\breve{r}_k(F)$ or $x_{Fk}(F) = \breve{z}_{Fk}(F)$, depending on the observation model. The decomposition of $\breve{r}_k(F)e^{-j\theta}$, $\breve{z}_k(F)e^{-j\theta}$ and $\breve{z}_{Fk}\ (F)e^{-j\theta}$ yields the same result as the decomposition of $\breve{r}_k(F)$, $\breve{z}_k(F)$ and $\breve{z}_{Fk}\ (F)$ in Section III-A [see (17)–(19)]. Note that $(x_{k_1}e^{-j\theta},x_{k_2}e^{-j\theta})$ depends only on $(a_{k_1},a_{k_2},N_{k_1},N_{k_2})$. Again, (35) and (36) can be further simplified by making use of the statistical properties of $x_{Fk_1}e^{-j\theta}$ and $x_{Fk_2}e^{-j\theta}$, conditioned on $(a_{k_1},a_{k_2},N_{k_1},N_{k_2})$, that turn out to be the same as the statistical properties of $x_{Fk_1}e^{-j\theta}$ and $x_{Fk_2}e^{-j\theta}$, conditioned on $\mathbf{x}$ in Section III-A. The average in (35) and (36) is computed by first taking the expectation conditioned on $(x_{k_1}e^{-j\theta},x_{k_2}e^{-j\theta})$, and then averaging over $(x_{k_1}e^{-j\theta},x_{k_2}e^{-j\theta})$. We obtain $J_{F\theta}=0$ and, consequently

$$\frac{1}{\mathrm{CRB}^{(ii),\mathrm{ra,si}}}=\frac{E_s}{N_0}\frac{4\pi^2L(L^2-1)T^2}{3}\cdot C\left(\frac{E_s}{N_0},M\right) \tag{37a}$$

$$\frac{1}{\mathrm{CRB}^{(ii),\mathrm{ra,co}}}=\frac{1}{\mathrm{CRB}^{(ii),\mathrm{ra,si}}}+\frac{E_s}{N_0}8\pi^2L\times\int_{-\infty}^{+\infty}t^2h^2(t)\,dt\cdot D\left(\frac{E_s}{N_0},M\right) \tag{37b}$$

where

$$C\left(\frac{E_s}{N_0},M\right)=E\left[\left(\Im\left(\sum_{i=1}^{M-1}G(\alpha_i,x_ke^{-j\theta})\alpha_i^*x_ke^{-j\theta}\right)\right)^2\right] \tag{38}$$

$$D\left(\frac{E_s}{N_0},M\right)=E\left[\left|\sum_{i=0}^{M-1}G(\alpha_i,x_ke^{-j\theta})\alpha_i^*\right|^2\right]. \tag{39}$$

We observe that $1/\mathrm{CRB}^{(ii),\mathrm{ra,co}}$ equals the sum of $1/\mathrm{CRB}^{(ii),\mathrm{ra,si}}$ and an additional positive term, which implies $\mathrm{CRB}^{(ii),\mathrm{ra,co}} < \mathrm{CRB}^{(ii),\mathrm{ra,si}}$. Again, the term $1/\mathrm{CRB}^{(ii),\mathrm{ra,si}}$ dominates at high E_s/N_0 (for large L), whereas at low SNR, the additional term becomes the largest (see the Appendix). The remarks from Section III-A, concerning the behavior of the CRBs at sufficiently high SNR and the effect of the pulse shape, are also valid here. Note that $\mathrm{CRB}^{(ii),\mathrm{ra,si}}$ from (37a) for BPSK and QPSK yields the CRB for frequency estimation presented in [5].

The low SNR limit based on the simplified observation model has been derived in [7]

$$\mathrm{ACRB}_F^{(ii),\mathrm{ra,si}}=\frac{3(M!)}{2\pi^2\left(\frac{E_s}{N_0}\right)^M M^2L(L^2-1)}\frac{1}{T^2}. \tag{40}$$

In contrast with (27), this ACRB is proportional to $(E_s/N_0)^{-M}$ and to L^{-3}. For the correct observation model, we must distinguish between $M>2$ and $M=2$. For $M>2$, $\mathrm{ACRB}_F^{(ii),\mathrm{ra,co}}$ resulting from (37b) is the same as $\mathrm{ACRB}_F^{(i),\mathrm{ra,co}}$ from (28); this indicates that estimating F independently of θ or jointly with θ yields the same ACRB when $M>2$ and the correct observation model is used. For $M=2$, we find that for large L, $\mathrm{ACRB}_F^{(ii),\mathrm{ra,co}}$ converges to $\mathrm{ACRB}_F^{(ii),\mathrm{ra,si}}$ from (40), evaluated for $M=2$; hence, for $M=2$, both observation models yield essentially the same ACRB.

D. Joint Estimation of F and θ; Known Data Symbols

In this specific case, it is a simple matter to compute $\mathrm{CRB}^{(ii),\mathrm{kn,si}}$, and $\mathrm{CRB}^{(ii),\mathrm{kn,co}}$ analytically, because there are no nuisance parameters. It is easily shown that the parameters F and θ are decoupled ($J_{\theta F}=0$). The resulting CRBs are nothing but the MCRB_F given by (9).

IV. Numerical Results and Discussion

As no further analytical simplification of (24), (25), (29), (30), (38), and (39) seems possible, we have to resort to numerical computation. This involves numerical integrations with respect to θ in (26), (29), and (30), and replacing the statistical expectations in (24), (25), (29), (30), (38), and (39) by arithmetical averages over a number of computer-generated vectors $\mathbf{x}$.

