

## Research Article

# Feedforward Data-Aided Phase Noise Estimation from a DCT Basis Expansion

Jabran Bhatti and Marc Moeneclaey

*Department of Telecommunications and Information Processing, Ghent University, 9000 Ghent, Belgium*

Correspondence should be addressed to Marc Moeneclaey, marc.moeneclaey@telin.ugent.be

Received 1 July 2008; Revised 5 November 2008; Accepted 25 December 2008

Recommended by Erchin Serpedin

This contribution deals with phase noise estimation from pilot symbols. The phase noise process is approximated by an expansion of discrete cosine transform (DCT) basis functions containing only a few terms. We propose a feedforward algorithm that estimates the DCT coefficients without requiring detailed knowledge about the phase noise statistics. We demonstrate that the resulting (linearized) mean-square phase estimation error consists of two contributions: a contribution from the additive noise, that equals the Cramer-Rao lower bound, and a noise independent contribution, that results from the phase noise modeling error. We investigate the effect of the symbol sequence length, the pilot symbol positions, the number of pilot symbols, and the number of estimated DCT coefficients on the estimation accuracy and on the corresponding bit error rate (BER). We propose a pilot symbol configuration allowing to estimate any number of DCT coefficients not exceeding the number of pilot symbols, providing a considerable performance improvement as compared to other pilot symbol configurations. For large block sizes, the DCT-based estimation algorithm substantially outperforms algorithms that estimate only the time-average or the linear trend of the carrier phase.

Copyright © 2009 J. Bhatti and M. Moeneclaey. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

## 1. Introduction

Phase noise refers to random perturbations in the carrier phase, caused by imperfections in both transmitter and receiver oscillators. Compensation of this phase noise is critical since these disturbances can considerably degrade the system performance. The phase noise process typically has a low-pass spectrum [1]. A description of the characteristics of oscillator phase noise is given in [2]. Discrete-time processes that have a bandwidth which is considerably less than the sampling frequency can often be modeled as an expansion of suitable basis functions, that contains only a few terms. Such a basis expansion has been successfully applied in the context of channel estimation and equalization in wireless communications, where the coefficients of the channel impulse response are low-pass processes with a bandwidth that is limited by the Doppler frequency [3–5]. Several methods trying to tackle the phase noise problem exist.

(i) Designing oscillators operating at low-phase noise reduces the need of accurate phase noise compensation algorithms. This, however, leads to expensive oscillators which are difficult to integrate on chip [6–8].

(ii) Phase noise can be tracked by means of a feedback algorithm that operates according to the principle of the phase-locked loop (PLL). As feedback algorithms give rise to rather long acquisition transients, they are not well suited to burst transmission systems [9, 10].

(iii) The observation interval is divided into subintervals and a feedforward algorithm is used to estimate within each subinterval the local time-average (or the linear trend) of the phase [9–11]. This corresponds to approximating the phase noise by a function that is constant (or linear) within each subinterval. Such algorithms avoid the long acquisition transients encountered with feedback algorithms. However, in order that the piecewise constant (or linear) approximation of the phase noise be accurate, the subintervals should

be short, in which case a high sensitivity to additive noise occurs.

(iv) Recently, iterative joint estimation and decoding/detection algorithms have been proposed that make use of the a priori statistics of the phase noise process. A factor graph approach for the estimation of the Markov-type phase noise has been presented in [12, 13], while in [14, 15] sequential Monte Carlo methods combined with Kalman filtering are used to perform detection in the presence of phase noise. These algorithms are computationally rather complex, prevent the use of off-the-shelf decoders, and assume detailed knowledge about the phase noise statistics at the receiver. Less complex iterative phase noise estimation algorithms based on Wiener filtering have been presented in [16], but still require knowledge about the phase noise autocorrelation function at the receiver.

