Carrier Phase and Frequency Estimation for Pilot-Symbol Assisted Transmission: Bounds and Algorithms

N. Noels, *Student Member, IEEE*, H. Steendam, *Member, IEEE*, M. Moeneclaey, *Fellow, IEEE*, and Herwig Bruneel

***Abstract*—In this paper we consider the Cramer–Rao lower bound (CRB) for the joint estimation of the carrier phase and the frequency offset from a noisy linearly modulated burst signal containing random data symbols (DSs) as well as known pilot symbols (PSs). We point out that the CRB depends on the location of the PSs in the burst, the number of PSs, the number of DSs, the signal-to-noise ratio (SNR), and the data modulation scheme. Distributing the PSs symmetrically about the center of the burst and estimating the carrier phase in the center of the burst interval decouples the frequency and phase estimation, making the CRB for phase estimation independent of the specific location of the PSs. At low and moderate SNR, the CRBs for both phase and frequency estimation decrease as the fraction of the PSs in the burst increases. In addition, the CRB for frequency estimation decreases as the PSs are separated with more DSs. Numerical evaluation of the CRB indicates that the carrier phase and frequency of a "hybrid" burst (i.e., containing PSs and DSs) can be estimated more accurately when exploiting both the presence of the DSs and the a priori knowledge about the PSs, instead of using only the knowledge about the PSs (and ignoring the DSs), or considering all the received symbols (PSs and DSs) as unknown (and ignoring the knowledge about the PSs). Comparison of the CRB with the performance of existing carrier synchronizers shows that the iterative soft-decision-directed (sDD) estimator with data-aided (DA) initialization performs very closely to the CRB and provides a large improvement over the classical non-data-aided (NDA) estimator at lower SNR.**

***Index Terms*—Carrier synchronization, Cramer–Rao bound, frequency estimation, phase estimation.**

I. Introduction

In burst digital transmission with coherent detection, the recovery of the carrier phase and the frequency offset is a key aspect. We assume that phase coherence over successive bursts cannot be maintained, so that the carrier phase and frequency offset have to be recovered on a burst-by-burst basis.

Most classical synchronizers belong to one of the following types: data-aided (DA) synchronization algorithms use known pilot symbols (PSs), while non-data-aided (NDA) and decision-directed (DD) estimators operate on modulated data symbols (DSs). DD estimators are similar to DA estimators, but use, instead of PSs, hard or soft decisions regarding the DSs that are provided by the detector; NDA estimators apply a nonlinearity to the received signal to remove the data modulation.

Assuming that the parameter estimate is unbiased, the variance of the estimation error is often used as a performance measure. The Cramer–Rao lower bound (CRB) is a fundamental lower bound on the variance of any unbiased estimate [1], and is also known to be asymptotically achievable for a large enough number of observations, under mild regularity conditions. The $\mathrm{CRB}_{\mathrm{PS}}(\mathrm{N_p})$ for phase and/or frequency estimation from $\mathrm{N_p}$ known PSs has been derived in [2] and [3]. The $\mathrm{CRB}_{\mathrm{DS}}(\mathrm{N}_d)$ related to joint carrier phase and frequency estimation from N_d random DSs has been addressed in [4]–[6]. In the latter case, the statistics of the observation depend not only on the vector parameter to be estimated but also on a nuisance vector parameter (i.e., the unknown DSs) 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 [7] and [8]. The MCRB is much easier to evaluate than the CRB, but is in general looser (i.e., lower) than the true CRB, especially at lower signal-to-noise ratio (SNR). In [9], the high-SNR limit of the $\mathrm{CRB}_{\mathrm{DS}}(\mathrm{N}_d)$ has been obtained analytically, and has been shown to coincide with $\mathrm{MCRB}_{\mathrm{DS}}(\mathrm{N}_d)$.

Very often it may be beneficial for carrier synchronizers to utilize information on both PSs and DSs in the estimation process. In [10], it has been shown that a frequency estimator that utilizes both PSs and DSs may provide the combined advantages of DA estimators and NDA estimators and allow more accurate synchronization at lower SNR. A similar observation holds for DA estimators and DD estimators. The proper operation of DD estimators requires an accurate initialization, which, at low SNR, can only be provided by a DA estimator using known PSs. At the same time, exploiting the DSs guarantees a good performance at high SNR. Note that the PSs also allows to resolve the ambiguity of the NDA and DD phase estimates caused by the rotational symmetry of the constellation.

In this paper, we derive the true $\mathrm{CRB}_{\mathrm{PS-DS}}(\mathrm{N_p}, \mathrm{N}_d)$ for joint phase and frequency estimation from the observation of a "hybrid" burst that contains $\mathrm{N_p}$ pilot symbols as well as N_d data symbols. These CRBs can be viewed as a generalization of the CRBs derived in [2]–[6]. Numerical results are reported for a quaternary phase-shift keying (QPSK) constellation, indicating that hybrid algorithms that exploit both PSs and DSs (in some intelligent way) are potentially more accurate to estimate the carrier phase and frequency from a hybrid burst than algorithms that only use the PSs (and ignore the DSs) or algorithms that use all burst symbols (PSs and DSs) but ignore the knowledge

Manuscript received November 7, 2003. This work was supported by the Interuniversity Attraction Poles Program P11/, Belgian Science Policy. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Jian Li.

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

Digital Object Identifier 10.1109/TSP.2005.859318

of the PSs. Comparing the obtained CRBs to the performance of the hybrid estimation algorithm from [10], it is concluded that more efficient hybrid algorithms may exist that perform more closely to the CRBs. We show that the iterative soft-DD (sDD) estimator with DA initialization yields a close agreement between the simulated performance and the new CRBs.

II. Problem Formulation

Consider the transmission, of a signal with digital linear modulation, over an additive white Gaussian noise channel with unknown carrier phase and frequency offset. Assuming ideal timing recovery, the matched filter output samples are given by

$$r_k = a_k e^{j\theta_k} + w_k, \quad \mathrm{k} \in \mathrm{I} = \{\mathrm{K}_1, \mathrm{K}_1 + 1, \ldots, \mathrm{K}_2\}. \quad (1)$$

In (1), $\{\mathrm{a}_k : \mathrm{k} \in \mathrm{I}\}$ is a sequence of $\mathrm{L} = \mathrm{K}_2 - \mathrm{K}_1 + 1$ transmitted phase-shift keying (PSK), quadrature amplitude modulation, or pulse amplitude modulation symbols. We assume a_k belongs to the symbol alphabet $\{\alpha_0, \alpha_1, \ldots, \alpha_{\mathrm{M}-1}\}$, with M denoting the number of constellation points and $\mathrm{E}[|\mathrm{a}_k|^2] = 1$. The symbol a_k denotes a known PS for k belonging to the set of indexes $\mathrm{I_p} = \{\mathrm{k}_0, \mathrm{k}_1, \ldots, \mathrm{k}_{\mathrm{Np}-1}\} \subseteq \mathrm{I}$, where $\mathrm{N_p}$ denotes the number of PSs. For $\mathrm{k} \in \mathrm{I}_d = \{\mathrm{I} \backslash \mathrm{I_p}\}$, a_k denotes an unknown DS. The N_d $(= \mathrm{L} - \mathrm{N_p})$ DSs are assumed to be statistically independent and uniformly distributed over the constellation, i.e., the transmitted DS can take any value from the symbol alphabet with equal probability. The sequence $\{\mathrm{w}_k : \mathrm{k} \in \mathrm{I}\}$ consists of zero-mean complex Gaussian noise variables, with independent real and imaginary parts each having a variance of $\mathrm{N}_0/2\mathrm{E}_s$. The quantities E_s and N_0 denote the symbol energy and the noise power spectral density (SNR $= \mathrm{E}_s/\mathrm{N}_0$), respectively. The quantity θ_k is defined as $(\theta + 2\pi \mathrm{kFT})$, where θ represents the carrier phase at $\mathrm{k} = 0$, F is the frequency offset, and T is the symbol duration. Both θ and F are unknown but deterministic parameters.

Let us denote by $\mathrm{p}(\mathbf{r}; \mathbf{u})$ the probability density function of the observation vector $\mathbf{r}$, where $\mathbf{u}$ is an unknown deterministic vector parameter. Suppose one is able to produce from $\mathbf{r}$ an unbiased estimate $\hat{\mathbf{u}}$ of the parameter $\mathbf{u}$. Then the estimation error covariance matrix $\mathbf{R}_{\hat{\mathbf{u}}-\mathbf{u}} = E[(\hat{\mathbf{u}} - \mathbf{u})(\hat{\mathbf{u}} - \mathbf{u})^T]$ satisfies

$$\mathbf{R}_{\hat{\mathbf{u}}-\mathbf{u}} - \mathbf{J}^{-1}(\mathbf{u}) \geq 0 \quad (2)$$

where the notation $\mathbf{A} \geq 0$ indicates that $\mathbf{A}$ is a positive semidefinite matrix (i.e., $\mathbf{x}^T\mathbf{A}\mathbf{x} \geq 0$, irrespective of $\mathbf{x}$), and $\mathbf{J}(\mathbf{u})$ is the Fisher information matrix (FIM) [1]. The (i, j)th element of $\mathbf{J}(\mathbf{u})$ is given by

$$\mathbf{J}_{\mathrm{ij}}(\mathbf{u}) = E_{\mathbf{r}}\left[\frac{\partial}{\partial u_i}\ln(p(\mathbf{r};\mathbf{u}))\frac{\partial}{\partial u_{\mathrm{j}}}\ln(p(\mathbf{r};\mathbf{u}))\right]. \quad (3)$$