A. Known Data Symbols

In the case of known data symbols, the CRB_F resulting from the joint estimation of F and θ equals MCRB_F. Hence, for any other scenario, the ratio CRB/MCRB is a measure of the penalty occurred by not knowing the data symbols and/or not estimating θ jointly with F.

Fig. 2 shows the ratio CRB/MCRB related to the estimation of F independently of θ, i.e., scenario 1), along with the corresponding ACRB/MCRB. We have pointed out in Section III-B that both observation models yield the same ACRB; our numerical results indicate that also the resulting CRBs are essentially the same (when using the approximation of $\mathrm{CRB}^{(i),\mathrm{kn,co}}$, outlined in Section III-B). For small (large) SNR, the CRBs converge to the corresponding ACRB (to the MCRB). Increasing the observation interval from L_1 to L_2 shifts the curve of CRB/MCRB to the left by an amount of $10\log(L_2/L_1)$ dB; hence, the value of E_s/N_0 at which the CRB comes close to the MCRB is shifted by the same amount. Note that for SNR values of practical interest, CRB/MCRB is close to one, even for a moderate length of the observation interval.

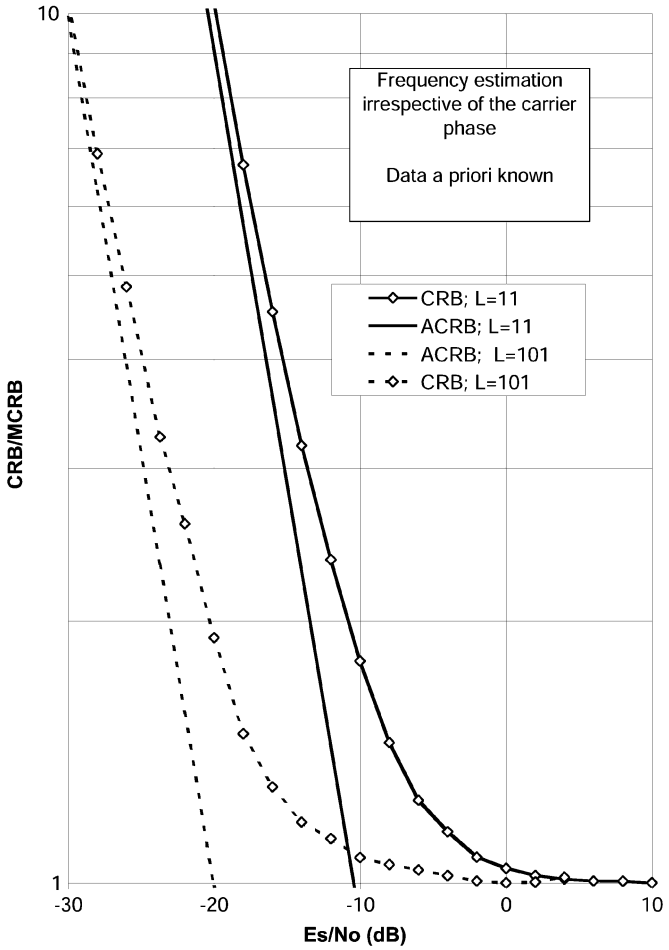

Fig. 2. CRB related to frequency estimation irrespective of the carrier phase, based on known data symbols.

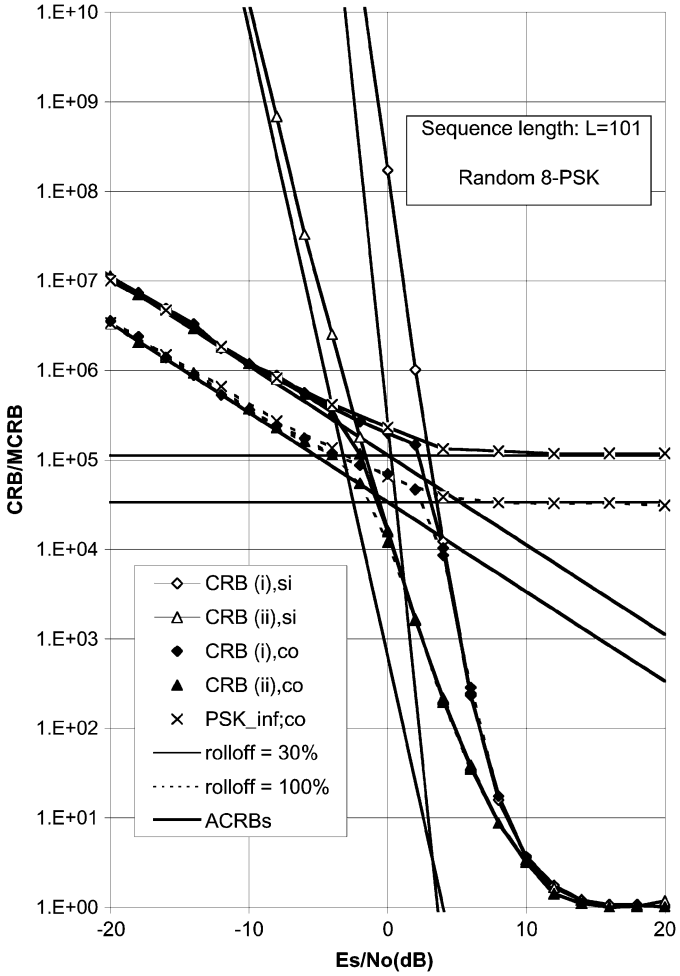

Fig. 3. CRB related to carrier frequency estimation irrespective of [scenario 1)] and jointly with [scenario 2)] the carrier phase for an 8-PSK constellation (si: simplified observation model, co: correct observation model).