In this contribution, we apply the basis expansion model to the problem of phase noise estimation from pilot symbols only, using the orthogonal basis functions from the discrete cosine transform (DCT). In contrast to the case of channel estimation, the phase noise does not enter the observation model in a linear way. Section 2 presents the system description which includes the observation model and a general phase noise model. Also, the phase noise estimation algorithm, based on the estimation of only a few DCT coefficients, is derived. Section 3 contains the performance analysis of the proposed algorithm in terms of the mean-square error (MSE) of the phase estimate. The behavior of the linearized model in the frequency domain is examined in Section 4. Analysis results are confirmed by computer simulations in Section 5, which consider both the mean-square phase estimation error and the associated bit error rate (BER) degradation. Section 6 gives a complexity analysis of our algorithm. Conclusions are drawn in Section 7.

## 2. System Description

We consider the transmission of a block of  $K$  data symbols over an AWGN channel that is affected by phase noise. The resulting received signal is represented as

$$r(k) = a(k)e^{j\theta(k)} + w(k) \quad \text{for } k = 0, \dots, K-1, \quad (1)$$

where the index  $k$  refers to the  $k$ th symbol interval of length  $T$ ,  $\{a(k)\}$  is a sequence of data symbols with symbol energy  $E[|a(k)|^2] = E_s$ , the additive noise  $\{w(k)\}$  is a sequence of i.i.d. zero-mean circularly symmetric complex-valued Gaussian random variables with  $E[|w(k)|^2] = N_0$ , and  $\theta(k)$  is a time-varying phase noise process with  $K \times K$  correlation matrix  $\mathbf{R}_\theta$ . The symbol sequence  $\{a(k)\}$  contains  $K_P$  known pilot symbols at positions  $k_i, i = 0, \dots, K_P - 1$ , with constant magnitude  $|a(k_i)|^2 = E_s$ . From the observation of the received signal at the pilot symbol positions  $k_i$ , an estimate  $\hat{\theta}(k)$  of the time-varying phase  $\theta(k)$  is to be produced. This phase estimate will be used to rotate the received signal before data detection, that is, the detection of the data symbols is based on  $\{z(k)\} = \{r(k)\exp(-j\hat{\theta}(k))\}$ . The detector is designed under the assumption of perfect carrier synchronization, that is,  $\hat{\theta}(k) = \theta(k)$ . For uncoded

transmission, the detection algorithm reduces to symbol-by-symbol detection:

$$\hat{a}(k) = \arg \min_{a \in A} |z(k) - a|^2, \quad k \notin \{k_i, i = 0, \dots, K_P - 1\} \quad (2)$$

with  $A$  denoting the symbol constellation. The phase  $\theta(k)$  can be represented as a weighed sum of  $K$  basis functions over the interval  $[0, K-1]$ :

$$\theta(k) = \sum_{n=0}^{K-1} x_n \psi_n(k), \quad k = 0, \dots, K-1. \quad (3)$$

As  $\theta(k)$  is essentially a low-pass process, it can be well approximated by the weighed sum of a *limited number*  $N (\ll K)$  of suitable basis functions:

$$\theta(k) \approx \sum_{n=0}^{N-1} x_n \psi_n(k), \quad k = 0, \dots, K-1. \quad (4)$$

In this contribution, we make use of the orthonormal discrete cosine transform (DCT) basis functions, that are defined as

$$\psi_n(k) = \begin{cases} \sqrt{\frac{1}{K}}, & n = 0, \\ \sqrt{\frac{2}{K}} \cos\left(\frac{\pi n}{K} \left(k + \frac{1}{2}\right)\right), & n > 0. \end{cases} \quad (5)$$

Hence, from (3),  $x_n$  is the  $n$ th DCT coefficient of  $\theta(k)$ . As  $\psi_n(k)$  has its energy concentrated near the frequencies  $n/2KT$  and  $-n/2KT$ , the DCT basis functions are well suited to represent a low-pass process by means of a small number of basis functions.