Note that $\mathbf{J}(\mathbf{u})$ is a symmetrical matrix. When the element $\mathbf{J}_{\mathrm{ij}}(\mathbf{u}) = 0$, the parameters u_i and $\mathrm{u_j}$ are said to be *decoupled*. The expectation $\mathrm{E}_{\mathbf{r}}[.]$ in (3) is with respect to $\mathrm{p}(\mathbf{r}; \mathbf{u})$. The probability density $\mathrm{p}(\mathbf{r}; \mathbf{u})$ of **r,** corresponding to a given value of $\mathbf{u}$, is called the *likelihood function* of $\mathbf{u}$; $\ln(\mathrm{p}(\mathbf{r}; \mathbf{u}))$ is the *log-likelihood function* of $\mathbf{u}$. When the observation $\mathbf{r}$ depends not only on the parameter $\mathbf{u}$ to be estimated but also on a *nuisance* vector parameter $\mathbf{v}$, the likelihood function of $\mathbf{u}$ is obtained by averaging the likelihood function $\mathrm{p}(\mathbf{r}|\mathbf{v}; \mathbf{u})$ of the vector $(\mathbf{u}, \mathbf{v})$ over the a priori distribution of the nuisance parameter: $p(\mathbf{r}; \mathbf{u}) = E_{\mathbf{v}}[p(\mathbf{r}|\mathbf{v}; \mathbf{u})]$. We refer to $\mathrm{p}(\mathbf{r}|\mathbf{v}; \mathbf{u})$ as the joint likelihood function, as $\mathrm{p}(\mathbf{r}|\mathbf{v}; \mathbf{u})$ is relevant to the joint maximum likelihood (ML) estimation of $\mathbf{u}$ and $\mathbf{v}$.

Considering the joint estimation of the carrier phase θ and the frequency offset F from the observation vector $\mathbf{r} = \{\mathrm{r}_k\}$ from (1), we take $\mathbf{u} = (\mathrm{u}_1, \mathrm{u}_2) = (\theta, \mathrm{F})$. The nuisance parameter vector $\mathbf{v} = \{\mathrm{a}_k : \mathrm{k} \in \mathrm{I}_d\}$ consists of the unknown DSs. Within a factor not depending on F, θ, and $\mathbf{a}$, the joint likelihood function $\mathrm{p}(\mathbf{r}|\mathbf{a}; \mathrm{F}, \theta)$ is given by

$$p(\mathbf{r}|\mathbf{a}; F, \theta) = \prod_{k \in I} F(a_k, \widetilde{r}_k) \quad (4)$$

where

$$F(a_k, \widetilde{r}_k) = e^{E_s/N_0\left(2\mathrm{Re}(a_k^* \widetilde{r}_k) - |a_k|^2\right)} \quad (5)$$

and $\widetilde{r}_k = r_k e^{-j(2\pi kFT + \theta)}$. Averaging (4) over the data symbols yields the likelihood function $\mathrm{p}(\mathbf{r}; \mathrm{F}, \theta)$. For the log-likelihood function $\ln(\mathrm{p}(\mathbf{r}; \mathrm{F}, \theta))$ we obtain, within a term that does not depend on (F, θ)

$$\ln p(\mathbf{r}; F, \theta) = 2\frac{E_s}{N_0}\mathrm{Re}\left(e^{-j\theta}\sum_{k \in I_{\mathrm{p}}} a_k^* \widetilde{r}_k\right) + \sum_{l \in I_d} \ln I(\widetilde{r}_l) \quad (6)$$

where

$$I(\widetilde{r}_k) = \sum_{i=0}^{M-1} F(\alpha_i, \widetilde{r}_k) \quad (7)$$

and $\{\alpha_0, \alpha_1, \ldots, \alpha_{\mathrm{M}-1}\}$ denotes the set of constellation points.

It follows from (2) that the error variance regarding the estimation of θ and F is lower bounded by the Cramer–Rao bound (CRB)

$$E_{\mathbf{r}}[(\hat{\theta} - \theta)^2] \geq \mathrm{CRB}^{\theta}_{\mathrm{PS-DS}}(N_{\mathrm{p}}, N_d) = \left(\mathbf{J}^{-1}_{(\theta,F)}\right)_{11} \quad (8)$$

$$E_{\mathbf{r}}[(\hat{F} - F)^2] \geq \mathrm{CRB}^{F}_{\mathrm{PS-DS}}(N_{\mathrm{p}}, N_d) = \left(\mathbf{J}^{-1}_{(\theta,F)}\right)_{22} \quad (9)$$

where $\mathbf{J}^{-1}$ denotes the inverse of the FIM. Similarly, (2) yields a lower bound on the variance of the estimation error on the instantaneous phase

$$\begin{aligned} E_r[(\hat{\theta}_k - \theta_k)^2] &\geq \mathrm{CRB}^{\theta_k}_{\mathrm{PS-DS}}(N_{\mathrm{p}}, N_d) \\ &= \left(\mathbf{J}^{-1}_{(\theta,F)}\right)_{11} + 4\pi kT\left(\mathbf{J}^{-1}_{(\theta,F)}\right)_{12} \\ &\quad + 4(\pi kT)^2\left(\mathbf{J}^{-1}_{(\theta,F)}\right)_{22}. \end{aligned} \quad (10)$$

The presence of the nuisance vector parameter $\mathbf{v} = \{\mathrm{a}_k : \mathrm{k} \in \mathrm{I}_d\}$ makes the analytical computation of the FIM $\mathbf{J}(\theta, \mathrm{F})$ very hard. In order to avoid the computational complexity caused by the nuisance parameters, a simpler lower bound, MCRB, has been derived in [7] and [8], i.e., $E[(\hat{x} - x)^2] \geq \mathrm{CRB}^{x}_{\mathrm{PS-DS}}(N_{\mathrm{p}}, N_d) \geq M\mathrm{CRB}^{x}_{\mathrm{PS-DS}}(N_{\mathrm{p}}, N_d)$, where $M\mathrm{CRB}^{x}_{\mathrm{PS-DS}}(N_{\mathrm{p}}, N_d)$ is defined in the same way as

$\mathrm{CRB}^{x}_{\mathrm{PS-DS}}(N_\mathrm{p}, N_d)$ in (8)–(10) but with the FIM $\mathbf{J}(\theta, \mathrm{F})$ replaced with the modified FIM (MFIM) $\mathbf{J}_\mathrm{M}(\theta, \mathrm{F})$ given by

$$\mathbf{J}_\mathrm{M} = \frac{2E_s}{N_0}\begin{bmatrix} \sum\limits_{k\in I}\gamma_k & 2\pi T\sum\limits_{k\in I} k\gamma_k \\ 2\pi T\sum\limits_{k\in I} k\gamma_k & (2\pi T)^2\sum\limits_{k\in I} k^2\gamma_k \end{bmatrix} \tag{11}$$

where

$$\gamma_k = \begin{cases} |a_k|^2, & k \in I_\mathrm{P} \\ 1, & k \in I_d \end{cases}.$$

III. True $\mathrm{CRB}_{\mathrm{PS-DS}}$: Analytical Results

Partial differentiation of the log-likelihood function (6) with respect to the carrier phase θ and the frequency offset F yields

$$\frac{\partial}{\partial\theta}\ln p(\mathbf{r}; F, \theta) = 2\frac{E_s}{N_0}\left(\sum_{k\in I_\mathrm{P}}\mathrm{Im}(a_k^*\tilde{r}_k) + \sum_{l\in I_d} M(\tilde{r}_l)\right) \tag{12}$$

$$\frac{\partial}{\partial F}\ln p(\mathbf{r}; F, \theta) = 4\pi T\frac{E_s}{N_0}\left(\sum_{k\in I_\mathrm{P}} k\mathrm{Im}(a_k^*\tilde{r}_k) + \sum_{l\in I_d} lM(\tilde{r}_l)\right) \tag{13}$$

where

$$M(\tilde{r}_k) = \frac{\sum\limits_{i=0}^{M-1} F(\alpha_i, \tilde{r}_k)\mathrm{Im}(\alpha_i^*\tilde{r}_k)}{I(\tilde{r}_k)}. \tag{14}$$

$\mathrm{F}(.,.)$ and $\mathrm{I}(.)$ are defined as in (5) and (7), respectively, and $\{\alpha_0, \alpha_1, \ldots, \alpha_{\mathrm{M}-1}\}$ denotes the set of constellation points. Substituting (12) and (13) into (3) yields

$$\begin{aligned}\mathbf{J} &= \frac{2E_s}{N_0}\begin{bmatrix} \sum\limits_{k\in I}\beta_k & 2\pi T\sum\limits_{k\in I} k\beta_k \\ 2\pi T\sum\limits_{k\in I} k\beta_k & (2\pi T)^2\sum\limits_{k\in I} k^2\beta_k \end{bmatrix} \\ &= \mathbf{J}_{11}\begin{bmatrix} 1 & 2\pi T k_G \\ 2\pi T k_G & (2\pi T)^2(k_G^2+\sigma_G^2) \end{bmatrix}\end{aligned} \tag{15}$$

where

$$\beta_k = \begin{cases} |a_k|^2, & k \in I_\mathrm{P} \\ \frac{2E_s}{N_0}E_r\left[M(r)^2\right], & k \in I_d \end{cases} \tag{16}$$

and

$$k_G = \frac{\sum\limits_I k\beta_k}{\sum\limits_I \beta_k} \qquad \sigma_G^2 = \frac{\sum\limits_I (k-k_G)^2\beta_k}{\sum\limits_I \beta_k}. \tag{17}$$