B. Random Data Symbols, $M > 2$

It follows from Sections III-A and III-C that in the case of random data symbols with $M > 2$, the correct observation model yields essentially the same ACRBs, irrespective of whether F is estimated jointly with or independently of θ. However, the ACRBs resulting from these two scenarios are much different when using the simplified observation model.

Figs. 3 and 4 show, for 8-PSK and 4-PSK, respectively, the ratios CRB/MCRB and the corresponding ACRB/MCRB, for scenarios 1) and 2), and for both observation models. The behavior of the various curves is as follows.

- For small SNR (say, $E_s/N_0 \leq -10$ dB), the CRBs are close to the corresponding ACRBs. We have verified that also for 8-PSK, the $\mathrm{CRB}^{(i),\mathrm{ra,si}}$ is very close to the $\mathrm{ACRB}^{(i),\mathrm{ra,si}}$ for $E_s/N_0 \leq -10$ dB, but because of the steepness of the ACRB (40) for higher constellation sizes M, the ACRB at $E_s/N_0 = -10$ dB takes a very large value (about 1E20) that is outside the range shown in Fig. 3. Hence, the convergence of the CRB to the ACRB is quite constellation-size independent.
- For a given observation model, scenario 1) yields the larger CRB. This indicates that estimating F jointly with θ is potentially more accurate than estimating F irrespective of θ. Indeed, as F and θ are uncoupled, the joint estimation of F and θ yields the same CRB_F as estimating F when θ is *a priori* known, and this CRB_F is smaller than the one resulting from estimating F irrespective of θ.
- For a given scenario, the simplified observation model yields the larger CRB. Indeed, as the transformation from $r(t)$ to $\{r_k\}$ cannot be inverted, estimating F from $\{r_k\}$ instead of $r(t)$ is suboptimum.
- We have pointed out in Section III-C that both observation models yield the same CRB for large values of E_s/N_0. Our numerical results indicate that for scenario 1), as well as for scenario 2), the two observation models yield essentially the same CRB for SNR values of practical interest. The difference between these observation models becomes apparent only at small or even very small SNR. Consequently, the shape of the transmit pulse has no effect on the CRB at moderate and high SNR.
- For a given scenario and observation model, the CRB at a fixed E_s/N_0 increases with increasing M. This indicates that frequency estimation becomes more difficult for larger constellations. This effect is more pronounced for the estimation of F irrespective of θ (with ACRB being proportional to $(E_s/N_0)^{-2M}$), than for joint estimation of F and θ (with ACRB being proportional to $(E_s/N_0)^{-M}$).

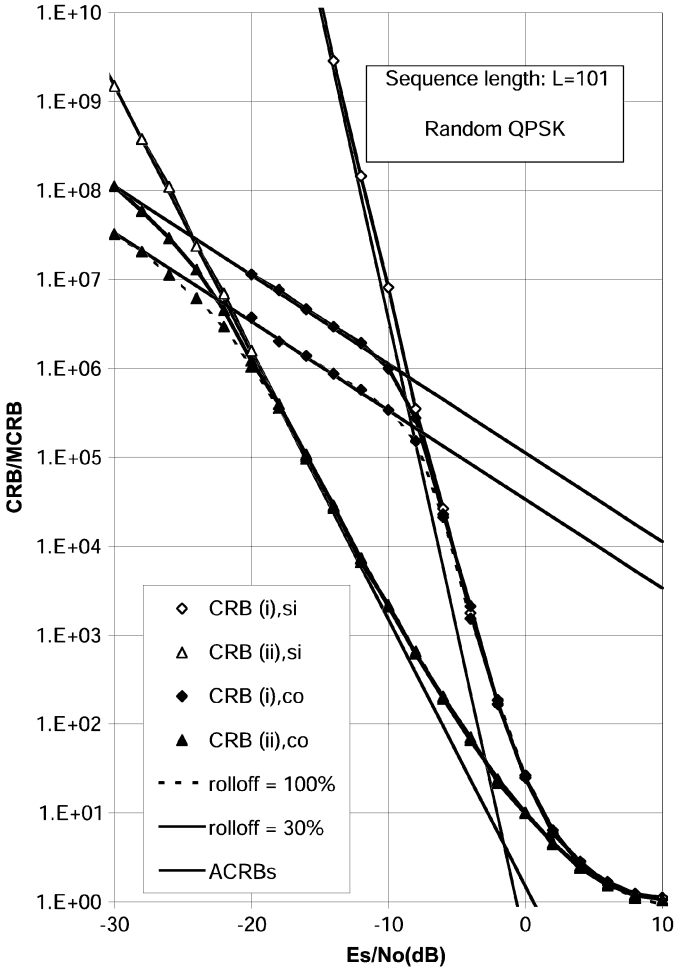

Fig. 4. CRB related to carrier frequency estimation irrespective of [scenario 1)] and jointly with [scenario 2)] the carrier phase for a QPSK constellation (si: simplified observation model, co: correct observation model).