In the following, we produce from the observation  $\{r(k_i)\}$  at the pilot symbol positions  $k_i$ , with  $i = 0, \dots, K_P - 1$ , an estimate  $\hat{x}_n$  of the coefficients  $x_n$ , with  $n = 0, \dots, N-1$ , using the phase model (4) with equality. The final estimate  $\hat{\theta}(k)$  is obtained by computing the inverse DCT of  $\{\hat{x}_n\}$ :

$$\hat{\theta}(k) = \sum_{n=0}^{N-1} \hat{x}_n \psi_n(k) \quad \text{for } k = 0, \dots, K-1. \quad (6)$$

However, as (4) is not an exact model of the true phase  $\theta(k)$ , the phase estimate is affected not only by the additive noise contained in the observation, but also by a phase noise modeling error. Considering the observations (1) at instants  $k_i$ , and assuming that (4) holds with equality, we obtain

$$\mathbf{r}_P = \mathbf{D}(\mathbf{x})\mathbf{a}_P + \mathbf{w}_P, \quad (7)$$

where for  $i = 0, \dots, K_P - 1$ ;  $(\mathbf{r}_P)_i = r(k_i)$ ,  $(\mathbf{w}_P)_i = w(k_i)$ ,  $(\mathbf{a}_P)_i = a(k_i)$ , and  $\mathbf{D}(\mathbf{x})$  is a  $K_P \times K_P$  diagonal matrix with

$$(\mathbf{D}(\mathbf{x}))_i = e^{j(\Psi_P \mathbf{x})_i} \quad (8)$$

and  $(\Psi_P)_{i,n} = \psi_n(k_i)$ ,  $(\mathbf{x})_n = x_n$ ,  $n = 0, \dots, N-1$  with  $N \leq K_P$ . The  $K_P \times 1$  vectors  $\mathbf{r}_P$ ,  $\mathbf{a}_P$ , and  $\mathbf{w}_P$  can be viewed as resulting from subsampling  $\{r(k)\}$ ,  $\{a(k)\}$ , and  $\{w(k)\}$  at

the instants  $k_i$  that correspond to the pilot symbol positions. Similarly, the  $n$ th column of the  $K_p \times N$  matrix  $\Psi_P$  is obtained by subsampling the  $n$ th DCT basis function  $\psi_n(k)$ . Maximum likelihood estimation of  $\mathbf{x}$  from  $\mathbf{r}_p$  results in

$$\hat{\mathbf{x}}_{ML} = \arg \min_{\mathbf{x}} \|\mathbf{r}_p - \mathbf{D}(\mathbf{x})\mathbf{a}_p\|^2. \quad (9)$$

As  $\mathbf{x}$  enters the observation  $\mathbf{r}_p$  in a nonlinear way, the ML estimate is not easily obtained. Therefore, we resort to a suboptimum ad hoc estimation of  $\mathbf{x}$ , which is based on the argument (angle) of the complex-valued observations. However, as the function  $\arg(z)$  reduces the argument of  $z$  to an interval  $[-\pi, \pi]$ , taking  $\arg(r(k_i))$  might give rise to phase wrapping, especially when the time-average of  $\theta(k)$  is close to  $-\pi$  or  $\pi$ . In order to reduce the probability of phase wrapping, we first rotate the observation  $\mathbf{r}$  over an angle  $\theta_{\text{avg}}$  that is close to the time-average of  $\theta(k)$ , then we estimate the DCT coefficients of the fluctuation  $\theta(k) - \theta_{\text{avg}}$  and finally we compute the phase estimate  $\hat{\theta}(k)$ . We select

$$\theta_{\text{avg}} = \arg \left( \sum_{i=0}^{K_p-1} r(k_i) \right) \quad (10)$$

and construct  $\mathbf{r}'$  with

$$\begin{aligned} (\mathbf{r}')_i &= r'(k_i) \\ &= \arg(r(k_i)a^*(k_i)\exp(-j\theta_{\text{avg}})) \\ &\quad \text{for } i = 0, \dots, K_p - 1. \end{aligned} \quad (11)$$