In (16), $\mathrm{E}_r[.]$ denotes the average over $r = a + n$, where a is a random variable that takes any value from the symbol alphabet with equal probability and n is complex zero-mean Gaussian noise with independent real and imaginary parts each having a variance equal to $\mathrm{N}_0/2\mathrm{E}_s$. The quantity k_G can be interpreted as the center of gravity of the sequence $\{\beta_k\}$. We obtain $\mathbf{J}_{12} \neq 0$, unless $\mathrm{k}_G = 0$, which is achieved if both PSs and DSs are each located symmetrically about zero, and the PSs satisfy $|\mathrm{a}_k| = |\mathrm{a}_{-k}|$. For $\mathrm{k}_G \neq 0$, the parameters θ and F are coupled, meaning that the inaccuracy in the carrier phase estimate has an impact on the frequency offset estimation and vice versa. Note that the FIM does not depend on θ or F. Substituting (15) into (8)–(10), we obtain

$$\begin{aligned} E[(\hat{\theta}-\theta)^2] &\geq \mathrm{CRB}^{\theta}_{\mathrm{PS-DS}}(N_\mathrm{p}, N_d) \\ &= \frac{1}{\mathbf{J}_{11}}\left(1 + \frac{k_G^2}{\sigma_G^2}\right) \end{aligned} \tag{18}$$

$$E[(\hat{F}T\text{–}FT)^2] \geq \mathrm{CRB}^{FT}_{\mathrm{PS-DS}}(N_\mathrm{p}, N_d) = \frac{1}{4\pi^2\sigma_G^2\mathbf{J}_{11}} \tag{19}$$

$$\begin{aligned} E[(\hat{\theta}_k-\theta_k)^2] &\geq \mathrm{CRB}^{\theta_k}_{\mathrm{PS-DS}}(N_\mathrm{p}, N_d) \\ &= \frac{1}{\mathbf{J}_{11}}\left(1 + \frac{(k-k_G)^2}{\sigma_G^2}\right). \end{aligned} \tag{20}$$

The lower bound on $E[(\hat{\theta}_k - \theta_k)^2]$ from (20) is *quadratic* in k. Its *minimum* value is achieved at $\mathrm{k} = \mathrm{k}_G$ and is equal to $1/\mathbf{J}_{11}$, which is the CRB for the estimation of the carrier phase when the frequency offset is a priori known. Note from (15) and (16) that $1/\mathbf{J}_{11}$ depends on the number ($\mathrm{N_p}$) of PSs, the number (N_d) of DSs, and the particular pilot sequence that was selected, but not on the specific position of the PSs in the burst. Let $\Delta_1 = |\mathrm{K}_1 - \mathrm{k}_G|$ and $\Delta_2 = |\mathrm{K}_2 - \mathrm{k}_G|$ represent the distance (in symbol intervals) between the position of the minimum value of the CRB (20) and the edges of the burst interval I. The bound (20) achieves its *maximum* value at

$$k_{\max} = \begin{cases} K_1, & \Delta_1 \geq \Delta_2 \\ K_2, & \Delta_1 \leq \Delta_2 \end{cases}$$

i.e., at one of the edges of the burst interval I (or at both edges if $\mathrm{k}_G = (\mathrm{K}_2 + \mathrm{K}_1)/2$). The difference between the minimum and the maximum value of (20) over the burst amounts to $\Delta^2 4\pi^2\mathrm{CRB}^{FT}_{\mathrm{PS-DS}}(\mathrm{N_p}, \mathrm{N}_d)$, where $\Delta = \max(\Delta_1, \Delta_2)$. Hence, for given values of $1/\mathbf{J}_{11}$ and $\mathrm{CRB}^{FT}_{\mathrm{PS-DS}}(\mathrm{N_p}, \mathrm{N}_d)$, the detection of symbols located near the edge $\mathrm{k}_{\max}$ suffers from a larger instantaneous phase error variance as Δ increases.

Let us define by $\mathbf{J}_\infty$ and $\mathbf{J}_0$ the high-SNR and low-SNR asymptotic FIM that are obtained as the limit of the FIM from (15) for $\mathrm{E}_s/\mathrm{N}_0 \to \infty$ and $\mathrm{E}_s/\mathrm{N}_0 \to 0$, respectively. It can be verified that $\mathbf{J}_0$ equals the FIM for estimation from the PSs only; it has been shown in [3] that this FIM is given by (15) in which the summation over I is replaced with a summation over $\mathrm{I_p}$ only. This indicates that at very low SNR, DA estimation techniques may perform close to optimal. The high SNR asymptotic FIM $\mathbf{J}_\infty$ equals the MFIM from (11).

Note that, for a PSK-type modulation $|\mathrm{a}_k| = 1$, some further simplification and interpretation of the above results is possible.

1) The ratio $\mathrm{CRB}^{\theta_{k_G}}_{\mathrm{PS-DS}}(N_\mathrm{p}, N_d)/M\mathrm{CRB}^{\theta_{k_G}}_{\mathrm{PS-DS}}(\mathrm{N_p}, N_d)$ only depends on the SNR and on the ratio $\mathrm{N_p}/\mathrm{N}_d$.
2) For very low SNR, k_G converges to the center of the pilot sequence. For very high SNR, k_G converges to the center of the complete burst, i.e., $(\mathrm{K}_2 + \mathrm{K}_1)/2$.
3) Independent of the presence of the PSs (number, value, location), the MFIM (11) related to a burst containing $\mathrm{N_p}$ pilot symbols and N_d data symbols equals the MFIM for transmitting a sequence of $\mathrm{L} = \mathrm{N_p} + \mathrm{N}_d$ unknown DSs

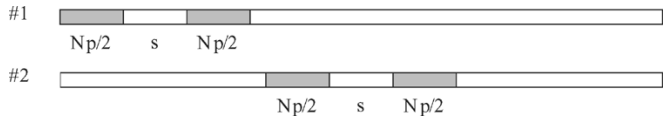


Fig. 1. Burst structure, location of the PSs.

that has been shown to coincide with the high-SNR limit of the FIM for the estimation from L unknown DSs in [9]. This implies that, at high SNR, estimation techniques that make no use of PSs may perform close to optimum.

4) The lower bound on $E[(\hat{FT} - FT)^2]$ from (19) does not depend on the choice of the time origin.

IV. True $\mathrm{CRB}_{\mathrm{PS-DS}}$: Numerical Results and Discussion

Numerical results were obtained for a QPSK constellation and a symmetrical observation interval, i.e., $\mathrm{I} = \{-\mathrm{K}, \ldots, \mathrm{K}\}$. In this case the MFIM from (11) becomes diagonal and the MCRBs reduce to [8]

$$M\mathrm{CRB}^{\theta_{k_G}}_{\mathrm{PS-DS}}(N_\mathrm{p}, N_d) = \mathrm{MCRB}^{\theta}_{\mathrm{DS}}(L) = \frac{N_0}{2LE_s}$$

$$\mathrm{MCRB}^{FT}_{\mathrm{PS-DS}}(N_\mathrm{p}, N_d) = \mathrm{MCRB}^{FT}_{\mathrm{DS}}(L) = \frac{3N_0}{2\pi^2 L(L^2-1)E_s}.$$

We assume a burst of $\mathrm{L} = 2\mathrm{K}+1 = \mathrm{N_p} + \mathrm{N}_d$ symbols, containing two parts of $\mathrm{N_p}/2$ PSs spaced with s DSs, as proposed in [10]. Two different burst structures are considered. They are shown in Fig. 1, where the shaded areas indicate the location of the PSs. In burst structure #1 the PSs are concentrated at the beginning of each burst, whereas burst structure #2 is symmetric yielding $\mathrm{k}_G = 0$, so that carrier phase and frequency estimation are decoupled (with $\theta_{k_G} = \theta$).

Figs. 2 and 3 show the ratio $\mathrm{CRB}_{\mathrm{PS-DS}}(\mathrm{N_p}, \mathrm{N}_d)/\mathrm{MCRB}_{\mathrm{PS-DS}}(\mathrm{N_p}, \mathrm{N}_d)$ as a function of the SNR, for the reference phase error in $\mathrm{k} = \mathrm{k}_G$ and for the frequency error, respectively. Results are presented for $\mathrm{L} = 321$, $\mathrm{N_p}/\mathrm{L}$ equal to (approximately) 10% and 20% ($\mathrm{N_p} = 32, 64$ if $\mathrm{s} \neq 0$ and $\mathrm{N_p} = 33, 65$ if $\mathrm{s} = 0$), and for $\mathrm{s} = 0$, $(\mathrm{s}-1) = \mathrm{N_p}/2$, and $(\mathrm{s}-1) = 3\mathrm{N_p}/2$. For comparison, the lower bound $\mathrm{CRB}_{\mathrm{DS}}(\mathrm{L}) = \mathrm{CRB}_{\mathrm{PS-DS}}(0, \mathrm{L})$ for the estimation from a burst without PSs is also displayed. The gray curves correspond to the lower bounds $\mathrm{CRB}_{\mathrm{PS}}(\mathrm{N_p}) = \mathrm{CRB}_{\mathrm{PS-DS}}(\mathrm{N_p}, 0)$ for the estimation from the PSs only, which are the low-SNR asymptotes of the $\mathrm{CRB}_{\mathrm{PS-DS}}(\mathrm{N_p}, \mathrm{N}_d)$.