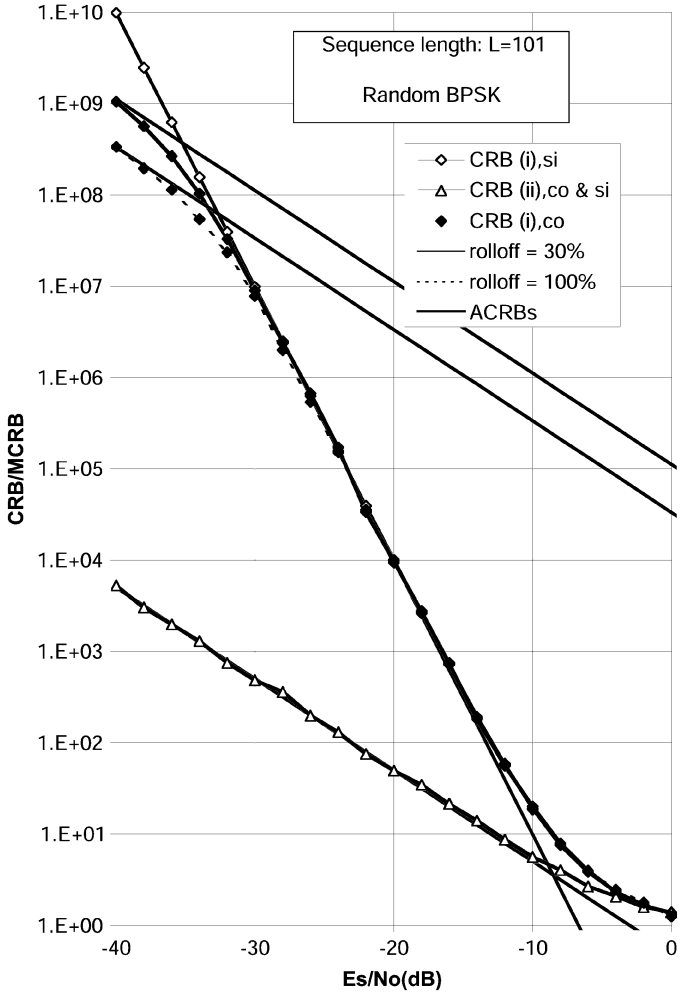

Fig. 5. CRB related to carrier frequency estimation irrespective of [scenario 1)] and jointly with [scenario 2)] the carrier phase for a BPSK constellation (si: simplified observation model, co: correct observation model).

- When E_s/N_0 is sufficiently large, the CRB converges to the MCRB. The value of E_s/N_0 at which the CRB comes close to the MCRB increases with the number M of constellation points.
- For small enough SNR, the CRB corresponding to the correct observation model is very close to the associated ACRB. Increasing SNR yields a CRB that for large M clearly exceeds the ACRB. In this context, it is instructive to consider the high SNR limit (i.e., $E_s/N_0 \to \infty$)ACRB_∞ of the CRB when the constellation size is infinite (which corresponds to continuous instead of discrete data symbols). Using the method outlined in [4], we obtain

$$\mathrm{ACRB}_\infty = \frac{N_0}{2E_s}\frac{1}{L4\pi^2\int\limits_{-\infty}^{+\infty} t^2h^2(t)\,dt} \tag{41}$$

which is proportional to $(E_s/N_0)^{-1}$ and to L^{-1}. Expression (41) indicates that for large E_s/N_0, the CRB in the case of infinite-size constellations does not approach the MCRB (9). According to [4], this is because of the nondiagonal nature of the Fisher information matrix, related to the joint estimation of the continuous data symbols and the frequency offset. For large but finite M, with increasing SNR, the CRB tends to ACRB_∞, but a further increase of SNR eventually makes the CRB approach the MCRB.

C. Random Data Symbols, $M = 2$

We have pointed out in Section III-A and III-C that in the case of random data symbols with $M = 2$, both observation models yield essentially the same ACRB when F is estimated jointly with θ. However, when estimating F irrespective of θ, the ACRBs resulting from the two observation models are much different from each other.

Assuming $M = 2$, Fig. 5 shows the ratios CRB/MCRB and the corresponding ACRB/MCRB, for scenarios 1) and 2), and for both observation models. The behavior of the various curves is as follows.

- As for $M > 2$, the CRBs converge to the corresponding ACRBs (to the MCRB) for small (large) SNR. The CRB resulting from scenario 1) is larger than the CRB corresponding to scenario 2), when the observation model is given. Also, when the scenario is given, the correct observation model yields the smaller CRB.

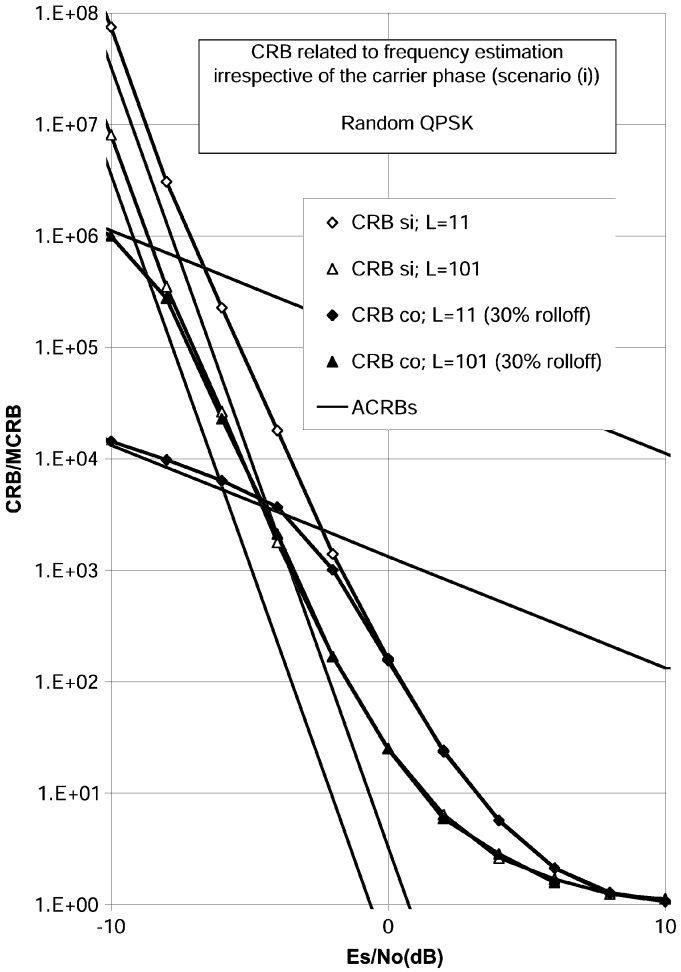

Fig. 6. CRB related to frequency estimation irrespective of the carrier phase [scenario 1)], based on random QPSK data symbols (si: simplified observation model, co: correct observation model).