We obtain an estimate  $\hat{\mathbf{x}}'$  of the DCT coefficients of the fluctuation  $\theta(k) - \theta_{\text{avg}}$  through a least-squares fit  $\hat{\mathbf{x}}' = \arg \min_{\mathbf{x}} \|\mathbf{r}' - \Psi_P \mathbf{x}\|^2$ , yielding

$$\hat{\mathbf{x}}' = (\Psi_P^T \Psi_P)^{-1} \Psi_P^T \mathbf{r}'. \quad (12)$$

In order that  $(\Psi_P^T \Psi_P)^{-1}$  exists, we need  $N \leq K_p$ . Finally, the phase estimate is given by

$$\begin{aligned} \hat{\boldsymbol{\theta}} &= \theta_{\text{avg}} \mathbf{1}_K + \Psi_K \hat{\mathbf{x}}' \\ &= \theta_{\text{avg}} \mathbf{1}_K + \mathbf{M} \mathbf{r}', \end{aligned} \quad (13)$$

where  $\mathbf{M} = \Psi_K (\Psi_P^T \Psi_P)^{-1} \Psi_P^T$  and  $(\hat{\boldsymbol{\theta}})_k = \hat{\theta}(k)$ ,  $(\mathbf{1}_K)_k = 1$ ,  $(\Psi_K)_{k,n} = \psi_n(k)$ ,  $k = 0, \dots, K - 1$ ;  $n = 0, \dots, N - 1$ . Note from (13) that the estimation algorithm does not need specific knowledge about the phase noise process. As  $r'(k_i)$  from (11) can be viewed as a noisy version of  $\theta(k_i) - \theta_{\text{avg}}$ , the phase estimate  $\hat{\theta}$  from (13), or, equivalently, the phase estimate  $\hat{\boldsymbol{\theta}}(k)$  from (6), can be interpreted as an interpolated version of the subsampled noisy phase trajectory. The estimation of the phase trajectory involves the inversion of the  $N \times N$  matrix  $\Psi_P^T \Psi_P$ , which depends on the pilot symbol positions  $\{k_i, i = 0, \dots, K_p - 1\}$ . Now, we point out that the pilot symbol positions can be selected such that  $\Psi_P^T \Psi_P$  is diagonal, or, equivalently, that the  $N$  columns of the  $K_p \times N$  matrix  $\Psi_P$  are orthogonal. Such selection of  $\{k_i\}$  avoids the need for matrix inversion in (12). Denoting by  $\phi_n(i)$  the

orthonormal DCT basis functions of length  $K_p$ , it is easily verified that selecting  $\{k_i\}$  such that

$$k_i = \frac{iK}{K_p} + \frac{K - K_p}{2K_p}, \quad i = 0, \dots, K_p - 1 \quad (14)$$

gives rise to

$$\psi_n(k_i) = \sqrt{\frac{K_p}{K}} \phi_n(i) \quad \text{for } n = 0, \dots, K_p - 1, \quad (15)$$

so that

$$\Psi_P^T \Psi_P = \frac{K_p}{K} \mathbf{I}_N \quad (16)$$

with  $\mathbf{I}_N$  denoting the  $N \times N$  identity matrix. Equations (12) and (13) then reduce to

$$\hat{\mathbf{x}}' = \frac{K}{K_p} \Psi_P^T \mathbf{r}', \quad (17)$$

$$\hat{\boldsymbol{\theta}} = \theta_{\text{avg}} \mathbf{1}_K + \frac{K}{K_p} \Psi_K \Psi_P^T \mathbf{r}'. \quad (18)$$

In order that all  $k_i$  from (14) be integer,  $K$  must be an odd multiple of  $K_p$ , that is,  $K = (2d + 1)K_p$ , yielding  $k_i = (2d + 1)i + d$ . The resulting pilot symbol configuration is suited for estimating any number of DCT coefficients not exceeding  $K_p$ . When  $K$  is not an odd multiple of  $K_p$ , rounding the right-hand side of (14) to the nearest integer gives rise to pilot symbol positions that still yield an essentially diagonal matrix  $\Psi_P^T \Psi_P$  in which case the simplified equations (17) and (18) can still be used.

### 3. Performance Analysis

As the observation vector  $\mathbf{r}_p$  is a nonlinear function of the carrier phase, an exact analytical performance analysis is not feasible. Instead, we will resort to a linearization of the argument function in (11) in order to obtain tractable results.