A. True CRB for the Estimation of the Reference Phase in $\mathrm{k} = \mathrm{k}_G$

Fig. 2 corresponds to the reference phase estimation error in $\mathrm{k} = \mathrm{k}_G$. As the ratio $\mathrm{CRB}^{\theta_{k_G}}_{\mathrm{PS-DS}}(N_\mathrm{p}, N_d)/\mathrm{MCRB}^{\theta_{k_G}}_{\mathrm{PS-DS}}(N_\mathrm{p}, N_d)$ is determined only by $\mathrm{N_p}/\mathrm{L}$ and $\mathrm{E}_s/\mathrm{N}_0$, the curves for all burst structures with the same ratio $\mathrm{N_p}/\mathrm{L}$ coincide. At low and intermediate SNR, the ratio $\mathrm{CRB}^{\theta_{k_G}}_{\mathrm{PS-DS}}(N_\mathrm{p}, N_d)/\mathrm{MCRB}^{\theta_{k_G}}_{\mathrm{PS-DS}}(N_\mathrm{p}, N_d)$ decreases as $\mathrm{N_p}/\mathrm{L}$ increases. At very low SNR, the $\mathrm{CRB}^{\theta_{k_G}}_{\mathrm{PS-DS}}(N_\mathrm{p}, N_d)$ converges to its low-SNR asymptote that is given by

$$\mathrm{CRB}^{\theta_{k_G}}_{\mathrm{PS}}(N_\mathrm{p}) = \frac{N_0}{2N_\mathrm{p}E_s} \tag{21}$$

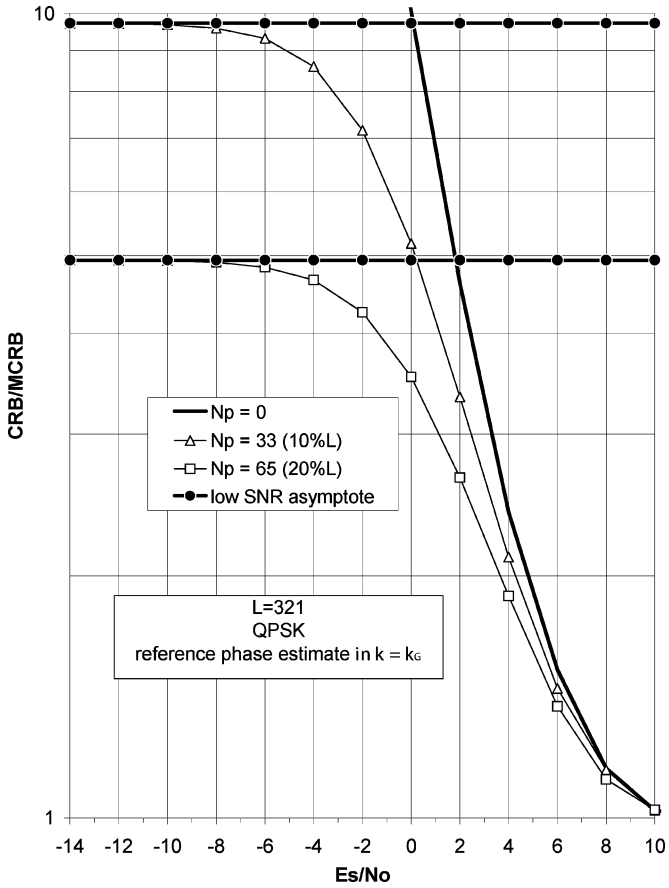


Fig. 2. CRB/MCRB for the reference phase estimate in $\mathrm{k} = \mathrm{k}_G$.

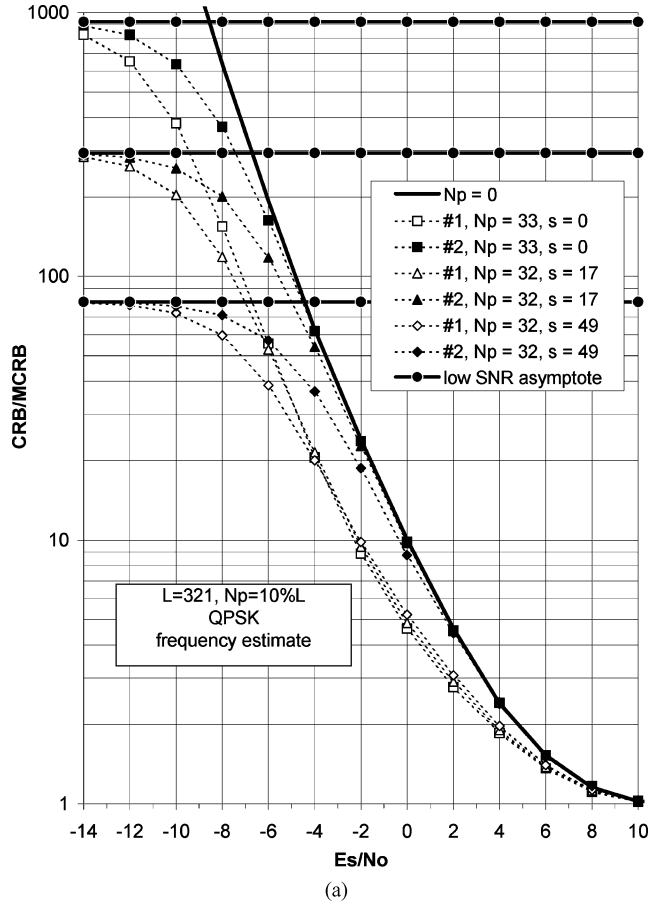
i.e., the ratio $\mathrm{CRB}^{\theta_{k_G}}_{\mathrm{PS-DS}}(N_\mathrm{p}, N_d)/\mathrm{MCRB}^{\theta_{k_G}}_{\mathrm{PS-DS}}(N_\mathrm{p}, N_d)$ converges to $\mathrm{L}/\mathrm{N_p}$. At very high SNR, the $\mathrm{CRB}^{\theta_{k_G}}_{\mathrm{PS-DS}}(N_\mathrm{p}, N_d)$ converges to its high-SNR asymptote, i.e., the ratio $\mathrm{CRB}^{\theta_{k_G}}_{\mathrm{PS-DS}}(N_\mathrm{p}, N_d)/\mathrm{MCRB}^{\theta_{k_G}}_{\mathrm{PS-DS}}(N_\mathrm{p}, N_d)$ converges to one.

B. True CRB for Frequency Estimation

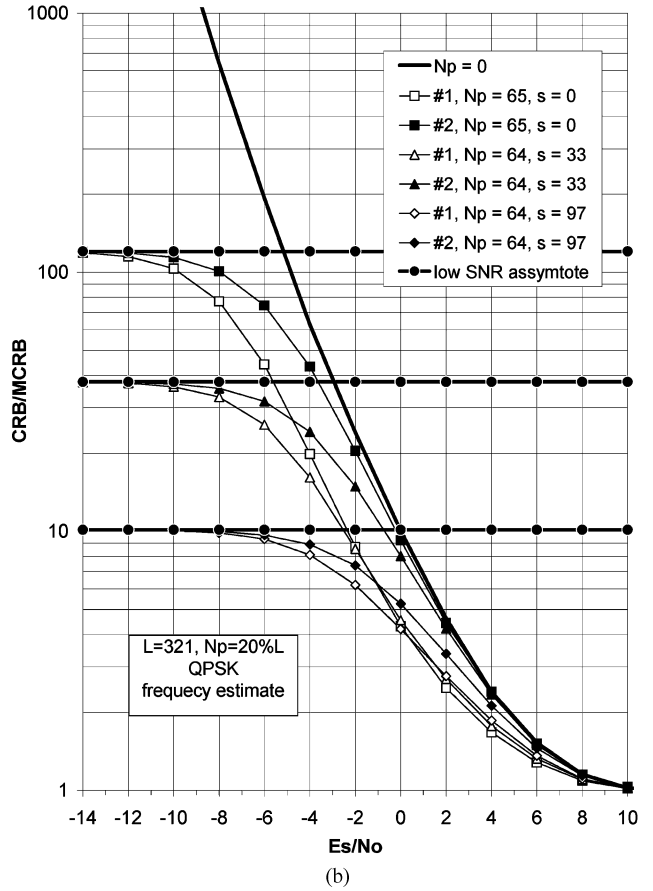
Fig. 3 corresponds to the frequency estimation error. At low and intermediate SNR, increasing the number of PSs ($\mathrm{N_p}$) decreases the ratio $\mathrm{CRB}^{FT}_{\mathrm{PS-DS}}(N_\mathrm{p}, N_d)/\mathrm{MCRB}^{FT}_{\mathrm{PS-DS}}(N_\mathrm{p}, N_d)$ [Fig. 3(a) versus Fig. 3(b)]. In contrast with $\mathrm{CRB}^{\theta_{k_G}}_{\mathrm{PS-DS}}(N_\mathrm{p}, N_d)/\mathrm{MCRB}^{\theta_{k_G}}_{\mathrm{PS-DS}}(N_\mathrm{p}, N_d)$, the ratio $\mathrm{CRB}^{FT}_{\mathrm{PS-DS}}(N_\mathrm{p}, N_d)/\mathrm{MCRB}^{FT}_{\mathrm{PS-DS}}(N_\mathrm{p}, N_d)$ also depends on the specific position of the PSs within the burst. For a fixed number of PSs ($\mathrm{N_p}$), the ratio $\mathrm{CRB}^{FT}_{\mathrm{PS-DS}}(N_\mathrm{p}, N_d)/\mathrm{MCRB}^{FT}_{\mathrm{PS-DS}}(N_\mathrm{p}, N_d)$ decreases as the spacing s increases. At very low SNR, the $\mathrm{CRB}^{FT}_{\mathrm{PS-DS}}(N_\mathrm{p}, N_d)$ becomes close to its low-SNR asymptote that is given by [3], [10]