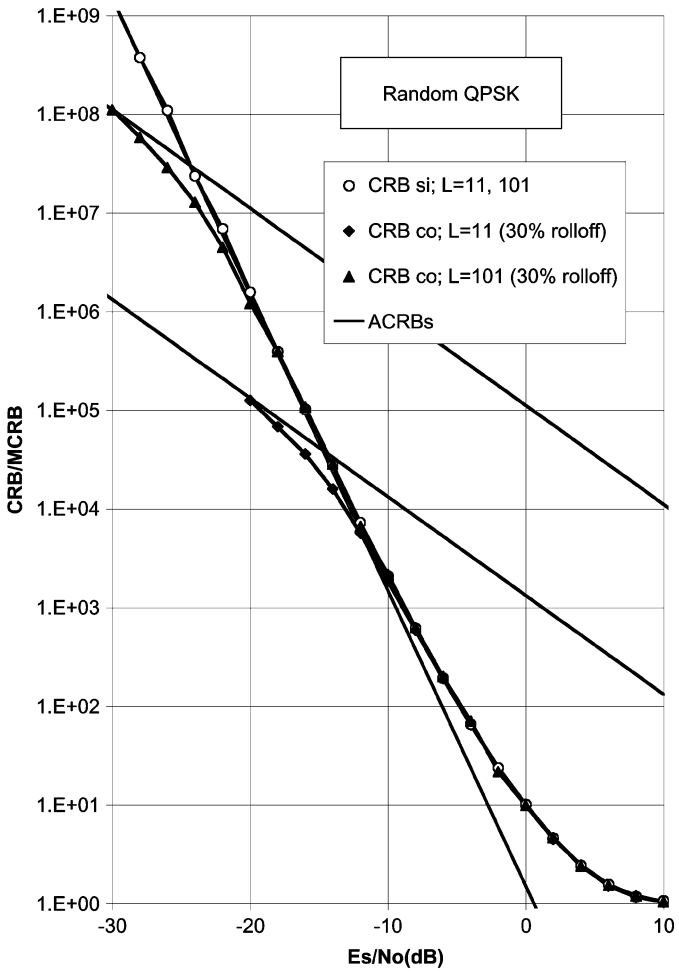

Fig. 7. CRB related to joint phase and frequency estimation [scenario 2)], based on random QPSK data symbols (si: simplified observation model, co: correct observation model).

- For scenario 2), both observation models yield essentially the same CRB.
- For scenario 2), the CRBs resulting from the two observation models behave in a similar way as for $M > 2$. When E_s/N_0 exceeds -30 dB, both observation models yield essentially the same CRB.
- For SNR values of practical interest, the CRB for any scenario and any observation model is very close to the MCRB.

D. Effect of Observation Interval (Random Data)

For both observation models and both scenarios, we consider the ratio CRB/MCRB and the corresponding ACRB/MCRB for different lengths L of the observation interval, assuming random QPSK modulation. We have verified that the same behavior applies to other M-PSK constellations.

For both observation models, the results related to scenarios 1) and 2) are shown in Figs. 6 and 7, respectively. In scenario 1), the ratio CRB/MCRB at SNR values of practical interest decreases with L, whereas in scenario 2), this ratio does not depend on L. This is consistent with the observation that for the simplified model, which is relevant for practical values of SNR, ACRB is proportional to L^{-4} (to L^{-3}) in scenario 1) [in scenario 2)]. Hence, taking into account that the MCRB is proportional to L^{-3}, it follows that for the simplified observation model, ACRB/MCRB is proportional to L^{-1} (independent of L) in scenario 1) [in scenario 2)].

Fig. 8 compares the ratios CRB/MCRB for both scenarios 1) and 2), assuming the simplified observation model. We have verified that, for the ranges of E_s/N_0 and of CRB/MCRB considered in Fig. 8, the correct observation model yields virtually the same curves. For given E_s/N_0, increasing L makes $\mathrm{CRB}^{(i),\mathrm{ra}}$ approach $\mathrm{CRB}^{(ii),\mathrm{ra}}$, and hence, reduces the penalty, caused by treating θ as a nuisance parameter; for given L, this penalty is considerably larger than in the case of known data symbols. The ratio $\mathrm{CRB}^{(ii),\mathrm{ra}}/\mathrm{MCRB}$ being independent of L indicates that the penalty, caused by treating the data symbols as nuisance parameters, cannot be reduced by increasing the observation interval.

V. Conclusions and Remarks

In this paper, we have considered the CRB related to the carrier frequency estimation of a noisy linearly modulated signal with arbitrary square-root Nyquist transmit pulse. We have handled both random and *a priori* known data symbols,