Linearization of the argument function yields

$$\begin{aligned} r'(i) &= \arg(r(k_i)a^*(k_i)e^{-j\theta_{\text{avg}}}) \\ &= \arg(e^{j(\theta(k_i) - \theta_{\text{avg}})}(E_s + a^*(k_i)w(k_i)e^{-j\theta(k_i)})) \\ &\approx \theta(k_i) - \theta_{\text{avg}} + n_p(i) \end{aligned} \quad (19)$$

for  $i = 0, \dots, K_p - 1$ , where  $\{n_p(i)\}$  is a sequence of i.i.d. zero-mean Gaussian random variables with variance  $N_0/2E_s$ . Note that (19) incorporates the true phase  $\theta(k_i)$  instead of the approximate model (4), so that our performance analysis will take the modeling error into account. In order that the linearization in (19) be valid, we need  $|\theta(k_i) - \theta_{\text{avg}}| < \pi$  (because  $|\arg(z)| < \pi$ ) and  $|w(k_i)|^2 \ll E_s$ ; hence, the phase noise fluctuations should not cause phase wrapping and  $E_s/N_0$  should be sufficiently large. Substituting (19) into (13) yields

$$\hat{\boldsymbol{\theta}} = \mathbf{M}(\boldsymbol{\theta}_p + \mathbf{n}_p) = \mathbf{M}\boldsymbol{\theta} + \mathbf{M}\mathbf{n}_p, \quad (20)$$

where  $(\mathbf{n}_p)_i = n_p(i)$ ,  $(\boldsymbol{\theta}_p)_i = \theta(k_i)$ , and the  $K_p \times K$  matrix  $\mathbf{S}$  is such that its  $i$ th row has a 1 at the  $k_i$ th column and zeroes elsewhere ( $i = 0, \dots, K_p - 1$ ). The estimation error resulting from (20) is given by

$$\hat{\boldsymbol{\theta}} - \boldsymbol{\theta} = (\mathbf{MS} - \mathbf{I}_K)\boldsymbol{\theta} + \mathbf{M}\mathbf{n}_p, \quad (21)$$

where  $\mathbf{I}_K$  denotes the  $K \times K$  identity matrix. If the model (4) was exact, we would have  $\boldsymbol{\theta} = \boldsymbol{\Psi}_K \mathbf{x}$  and  $\boldsymbol{\theta}_p = \boldsymbol{\Psi}_p \mathbf{x}$ , yielding

$$\hat{\boldsymbol{\theta}} = \boldsymbol{\theta} + \mathbf{M}\mathbf{n}_p, \quad (22)$$

in which case the estimation error would be caused only by the additive noise.

As a performance measure of the estimation algorithm, we consider the mean-square error (MSE), defined as

$$\text{MSE} = \frac{1}{K} E[\text{trace}((\hat{\boldsymbol{\theta}} - \boldsymbol{\theta})(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta})^T)]. \quad (23)$$

Substituting (21) into (23) yields

$$\text{MSE} = \frac{1}{K} \frac{N_0}{2E_s} \text{trace}((\boldsymbol{\Psi}_p^T \boldsymbol{\Psi}_p)^{-1}) + \text{MSE}_\infty, \quad (24)$$

where

$$\text{MSE}_\infty = \frac{1}{K} \text{trace}((\mathbf{MS} - \mathbf{I}_K)\mathbf{R}_\theta(\mathbf{MS} - \mathbf{I}_K)^T). \quad (25)$$

The first term in (24) denotes the contribution from the additive noise, whereas the second term in (24) constitutes an MSE floor, caused by the phase noise modeling error. The phase noise statistics affect the MSE floor through the autocorrelation matrix  $\mathbf{R}_\theta$ . The MSE floor decreases with increasing  $N$  (because the modeling error is reduced when more DCT coefficients are taken into account), whereas the additive noise contribution to the MSE increases with  $N$  (because  $N$  parameters need to be estimated). Hence, there is an optimum value of  $N$  that minimizes the MSE.