$$\mathrm{CRB}^{FT}_{\mathrm{PS}}(N_\mathrm{p}) = \frac{3N_0}{2\pi^2 E_s[N_\mathrm{p}(N_\mathrm{p}^2-1) + 3N_\mathrm{p}s(s+N_\mathrm{p})]} \tag{22}$$

i.e., assuming $\mathrm{N_p}, \mathrm{L} \gg 1$, the ratio $\mathrm{CRB}^{FT}_{\mathrm{PS-DS}}(N_\mathrm{p}, N_d)/s\mathrm{MCRB}^{FT}_{\mathrm{PS-DS}}(N_\mathrm{p}, N_d)$ converges to $(1+3\mathrm{q}+3\mathrm{q}^2)^{-1}(\mathrm{L}/\mathrm{N_p})^3$, where $\mathrm{q} = \mathrm{s}/\mathrm{N_p}$. For fixed $\mathrm{N_p}$ and fixed s, the low SNR asymptote of the $\mathrm{CRB}^{FT}_{\mathrm{PS-DS}}(N_\mathrm{p}, N_d)$ is the same for burst structures #1 and #2, as (22) is not affected by a time-shift of the pilot sequence within the burst. At very high SNR, the



(a)



(b)

Fig. 3. CRB/MCRB for the frequency estimate.

$\mathrm{CRB}^{FT}_{\mathrm{PS-DS}}(N_\mathrm{p}, N_d)$ converges to its high-SNR asymptote, i.e., the ratio $\mathrm{CRB}^{FT}_{\mathrm{PS-DS}}(N_\mathrm{p}, N_d)/\mathrm{MCRB}^{FT}_{\mathrm{PS-DS}}(N_\mathrm{p}, N_d)$ converges to one. For both burst structures #1 and #2, it can be easily verified that, assuming $\mathrm{N_p, L} \gg 1$ and for a fixed value of $\mathrm{q} = \mathrm{s}/\mathrm{N_p}$, the ratio $\mathrm{CRB}^{FT}_{\mathrm{PS-DS}}(N_\mathrm{p}, N_d)/\mathrm{MCRB}^{FT}_{\mathrm{PS-DS}}(N_\mathrm{p}, N_d)$ is mainly determined by the ratio $\mathrm{N_p/L}$. Simulation results, not reported here, indicate that the assumption $\mathrm{N_p, L} \gg 1$ holds even for the relatively small values of $\mathrm{N_p}$ and L from Fig. 3.

C. $CRB_{\mathrm{PS-DS}}(\mathrm{N_p}, \mathrm{N}_d)$ Versus $CRB_{\mathrm{PS}}(\mathrm{N_p})$ and $CRB_{\mathrm{DS}}(\mathrm{N_p} + \mathrm{N}_d)$

It follows from both Figs. 2 and 3 that the $\mathrm{CRB_{PS-DS}}(\mathrm{N_p}, \mathrm{N}_d)$ is smaller than both $\mathrm{CRB_{PS}}(\mathrm{N_p})$ and $\mathrm{CRB_{DS}}(\mathrm{N_p} + \mathrm{N}_d)$. This indicates that it is potentially more accurate to estimate the carrier phase and frequency of a hybrid burst with a hybrid algorithm that exploits both PSs and DSs (in some intelligent way) than with an algorithm that only uses the PSs (and ignores the DSs) or with an algorithm that uses all received symbols (PSs+DSs) but ignores the a priori knowledge about the PSs. The ratio $\mathrm{CRB_{PS}}(\mathrm{N_p})/\mathrm{CRB_{PS-DS}}(\mathrm{N_p}, \mathrm{N}_d)$ ($\mathrm{CRB_{DS}}(\mathrm{N_p} + \mathrm{N}_d)/\mathrm{CRB_{PS-DS}}(\mathrm{N_p}, \mathrm{N}_d)$) depends on the operating SNR and on the burst structure, and indicates to what extent synchronizer performance can be improved by making clever use of the presence of the DSs (of the knowledge about the PSs) in the estimation process.

D. Effect of the Burst Structure

For a fixed $\mathrm{N_p}$ and fixed s, burst structures #1 and #2 yield the same $\mathrm{CRB}^{\theta_{kG}}_{\mathrm{PS-DS}}(N_\mathrm{p}, N_d)$, while the asymmetric burst structure #1 yields the smallest $\mathrm{CRB}^{FT}_{\mathrm{PS-DS}}(N_\mathrm{p}, N_d)$ (at any SNR). However, as the following example illustrates, we should be very careful when interpreting these results. Fig. 4 depicts the $\mathrm{CRB}^{\theta_k}_{\mathrm{PS-DS}}(N_\mathrm{p}, N_d)$ for the reference phase error as a function of the symbol index k at $\mathrm{E}_s/\mathrm{N}_0 = 2$ dB for burst structures #1 and #2 with $\mathrm{N_p} = 64$ and $\mathrm{s} = 33$. The following observations can be made.

1) Although burst structure #1 yields the smallest $\mathrm{CRB}^{FT}_{\mathrm{PS-DS}}(N_\mathrm{p}, N_d)$, its CRB on the reference phase error variance at $\mathrm{k} = \mathrm{K}$ is larger than for burst structure #2. This can be explained by noting that, at a value of SNR as low as 2 dB, the distance (Δ) between the positions of the minimum and maximum value of the $\mathrm{CRB}^{\theta_k}_{\mathrm{PS-DS}}(N_\mathrm{p}, N_d)$ is significantly larger for burst structure #1 than for burst structure #2.
2) Although burst structure #2 results in the smallest maximum for $\mathrm{CRB}^{\theta_k}_{\mathrm{PS-DS}}(N_\mathrm{p}, N_d)$ over the burst, other than for burst structure #1, this maximum value is reached near both edges of the burst interval. This implies that in burst structure #2 more symbols are affected by a large instantaneous phase error variance than in burst structure #1.

Hence, the "best" burst structure depends strongly on the operating SNR and on the maximum allowable phase error variance for proper symbol detection.

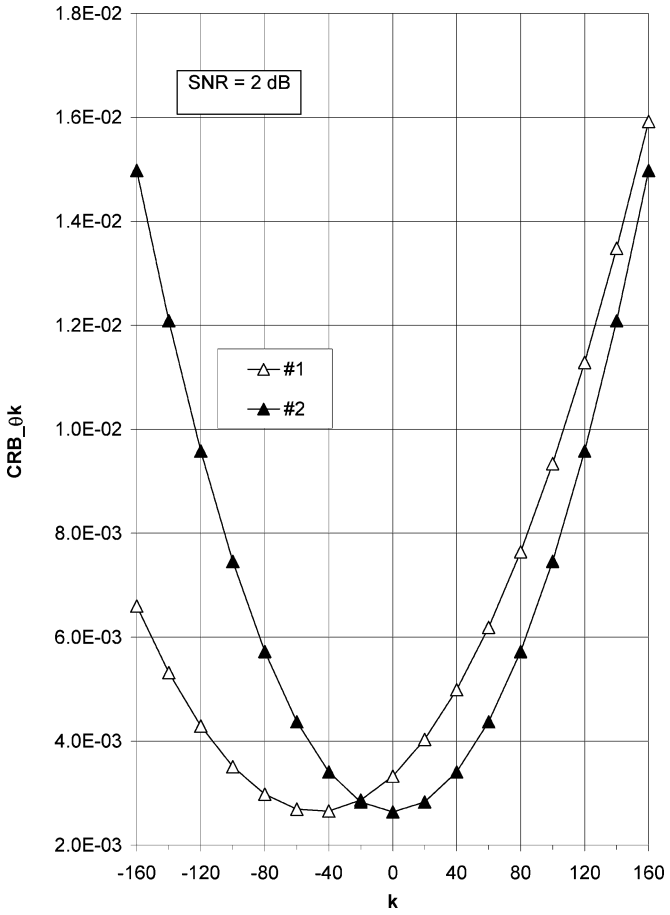


Fig. 4. CRB for the reference phase estimate at $\mathrm{E}_s/\mathrm{N}_0 = 2$ dB.

V. Practical Estimators for Hybrid Burst Structures

In this section, we consider some practical joint carrier phase and frequency estimators for hybrid burst structures.

A. DA Synchronization

The DA estimates are given by [2]

$$\hat{F} = \arg\max_{\widetilde{F}} \left(\left| \sum_{k \in I_{\mathrm{p}}} a_k^* r_k e^{-j2\pi k \widetilde{F} T} \right| \right) \tag{23}$$

$$\hat{\theta} = \arg \left\{ \sum_{k \in I_{\mathrm{p}}} a_k^* r_k e^{-j2\pi k \hat{F} T} \right\}. \tag{24}$$

Of the $\mathrm{L} = \mathrm{N}_d + \mathrm{N}_{\mathrm{p}}$ received samples, only the N_{p} known PSs are used. For $\mathrm{FT} \ll 1$, the estimates $\hat{F}$ and $\hat{\theta}$ from (23) and (24) are unbiased [2]. The mean square estimated errors (MSEEs) of (23) and (24) are lower bounded by the $\mathrm{CRB}_{\mathrm{PS}}(\mathrm{N}_{\mathrm{p}})$ (with $\mathrm{CRB}_{\mathrm{PS}}(\mathrm{N}_{\mathrm{p}}) \geq \mathrm{CRB}_{\mathrm{PS-DS}}(\mathrm{N}_{\mathrm{p}}, \mathrm{N}_d)$). This implies that DA synchronization is intrinsically suboptimal, especially at high SNR where $\mathrm{CRB}_{\mathrm{PS}}(\mathrm{N}_{\mathrm{p}}) \gg \mathrm{CRB}_{\mathrm{PS-DS}}(\mathrm{N}_{\mathrm{p}}, \mathrm{N}_d)$ for $\mathrm{N}_d \gg 1$. Still, it is interesting to understand the behavior of the DA algorithm since multistage synchronization procedures are often initialized with a DA estimate (see Section V-B and -C).