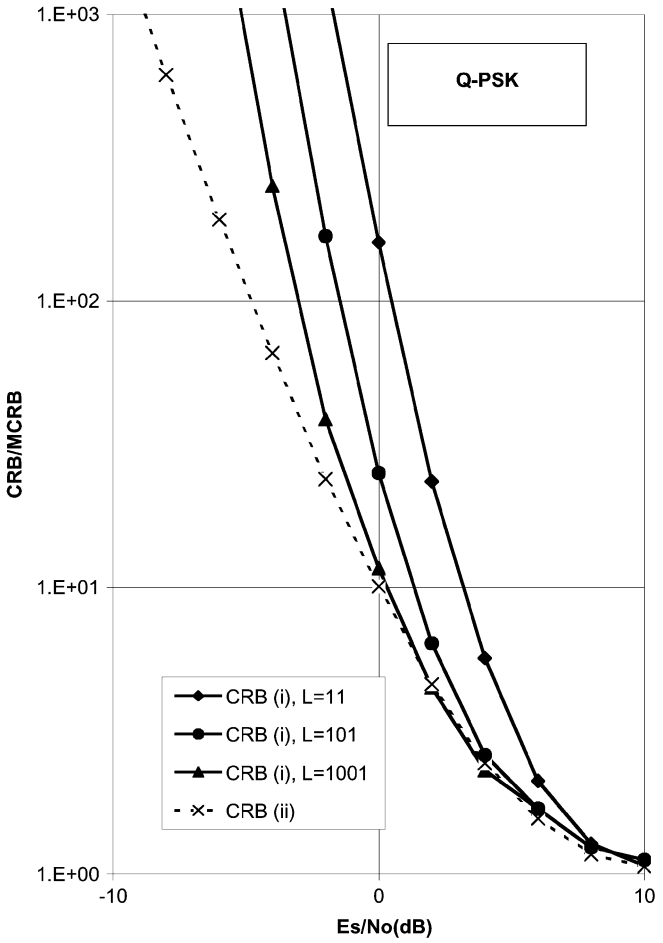

Fig. 8. Effect of the observation interval on the CRB related to frequency estimation irrespective of the carrier phase [scenario 1)] and to joint phase and frequency estimation [scenario 2)].

and have contrasted the results obtained from the correct observation model with those resulting from an approximation of the correct model, used in [5]–[7]. The CRB for joint phase and frequency estimation [scenario 2)] has been compared to the CRB for frequency estimation irrespective of the carrier phase [scenario 1)]. The numerical evaluation of the CRBs requires the approximation of each statistical expectation by an arithmetical average and, for scenario 1), also numerical integration is needed. These averages and numerical integration depend on E_s/N_0 and on the constellation size (and, for scenario 1), on the transmitted sequence length), but not on the pulse shape. The effect of the pulse shape is analytically accounted for.

The numerical results indicate that the correct and the simplified model yield CRBs that are substantially different only at low E_s/N_0. The influence of the pulse shape is restricted to these low SNR values. For both small E_s/N_0 and very large E_s/N_0, the effect of the constellation on the CRB is small. For moderate E_s/N_0, the CRB increases with increasing constellation size. Frequency estimation irrespective of the carrier phase yields a larger CRB than joint frequency and phase estimation, which implies that the latter strategy is potentially the better one. For given SNR, the penalty of the former strategy with respect to the latter decreases with increasing observation interval. When the data symbols are known, this penalty can be neglected for practical values of SNR, even for moderate observation intervals. In the case of random data symbols, considerably longer observation intervals are required to make the penalty very small.

When the carrier frequency and phase are estimated jointly, we have considered the CRB for frequency estimation only. As for both observation models the frequency and phase are not coupled (i.e., $J_{F\theta} = 0$), the CRB for phase estimation is the same as in the case where F is *a priori* known to be zero. For $F = 0$, the two observation models are equivalent, and the resulting CRB for phase estimation is the same as in [5] and [6].

The simplified model of the matched-filter output samples ignores the ISI and the reduction of the useful symbol magnitude, which are caused by the frequency offset F at the input of the matched filter. For E_s/N_0 up to 30 dB, we have verified (details not reported here) that the useful signal magnitude is reduced by less than 0.01 dB and the ISI power is at least 20 dB below the noise power, provided that $|FT| < 0.015$ ($|FT| < 0.030$) for a rolloff factor of 20% (of 100%). Hence, the simplified observation model is valid as long as the maximum frequency offset is on the order of 1% of the symbol rate.

The behavior of the true CRB for frequency estimation in the presence of coding is a topic for further research.

APPENDIX
HIGH AND LOW SNR LIMITS OF THE TRUE CRBS

A. High SNR Limits of the True CRBs

For large E_s/N_0, we obtain an approximation of the CRB by keeping in the summation over the constellation points only the dominant term, i.e., the term with $\alpha_i = a_k$.

1) Scenario 1): Keeping only the dominant terms and taking into account the decomposition (17) of x_k, we obtain for the quantities $A(E_s/N_0, M, L)$, and $B(E_s/N_0, M, L)$ from (24) and (25)

$$A\left(\frac{E_s}{N_0}, M, L\right) = E[\Im^2(a_k^* x_k)] - (E[\Im(a_k^* x_k)])^2 = \frac{1}{2} \tag{A.1}$$

$$B\left(\frac{E_s}{N_0}, M, L\right) = E[|a_k|^2] = 1. \tag{A.2}$$

Substitution of (A.1) in (23a) yields $1/\mathrm{MCRB}_F$; substitution of (A.1) and (A.2) in (23b) yields the sum of $1/\mathrm{MCRB}_F$ and an additional term, which can be neglected for large L.

2) Scenario 2): Keeping only the dominant terms and taking into account the decomposition (17) of $x_k e^{-j\theta}$, we obtain for the quantities $C(E_s/N_0, M)$ and $D(E_s/N_0, M)$ from (38) and (39)

$$C\left(\frac{E_s}{N_0}, M\right) = E[\Im^2(x_k a_k^*)] = \frac{1}{2} \tag{A.3}$$

$$D\left(\frac{E_s}{N_0}, M\right) = E[|a_k|^2] = 1. \tag{A.4}$$

Substitution of (A.3) in (37a) yields $1/\mathrm{MCRB}_F$; substitution of (A.3) and (A.4) in (37b) yields the sum of $1/\mathrm{MCRB}_F$ and an additional term, which again can be neglected for large L.