From the nonlinear observation model (7), which assumes that (4) holds with equality, we compute the Cramer-Rao lower bound on the MSE (23) resulting from any unbiased estimate  $\hat{\mathbf{x}}$  of the DCT coefficients of  $\theta(k)$ :

$$\text{MSE} \geq \frac{1}{K} \text{trace}(\mathbf{J}^{-1}). \quad (26)$$

In (26),  $\mathbf{J}$  denotes the Fisher information matrix related to the estimation of  $\mathbf{x}$  from (7), which is found to be

$$(\mathbf{J})_{n,n'} = \frac{2E_s}{N_0} ((\boldsymbol{\Psi}_p^T \boldsymbol{\Psi}_p)^{-1})_{n,n'}. \quad (27)$$

Combining (26) with (27) yields the following performance bound:

$$\text{MSE} \geq \frac{1}{K} \frac{N_0}{2E_s} \text{trace}((\boldsymbol{\Psi}_p^T \boldsymbol{\Psi}_p)^{-1}). \quad (28)$$

Comparison of (24) and (28) indicates that our ad hoc algorithm (13) yields the minimum possible (over all unbiased estimates) noise contribution to the MSE (assuming that the linearization of the observation model is valid).

When the pilot symbol positions  $\{k_i\}$  are selected according to (14), the Cramer-Rao bound (28) reduces to

$$\text{MSE} \geq \frac{N_0}{2E_s} \frac{N}{K_p}, \quad (29)$$

which indicates that the sensitivity to additive noise increases with the number ( $N$ ) of estimated DCT coefficients.

## 4. Frequency-Domain Analysis

After linearization, (20) relates the phase estimate  $\hat{\boldsymbol{\theta}}$  to the actual phase  $\boldsymbol{\theta}$  and the additive noise  $\mathbf{n}_p$ . In the absence of additive noise, the estimator can be viewed as a linear system that transforms  $\boldsymbol{\theta}$  into  $\hat{\boldsymbol{\theta}}$  by means of the transfer matrix  $\mathbf{MS}$ . In order to analyze this system in the frequency domain, we consider an input  $\boldsymbol{\theta}_n$  with  $(\boldsymbol{\theta}_n)_k = \exp(j2\pi kn/K)$ , that is,  $\boldsymbol{\theta}_n$  contains only the frequency  $n/K$ . We investigate the mean-square error  $\text{MSE}_n$  between the input  $\boldsymbol{\theta}_n$  and the output  $\hat{\boldsymbol{\theta}} = \mathbf{MS}\boldsymbol{\theta}_n$ ;  $\text{MSE}_n$  is given by (25), with  $\mathbf{R}_\theta$  replaced by  $\boldsymbol{\theta}_n \boldsymbol{\theta}_n^H$ , where the superscript  $H$  indicates conjugate transpose.

As  $\boldsymbol{\theta}_n$  is periodic in  $n$  with period  $K$ , the same periodicity holds for  $\text{MSE}_n$ . Assuming the pilot symbol positions are according to (14) with  $K = 105$  and  $K_p = 15$ , Figure 1 shows  $\text{MSE}_n$  as a function of  $n/K$ , with  $n/K$  in the interval  $[-1/2, 1/2]$  and  $N = 7$ . The behavior of  $\text{MSE}_n$  is explained by noting that subsampling  $\boldsymbol{\theta}_n$  at the instants  $k_i$  (with spacing  $K_p$ ) gives rise to aliasing. Frequencies  $n/K$  and  $(n + K_p)/K$  yield the same subsampled phase trajectory. In the following discussion, the intervals  $I_{K_p}$  and  $I_N$  are defined as  $[-(K_p - 1)/(2K), (K_p - 1)/(2K)]$  and  $[-(N - 1)/(2K), (N - 1)/(2K)]$ , respectively; note that  $I_N \subset I_{K_p}$ .