The DA estimates $\hat{F}$ and $\hat{\theta}_{k_G} = \hat{\theta} + 2\pi k_G \hat{F} T$, resulting from (23) and (24), are not affected by a time-shift of the pilot sequence within the burst. This implies that, for a given value of N_{p} and s, the corresponding MSEEs are the same for burst structures #1 and #2. It is well known that, at high SNR, the MSEE reaches the $\mathrm{CRB}_{\mathrm{PS}}(\mathrm{N}_{\mathrm{p}})$. Considering the burst structures from Fig. 1, $\mathrm{CRB}_{\mathrm{PS}}^{\theta_{k_G}}(N_{\mathrm{p}})$ and $\mathrm{CRB}_{\mathrm{PS}}^{FT}(N_{\mathrm{p}})$ are given by (21) and (22). Increasing the number of PSs (N_{p}) decreases $\mathrm{CRB}_{\mathrm{PS}}^{\theta_{k_G}}(N_{\mathrm{p}})$, thus improving the performance of the DA reference phase estimate at $\mathrm{k} = \mathrm{k}_G$ at high SNR. Increasing the number of PSs (N_{p}) and/or the spacing (s) decreases $\mathrm{CRB}_{\mathrm{PS}}^{FT}(N_{\mathrm{p}})$, thus improving the performance of the DA frequency estimate at high SNR. However, below a certain SNR threshold, the performance dramatically degrades across a narrow SNR interval, resulting in an MSEE much larger than the $\mathrm{CRB}_{\mathrm{PS}}(\mathrm{N}_{\mathrm{p}})$. This so-called *threshold phenomenon* results from the occurrence of estimates with large errors, i.e., outlier estimates [2]. The presence of important secondary peaks in the likelihood function results in a large probability of generating outlier frequency estimates at lower SNR, because these secondary peaks can more easily exceed the central peak when noise is added. The SNR threshold decreases with the number of available signal samples N_{p}. For N_{p} consecutive PSs (as in burst structure #1), the threshold is very low so that the DA estimator usually operates above threshold. However, the SNR threshold tends to increase as the PSs are separated by DSs [3], [10].

The simulation results, reported in Fig. 5, illustrate this behavior and show that $\mathrm{s} \approx \mathrm{N}_{\mathrm{p}}/2$ provides a good compromise between a small value of $\mathrm{CRB}_{\mathrm{PS}}^{FT}(N_{\mathrm{p}})$ and a low SNR threshold. A minimum of 10^5 trials have been run to ensure accuracy. Each trial a new phase θ and frequency offset FT is taken from a random uniform distribution over $[-\pi, \pi]$ and $[-0.1, 0.1]$, respectively. The estimated reference phase error was measured modulo 2π, i.e., in the interval $[-\pi, \pi]$.

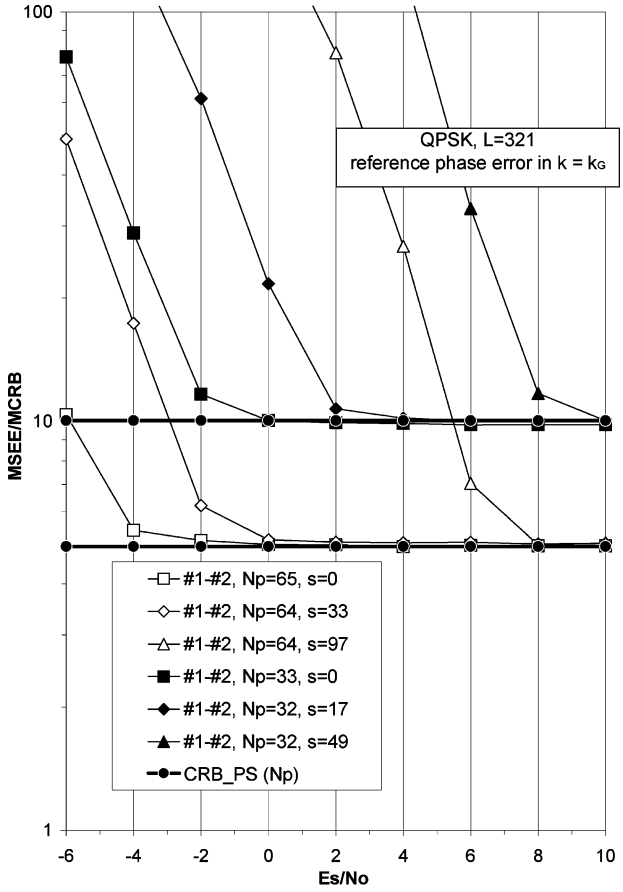
B. NDA Synchronization

Assuming a QPSK constellation, the NDA estimates are given by [2], [11]

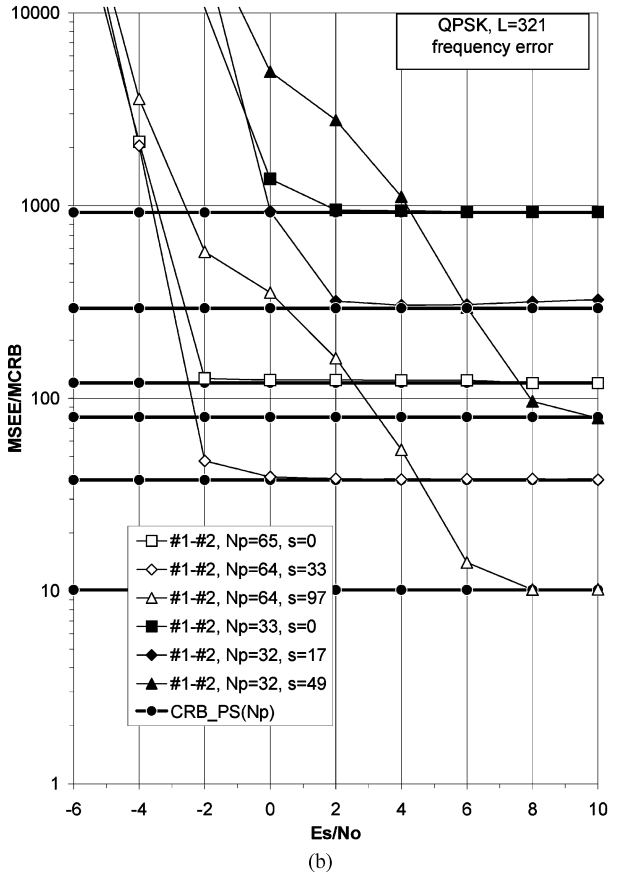
$$\hat{F} = \frac{1}{M} \arg\max_{\widetilde{F}} \left| \sum_{k \in I} |r_k|^2 e^{j4\arg\{r_k\}} e^{-j2\pi k \widetilde{F} T} \right| \tag{25}$$

$$\hat{\theta} = \frac{1}{M} \arg \left\{ \sum_{k \in I} |r_k|^2 e^{j4\arg\{r_k\}} e^{-j8\pi k \hat{F} T} \right\}. \tag{26}$$

All (L) received samples are taken into account, but the knowledge of the PSs is disregarded. For $\mathrm{FT} \ll 1$ and $\theta \in [-\pi/4, \pi/4]$, the estimates $\hat{F}$ and $\hat{\theta}$ from (25) and (26) are unbiased [2]. The resulting MSEE converges to $\mathrm{CRB}_{\mathrm{DS}}(\mathrm{L})$ at high SNR. Simulation results indicate, however, that the value of SNR at which the MSEE becomes close to the $\mathrm{CRB}_{\mathrm{DS}}(\mathrm{L})$ may be quite large. The SNR threshold for the NDA estimator is much higher than for the DA estimator, as the nonlinearity increases the noise level. To cope with this problem, a two-stage coarse–fine DA-NDA estimator has been proposed in [10]. An ML DA estimator is used to coarsely locate the frequency offset, and then the more accurate NDA estimator attempts to improve the estimate within the uncertainty of the coarse estimator. In fact, the search range of the NDA estimator is restricted to the neighborhood of the peak of the DA based likelihood function. This excludes a large percentage of secondary peaks from the search range of the NDA estimator, and thus considerably



(a)



(b)

Fig. 5. The ratio MSEE/MCRB for DA synchronization.

reduces the probability to estimate an outlier frequency. Assuming the MSEE of the initial DA estimate equals the $\mathrm{CRB}_{\mathrm{PS}}^{FT}(N_{\mathrm{p}})$, this uncertainty range can be determined as $\pm m\sqrt{\mathrm{CRB}_{\mathrm{PS}}^{FT}(N_{\mathrm{p}})}$, where m should be carefully chosen. When the parameter m increases, the search region increases, as well as the probability of comprising outlier peaks, which may result in a degradation of the performance *at low SNR* (outlier effect). However, if m decreases, the search region decreases, as well as the probability of comprising the (correct) central peak, which in turn may result in a degradation of the performance *at high SNR*. After frequency and phase correction, the samples for $\mathrm{k} \in \mathrm{I_p}$ are compared to the original PSs and, if necessary, an extra multiple of $\pi/2$ is compensated for.