B. Low SNR Limits of the LLFs

For small E_s/N_0, we obtain an approximation of the LLF by expanding the exponential function in (8) into a Taylor series, and keeping only the relevant terms that correspond to the smallest powers of ε

$$\exp(\varepsilon a_k^* x_k e^{-j\theta} + \varepsilon a_k x_k^* e^{j\theta}) = 1 + \sum_{p=1}^{\infty}\sum_{q=0}^{p} P(p,q,x_k e^{-j\theta})\varepsilon^p (a_k^*)^{p-q} a_k^q \quad \text{(A.5)}$$

where

$$P(p,q,x_k e^{-j\theta}) = \frac{x_k^{p-q} e^{-j(p-q)\theta} x_k^{*q} e^{jq\theta}}{q!(p-q)!}. \quad \text{(A.6)}$$

1) Random Data Symbols: The random data symbols have to be considered as nuisance parameters. Consequently, in (8), averaging over the data is required

$$E_{a_k}[\exp(\varepsilon a_k^* x_k e^{-j\theta} + \varepsilon a_k x_k^* e^{j\theta})] = 1 + \sum_{p=1}^{\infty}\sum_{q=0}^{p} P(p,q,x_k e^{-j\theta})\varepsilon^p E[(a_k^*)^{p-q} a_k^q]. \quad \text{(A.7)}$$

Since we assume a M-PSK constellation, we obtain

$$E[(a_k^*)^{p-q} a_k^q] = 0, \quad 2q-p \notin \{0, \pm M, \pm 2M, \ldots\}$$
$$E[(a_k^*)^{p-q} a_k^q] = 1, \quad 2q-p \in \{0, \pm M, \pm 2M, \ldots\}.$$

Hence, the LLF $\ln p(\mathbf{r}|\mathbf{u})$ can be approximated as shown in (A.8) at the bottom of the page, where $\mathbf{w}$ is a vector of nuisance parameters other than the data. All other nonzero terms in (A.8) contain a power of ε larger than M.

a) Scenario 1); Simplified Observation Model: Take in (A.8) $\mathbf{w} = \theta$ and $x_k = \breve{r}_k = r_k e^{-j2\pi F kT}$. The terms in $|x_k|^{2l}$ are independent of F, and for small ε, these terms can be neglected, as compared with the term 1. Keeping only the terms that correspond to the smallest powers of ε and taking into account that $E_\theta[e^{jx\theta}] = 0, \forall x \neq 0$ (with x integer), we obtain for the LLF

$$\ln p(\mathbf{r}\,|\,F) = \ln\left(E_\theta\left[\prod_{k=-K}^{K}\begin{pmatrix} 1 + \frac{1}{M!}\varepsilon^M \breve{r}_k^M e^{-jM\theta} \\ + \frac{1}{M!}\varepsilon^M \breve{r}_k^{*M} e^{jM\theta}\end{pmatrix}\right]\right) = \ln\left(1 + \sum_{k=-K}^{K}\sum_{\substack{m=-K \\ m\neq k}}^{K} \frac{\varepsilon^{2M}}{(M!)^2} r_k^M r_m^{*M} e^{j2\pi F(m-k)TM}\right) \cong \sum_{k=-K}^{K}\sum_{\substack{m=-K \\ m\neq k}}^{K} \frac{\varepsilon^{2M}}{(M!)^2} r_k^M r_m^{*M} e^{j2\pi F(m-k)TM}. \quad \text{(A.9)}$$

b) Scenario 1); Correct Observation Model: Take in (A.8) $\mathbf{w} = \theta$ and $x_k = \breve{z}_k$. In this case, $|x_k|^{2l}$ is function of F, so that higher order terms can be neglected, as compared with the term in ε^2. Taking into account that $E_\theta[e^{jx\theta}]=0, \forall x \neq 0$ (with x integer), we obtain for the LLF

$$\ln(p(\mathbf{r}\,|\,F,\theta)) = \ln\left(E_\theta\left[\prod_{k=-K}^{K}(1+\varepsilon^2|\breve{z}_k|^2)\right]\right) \cong \ln\left(1 + \sum_{k=-K}^{K}\varepsilon^2|\breve{z}_k|^2\right) \cong \sum_{k=-K}^{K}\varepsilon^2|\breve{z}_k|^2. \quad \text{(A.10)}$$

c) Scenario 2); Simplified Observation Model: Computed in [7].

d) Scenario 2); Correct Observation Model: Take $x_k = \breve{z}_k$; no average has to be taken. Again, $|x_k|^{2l}$ is function of F, so that higher order terms can be neglected, as compared with the term in ε^2. We must distinguish between $M > 2$ and $M = 2$. For $M > 2$, the terms in ε^M can be neglected, as compared with the term in ε^2. We obtain for the LLF

$$\ln(p(\mathbf{r}\,|\,F,\theta)) = \sum_{k=-K}^{K}\ln(1+\varepsilon^2|\breve{z}_k|^2) \cong \sum_{k=-K}^{K}\varepsilon^2|\breve{z}_k|^2. \quad \text{(A.11)}$$

For $M = 2$, the LLF becomes

$$\ln p(\mathbf{r}\,|\,F,\theta) \cong \sum_{k=-K}^{K}\left(\varepsilon^2|\breve{z}_k|^2 + \frac{1}{2}\varepsilon^2\,\breve{z}_k^2\,e^{-j2\theta} + \frac{1}{2}\varepsilon^2(\breve{z}_k^*)^2 e^{j2\theta}\right) = 2\varepsilon^2\sum_{k=-K}^{K}\Re^2[\breve{z}_k e^{-j\theta}]. \quad \text{(A.12)}$$

This indicates that for $M = 2$, only the in-phase component of $\breve{z}_k$ is needed to estimate (F,θ).