- (i) As the first  $N$  basis functions of the DCT transform cover the frequency interval  $I_N$ , we get  $\hat{\boldsymbol{\theta}}_n \approx \boldsymbol{\theta}_n$  and  $\text{MSE}_n \approx 0$  when  $n/K$  is in  $I_N$ .
- (ii) When  $n/K$  is in the interval  $I_{K_p}$ , but outside  $I_N$ , we get  $\hat{\boldsymbol{\theta}}_n \approx 0$  and  $\text{MSE}_n \approx 1$ .
- (iii) Suppose  $n = mK_p + n'$ , with  $m \neq 0$ ,  $|m| < K/(2K_p)$  and  $n'$  in  $I_{K_p}$ , because of aliasing,  $\boldsymbol{\theta}_n$  is interpreted as  $\boldsymbol{\theta}_{n'} \exp(j\phi_m)$  with  $\phi_m = 2\pi m(K - K_p)/(2K)$ . When  $n'$  is in the interval  $I_N$ , we get  $\hat{\boldsymbol{\theta}}_n \approx \boldsymbol{\theta}_{n'} \exp(j\phi_m)$ . The resulting estimation error is the sum of two complex exponentials with frequencies  $n/K$  and  $n'/K$ , yielding  $\text{MSE}_n \approx 2$ . When  $n'$  is not in  $I_N$ , we get  $\hat{\boldsymbol{\theta}}_n \approx 0$  and  $\text{MSE}_n \approx 1$ .

It follows from Figure 1 that the estimator can be viewed as a low-pass system with bandwidth  $B = (N - 1)/(2K)$ . Basically, the frequency components  $n/K$  of  $\boldsymbol{\theta}$  with  $|n/K| < B$  are tracked by the estimator, whereas the components with  $|n/K| > B$  contribute to the MSE.

## 5. Simulation Results

In this section, we assess the performance of the proposed technique in terms of the MSE of the phase estimate and the resulting BER degradation by means of computer

simulations. In our simulations, we will consider two types of phase noise, that is, Wiener phase noise and first-order phase noise. The (discrete time) first-order phase noise process  $\theta(k)$  can be viewed as the output of a one-pole filter driven by white Gaussian noise:

$$\theta(k+1) = (1 - \alpha)\theta(k) + \Delta(k), \quad (30)$$

where  $\{\Delta(k)\}$  is a sequence of i.i.d. zero-mean Gaussian random variables with variance  $\sigma_\Delta^2$ . The corresponding phase noise power spectrum and phase noise variance are given by

$$\begin{aligned} S_\theta^{\text{1st-order}}(e^{j2\pi fT}) &= \frac{\sigma_\Delta^2}{|\exp(j2\pi fT) - 1 + \alpha|^2} \\ &\approx \frac{\sigma_\Delta^2}{|j2\pi fT + \alpha|^2}, \end{aligned} \quad (31)$$

$$\sigma_\theta^2 = \frac{\sigma_\Delta^2}{\alpha(2 - \alpha)} \approx \frac{\sigma_\Delta^2}{2\alpha}. \quad (32)$$

The approximations in (31) and (32) hold for  $fT \ll 1/2$  and  $\alpha \ll 1$ . It follows from (31) that  $\alpha/2\pi T$  is the 3 dB frequency of the power spectrum. The first-order phase noise models the phase instabilities of an oscillator signal that results from a phase-locked loop (PLL) circuit. The (discrete-time) Wiener phase noise  $\theta(k)$  is described by the following system equation:

$$\theta(k+1) = \theta(k) + \Delta(k), \quad k = 0, \dots, K-2, \quad (33)$$

where the initial phase noise value  $\theta(0)$  is uniformly distributed in  $[-\pi, \pi]$  and  $\Delta(k)$  has the same meaning as in (30). Hence,  $\theta(k)$  can be viewed as the output of an integrator with a white noise input. From (33), it follows that the variance of the Wiener phase noise increases linearly with the time index  $k$ , which indicates that the process is nonstationary.