A major disadvantage of this DA-NDA algorithm is that it does not exploit the knowledge of the PSs in the NDA fine estimation step. Therefore, its MSEE is lower bounded by the $\mathrm{CRB_{DS}}(\mathrm{N_p} + \mathrm{N}_d)$ (with $\mathrm{CRB_{DS}}(\mathrm{N_p} + \mathrm{N}_d) \geq \mathrm{CRB_{PS-DS}}(\mathrm{N_p}, \mathrm{N}_d)$). This implies that the DA-NDA algorithm is intrinsically suboptimal (especially at low and intermediate SNR) in the sense that under no circumstances its performance may meet the $\mathrm{CRB_{PS-DS}}$. Some other estimator may yield an MSEE between $\mathrm{CRB_{DS}}$ and $\mathrm{CRB_{PS-DS}}$, but it should fully exploit the knowledge of the PSs.

C. Iterative DD Synchronization

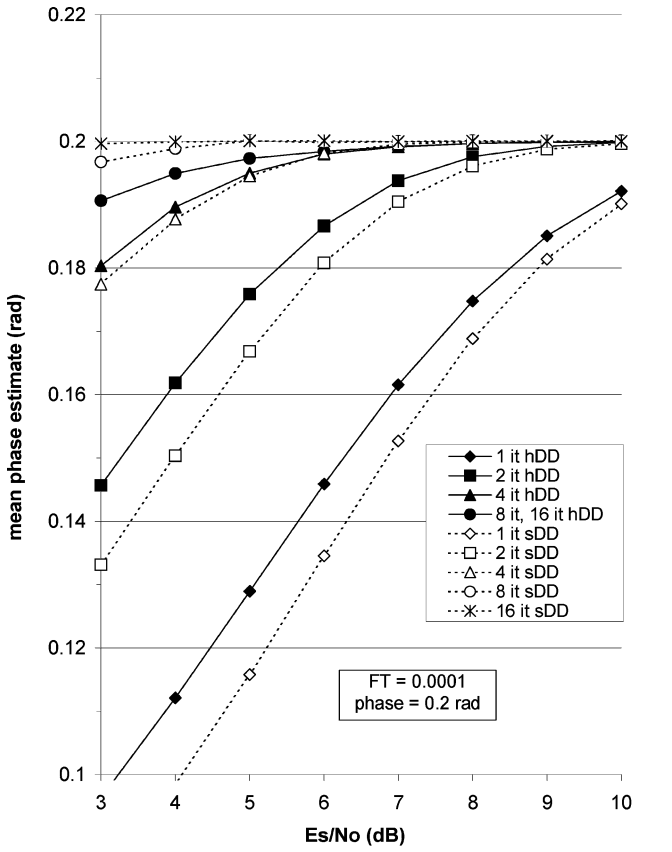
DD estimators extend the sum over $\mathrm{I_p}$ in (23) and (24) with terms over I_d in which the quantities a_k are replaced by hard (hDD) or soft (sDD) decisions that are based upon a previous estimate of (θ, F). For QPSK $\{1, \mathrm{j}, -1, -\mathrm{j}\}$, the soft decisions are given by [4]

$$\hat{a}_k = \frac{\sinh\left[\frac{2E_{\mathrm{s}}}{N_0}\mathrm{Re}(\tilde{r}_k^{(n-1)})\right] + j\sinh\left[\frac{2E_{\mathrm{s}}}{N_0}\mathrm{Im}(\tilde{r}_k^{(n-1)})\right]}{\cosh\left[\frac{2E_{\mathrm{s}}}{N_0}\mathrm{Re}(\tilde{r}_k^{(n-1)})\right] + \cosh\left[\frac{2E_{\mathrm{s}}}{N_0}\mathrm{Im}(\tilde{r}_k^{(n-1)})\right]}. \tag{27}$$

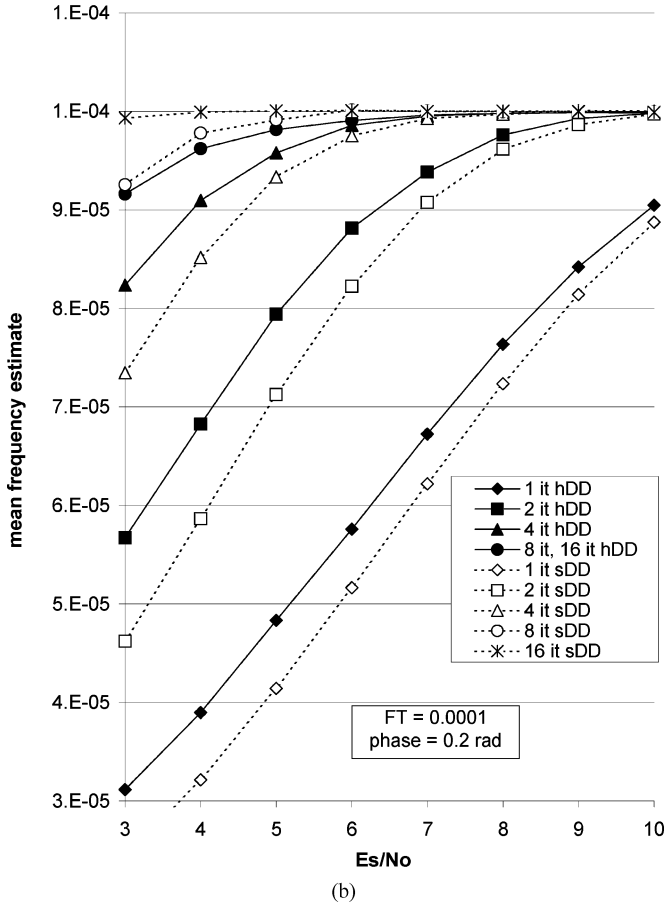
In (27), $\tilde{r}_k^{(n-1)} = r_k e^{-j(\hat{\theta}^{(n-1)} + 2\pi k \hat{F}^{(n-1)} T)}$. The hard decisions $\hat{a}_k$ are determined as the constellation points closest to $\tilde{r}_k^{(n-1)}$. The normal operating SNR of the DD estimators is situated above threshold (large number of available samples (L), no noise enhancement). The required initial estimate $(\hat{\theta}^{(0)}, \hat{F}^{(0)})$ can be obtained from the NDA method; however, the performance below the NDA threshold rapidly degrades, because of an inaccurate initial estimate. If PSs are available, it is better to use DA initialization. We will further refer to these schemes as DA-hDD and DA-sDD. After phase and frequency correction, the samples for $\mathrm{k} \in \mathrm{I_p}$ are compared to the original PSs and, if necessary, an extra multiple of $\pi/2$ is compensated for.

The DD estimators take advantage of both the $(\mathrm{N_p})$ PSs and the (N_d) DSs. At high SNR, the DD estimates $\hat{F}$ and $\hat{\theta}$ are unbiased and their MSEE equals the $\mathrm{CRB_{PS-DS}}(\mathrm{N_p}, \mathrm{N}_d)$. However, the DD estimates become biased at low SNR, implying that the CRB is no longer a valid lower bound on the estimators' performance in this region.

Fig. 6 illustrates this behavior. The presented results are for burst structure #2 with $\mathrm{N_p} = 64$, $\mathrm{s} = 33$, and $\mathrm{L} = 321$. Note that F and θ are decoupled (with $\theta_{k_G} = \theta$). The initial phase error θ was set to 0.2 rad and the initial frequency offset FT was set to 10^{-4}. The mean estimate is plotted versus the SNR. We observe that increasing the number of DD iterations enlarges the



(a)



(b)

Fig. 6. Mean estimate after n iterations of the sDD and hDD algorithms.

SNR range for which the DD estimates are unbiased. The hDD algorithm converges somewhat faster to an unbiased estimate at intermediate SNR, but only the sDD algorithm yields unbiased estimates up to values of $\mathrm{E}_s/\mathrm{N}_0$ as low as 3 dB.

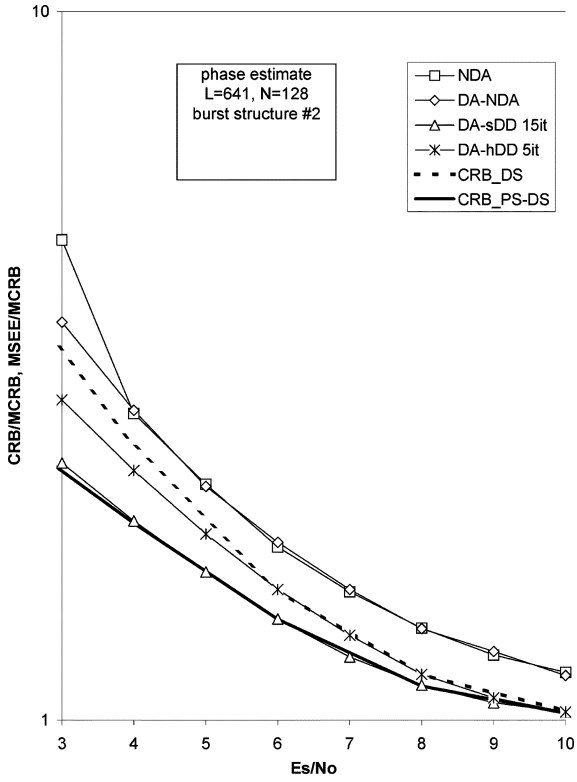
It can be easily shown that the sDD algorithm proposed in [4] for carrier phase estimation and extended here to joint carrier phase and frequency estimation involves a practical implementation of the ML estimator by means of the expectation-maximization (EM) algorithm. This algorithm converges iteratively to the ML estimate provided that the initial estimate is sufficiently accurate [12].