2) Known Data Symbols: No averaging over the data symbols is required. This implies that in (A.5), the first-order terms are the dominant terms. Hence, keeping only the terms that correspond to the smallest powers of ε, (8) can be approximated as follows:

$$\ln p(\mathbf{r}\,|\,\mathbf{u}) = \ln\left(E_{\mathbf{v}}\left[\prod_{k=-K}^{K}\begin{pmatrix}1 + \varepsilon a_k^* x_k e^{-j\theta} \\ + \varepsilon a_k x_k^* e^{j\theta}\end{pmatrix}\right]\right) \cong E_{\mathbf{v}}\left[\begin{matrix}\varepsilon\sum_{k=-K}^{K}(a_k^* x_k e^{-j\theta} + a_k x_k^* e^{j\theta}) \\ +\varepsilon^2\sum_{k=-K}^{K}\sum_{\substack{m=-K \\ m\neq k}}^{K} a_k^* a_m x_k x_m^*\end{matrix}\right]. \quad \text{(A.13)}$$

$$\ln p(\mathbf{r}\,|\,\mathbf{u}) \cong \ln\left(E_{\mathbf{w}}\left[\prod_{k=-K}^{K}\begin{pmatrix}1 + \sum_{l=1}^{\frac{M}{2}}\frac{\varepsilon^{2l}|x_k|^{2l}}{(l!)^2} + \frac{1}{M!}\varepsilon^M x_k^M e^{-jM\theta} \\ + \frac{1}{M!}\varepsilon^M (x_k^*)^M e^{jM\theta}\end{pmatrix}\right]\right) \quad \text{(A.8)}$$

e) Scenario 1); Simplified Observation Model: Take in (A.13) $\mathbf{v} = \theta$ and $x_k = \breve{r}_k = r_k e^{-j2\pi F\, kT}$. As the average of the term in ε over θ is zero, we obtain for the LLF

$$\ln p(\mathbf{r} \mid F) = \varepsilon^2 \sum_{k=-K}^{K} \sum_{\substack{m=-K \\ m\neq k}}^{K} a_k^* a_m r_k r_m^* e^{j2\pi F(m-k)T}. \tag{A.14}$$

f) Scenario 1); Correct Observation Model: Take in (A.13) $\mathbf{v} = \theta$ and $x_k = \breve{z}_k$. Again, only the terms in ε^2 should be considered, as the average of the terms in ε over θ is zero. We obtain for the LLF

$$\ln p(\mathbf{r} \mid F) = \varepsilon^2 \sum_{k=-K}^{K} \sum_{\substack{m=-K \\ m\neq k}}^{K} a_k^* a_m \breve{z}_k\, \breve{z}_m^* . \tag{A.15}$$

g) Scenario 2); Correct and Simplified Observation Model: In this case, $\mathrm{CRB} = \mathrm{ACRB} = \mathrm{MCRB}$.

C. Low SNR Limits of the True CRBs

The application of (3) to (A.9), (A.10), (A.11), (A.12), (A.14), and (A.15) is straightforward but tedious. Keeping only the relevant terms corresponding to the smallest powers of ε, and taking into account the decomposition (17) of x_k or $x_k e^{-j\theta}$, we obtain (the dominant part of) $\mathbf{J}$. Inverting $\mathbf{J}$ then gives rise to the low SNR limits of the CRBs.

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. 24, pp. 1391–1399, Feb.–Apr. 1994.

[3] F. Gini, R. Reggiannini, and U. Mengali, "The modified Cramer–Rao bound in vector parameter estimation," *IEEE Trans. Commun.*, vol. 46, pp. 52–60, Jan. 1998.

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

[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] H. 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 Commun. Lett.*, vol. 5, pp. 215–217, May 2001.

[8] N. Noels, H. Steendam, and M. Moeneclaey, "The true Cramer–Rao bound for timing recovery from a bandlimited linearly modulated wafeform," in *Proc. IEEE Int. Conf. Communications*, Paper D08-1, New York, Apr. 2002.

Nele Noels received the diploma of electrical engineering from Ghent University, Gent, Belgium in 2001. She is currently working toward the Ph.D. degree at the Department of Telecommunications and Information Processing, Ghent University.

Her main research interests are in carrier and symbol synchronization. She is the author of several papers in international journals and conference proceedings.

Heidi Steendam (M'00) received the Diploma and the Ph.D. degree, both in electrical engineering, from Ghent University, Gent, Belgium, in 1995 and 2000, respectively.

She is currently a Professor in the Department of Telecommunications and Information Processing, Ghent University. Her main research interests are in statistical communication theory, carrier and symbol synchronization, bandwidth-efficient modulation and coding, spread spectrum (multicarrier spread spectrum), satellite and mobile communication. She is the author of more than 50 scientific papers in international journals and conference proceedings.

Marc Moeneclaey (M'93–SM'99–F'02) received the Diploma and the Ph.D. degree, both in electrical engineering, from Ghent University, Gent, Belgium, in 1978 and 1983, respectively.

He is currently a Professor in the Department of Telecommunications and Information Processing, Ghent University. His main research interests are in statistical communication theory, carrier and symbol synchronization, bandwidth-efficient modulation and coding, spread spectrum, satellite and mobile communication. He is the author of more than 200 scientific papers in international journals and conference proceedings. Together with H. Meyr (RWTH Aachen) and S. Fechtel (Siemens AG), he is a coauthor of the book *Digital Communication Receivers—Synchronization, Channel Estimation, and Signal Processing* (New York: Wiley, 1998).

Between 1992–1994, Dr. Moeneclaey served as Editor for Synchronization for the IEEE TRANSACTIONS ON COMMUNICATIONS.