VI. Comparing Performance With True $\mathrm{CRB}_{\mathrm{PS-DS}}$

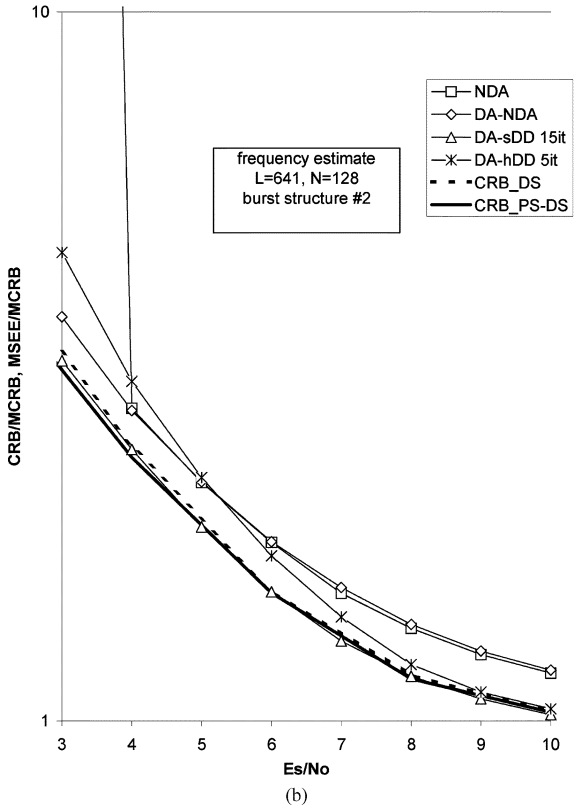
The true ML estimator is known to be asymptotically optimal in the sense that it achieves the performance predicted by the CRB for large data records. However, the performance for finite signal durations cannot be determined analytically. In this section we compare the simulated MSEE of the different estimators listed in Section V to the $\mathrm{CRB}_{\mathrm{PS-DS}}(\mathrm{N_p}, \mathrm{N}_d)$ derived in Section III.

Numerical results pertaining to the different algorithms are obtained in Fig. 7. We assume a burst with $\mathrm{L} = 641$ QPSK symbols, $\mathrm{N_p} = 128$ and $\mathrm{s} = 65$. The PSs are organized as in burst structure #2 from Fig. 1. Note that F and θ are decoupled (with $\theta_{k_G} = \theta$. A minimum of 10^5 simulations has been run to ensure accuracy. Each simulation a different phase and frequency offset was randomly generated from $[-\pi, \pi]$ and $[-0.1, 0.1]$, respectively. We have plotted the ratio MSEE/MCRB for the estimation of F and θ as a function of the SNR. The phase error is measured modulo 2π and supported in the interval $[-\pi, \pi]$, except for the NDA estimator. The phase error of the NDA estimator was estimated modulo $\pi/2$, i.e., in the interval $[-\pi/4, \pi/4]$, as this estimator gives a four-fold phase ambiguity. For the DA-NDA estimation, we chose $\mathrm{m} = 3$. The performance of the DA estimator is not displayed: as $\mathrm{N_p} \ll \mathrm{L}$, the MSEE of the DA estimates is much larger than the MSEE resulting from the other estimators. The MSEE resulting from the DA-hDD estimator reaches a steady state after about five iterations. The MSEE resulting from the DA-sDD estimator reaches a steady state after 10 to 15 iterations. We note that using a combined DA-NDA initialization instead of a DA initialization, the same steady-state performance as for DA initialization is obtained after considerably less (no more than five) iterations, which indicates the importance of an accurate initial estimate to speed up convergence. Further, our results show the following.

1) Above its SNR threshold (at about 4 dB), the NDA estimator performs more or less closely to the $\mathrm{CRB}_{\mathrm{DS}}(\mathrm{L})$.
2) At high SNR, the performance of the DA-NDA estimator matches that of the NDA estimator, but the performance below the SNR threshold degrades less rapidly and is still adequate for reliable receiver operation.
3) At low SNR, the DA-hDD estimator performs worse than the DA-NDA estimator [Fig. 7(b) for frequency estimation]. At (very) low SNR the initial DA estimates become even more accurate than the steady-state hDD estimates. Hence, hard decisions are not useful at low SNR.



(a)



(b)

Fig. 7. MSEE performance of some algorithms using PSs and DSs. Burst structure #2, $\mathrm{L} = 641$, $\mathrm{N_p} = 128$, $\mathrm{s} = 65$.

4) The DA-sDD estimator outperforms by far the DA-hDD estimator, especially at low SNR, and provides a considerable improvement over the DA-NDA estimator. For large L, the DA-sDD estimator performs very closely to the $\mathrm{CRB_{PS-DS}}$.

VII. Conclusion

In this paper, we have investigated the joint phase and frequency estimation from the observation of a "hybrid burst" that contains PSs as well as DSs. We have compared the $\mathrm{CRB_{PS-DS}}$ with the performance of existing carrier synchronizers. Numerical evaluation of the CRB shows how much can be gained in estimator performance by using a "hybrid algorithm" that exploits both PSs and DSs (in some intelligent way), rather than an algorithm that only uses the PSs (and ignores the DSs) or an algorithm that uses all received symbols but ignores the a priori knowledge about the PSs. We have pointed out that the hybrid DA-NDA estimator proposed in [10] is suboptimal because it does not fully exploit the knowledge about the PSs. Further, we have shown that the iterative sDD estimator with DA initialization outperforms the DA-NDA estimator and operates very closely to the $\mathrm{CRB_{PS-DS}}$.

References

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

[2] D. C. Rife and R. R. Boorstyn, "Single-tone parameter estimation from discrete-time observations," *IEEE Trans. Inform. Theory*, vol. IT-20, no. 5, pp. 591–597, Sep. 1974.

[3] J. A. Gansman, J. V. Krogmeier, and M. P. Fitz, "Single frequency estimation with nonuniform sampling," in *Proc. 13th Asilomar Conf. Signals, Systems Computers*, Pacific Grove, CA, Nov. 1996, pp. 878–882.

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

[5] 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, Sep. 2001.

[6] N. Noels, H. Steendam, and M. Moeneclaey, "The true Cramer-Rao bound for phase-independent carrier frequency estimation from a PSK signal," in *Proc. IEEE Globecom 2002*, Taipei, Taiwan, R.O.C., Nov. 2002, CTS-04-2.

[7] 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.

[8] 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.

[9] 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.

[10] B. Beahan, "Frequency estimation of partitioned reference symbol sequences," master's degree , Univ. of South Australia, 2001.

[11] A. J. Viterbi and A. M. Viterbi, "Nonlinear estimation of PSK-modulated carrier phase with application to burst digital transmission," *IEEE Trans. Inform. Theory*, vol. IT-29, pp. 543–551, Jul. 1983.

[12] R. A. Boyles, "On the convergence of the EM algorithm," *J. Roy. Statist. Soc. B*, vol. 45, no. 1, pp. 47–50, 1983.

N. Noels (S'03) received the diploma in electrical engineering from Ghent University, Gent, Belgium, in 2001, where she is currently pursuing the Ph.D. degree at the Department of Telecommunications and Information Processing.

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

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

She is 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.

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

In 1978 to 1999, he held various positions at Ghent University for the Belgian National Fund for Scientific Research (NFWO), from Research Assistant to Research Director. He is presently a Professor in the Department of Telecommunications and Information Processing (TELIN). His research interests include statistical communication theory, carrier and symbol synchronization, bandwidth-efficient modulation and coding, and spread-spectrum, satellite, and mobile communication. He is the author of more than 200 scientific papers in international journals and conference proceedings. He is coauthor (with H. Meyr and S. Fechtel of *Digital Communication Receivers—Synchronization, Channel Estimation, and Signal Processing* (New York: Wiley, 1998). He has been active in various international conferences as Technical Program Committee Member and Session Chairman.

From 1992 to 1994, he was Editor for Synchronization for the IEEE TRANSACTIONS ON COMMUNICATIONS. He was Co-Guest Editor for the December 2001 IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS Special Issue on Signal Synchronization in Digital Transmission Systems. From 1993 to 2002, he was an Executive Committee Member of the IEEE Communications and Vehicular Technology Society Joint Chapter, Benelux Section.

Herwig Bruneel was born in Zottegem, Belgium, in 1954. He received the M.S. degree in electrical engineering, the Licentiate degree in Computer Science, and the Ph.D. degree in computer science from Ghent University, Gent, Belgium, in 1978, 1979, and 1984, respectively.

He is a full Professor in the Faculty of Applied Sciences and Head of the Department of Telecommunications and Information Processing at the same university. He also leads the SMACS Research Group within this department. His main personal research interests include stochastic modeling and analysis of communication systems, discrete-time queueing theory, and the study of ARQ protocols. He has published more than 200 papers on these subjects and is coauthor (with H. Bruneel and B. G. Kim) *Discrete-Time Models for Communication Systems Including ATM* (Boston, MA: Kluwer Academic, 1993). From October 2001 to September 2003, he was Academic Director for Research Affairs at Ghent University.