Comparing (33) and (30), it follows that the Wiener phase noise can be interpreted as a limiting case of first-order phase noise, in the limit for  $\alpha \rightarrow 0$ . Hence, one can *formally* define the Wiener phase noise spectrum as the limit of the first-order spectrum (31); for  $\alpha \rightarrow 0$ ,

$$S_\theta^{\text{Wiener}}(e^{j2\pi fT}) = \frac{\sigma_\Delta^2}{|\exp(j2\pi fT) - 1|^2} \approx \frac{\sigma_\Delta^2}{4\pi^2 f^2 T^2}, \quad (34)$$

where the approximation in (34) holds for  $|fT| \ll 1/2$ . Note that the Wiener phase noise spectrum becomes unbounded at  $f = 0$ , which is a consequence of the variance increasing linearly with time. In contrast, the complex envelope  $\exp(j\theta)$  of the oscillator signal can be shown to be a stationary process (with [1, the Lorentzian power spectrum]). The Wiener phase noise model is often used to describe the phase noise process of a free-running oscillator, although also more elaborate models exist, involving a phase noise spectrum that consists of a sum of terms of the form  $A_m f^{-m}$ ,  $m = 0, \dots, 4$  [10, 17–19]. In order to reduce the strong low-frequency components of the phase noise resulting from a free-running oscillator, the oscillator is often incorporated in a PLL circuit;

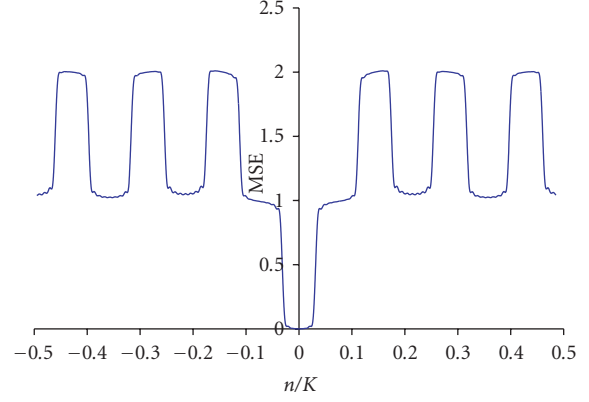


FIGURE 1: MSE as a function of  $n/K$  for  $K = 105$ ,  $K_p = 15$ , and  $N = 7$ .

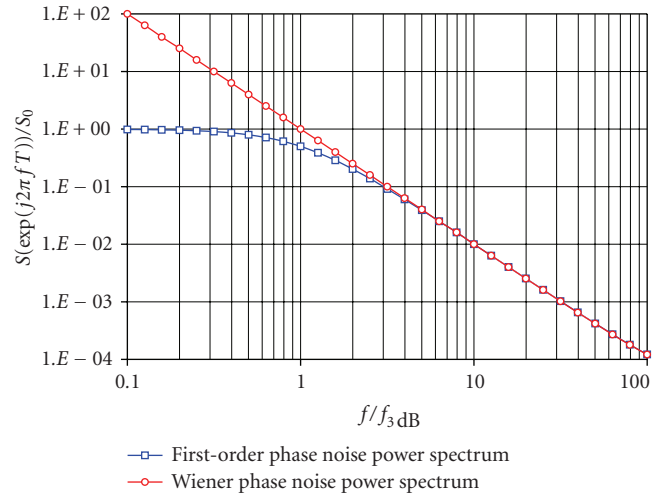


FIGURE 2: Power spectrum for Wiener phase noise and first-order phase noise.

a first-order PLL gives rise to the first-order phase noise process (30) [17].

Figure 2 shows the first-order phase noise power spectrum, normalized by its value  $S_0$  at  $f = 0$ , as a function of the normalized frequency  $f/f_{3\text{dB}}$ , with  $f_{3\text{dB}} = \alpha/(2\pi T)$ ; also displayed is the Wiener phase noise power spectrum (normalized by the same  $S_0$ ). As for both types of phase noise, the same value of  $\sigma_\Delta^2$  has been used, both spectra have the same high-frequency content.

In the following simulations, Wiener phase noise is assumed, unless noted otherwise. First, we assume transmission of a block of length  $K = 105$  symbols, consisting of  $K_D = 90$  uncoded QPSK data symbols and  $K_P = 15$  constant-energy pilot symbols that are inserted into the sequence according to (14).

(i) Figure 3 shows the MSE of the phase estimate in the absence of phase noise as a function of  $E_s/N_0$  when  $N = 1, 4$  and 10 DCT coefficients are estimated; in addition, these simulation results are compared to the corresponding CRB (29). We observe that the CRB is achieved for sufficiently

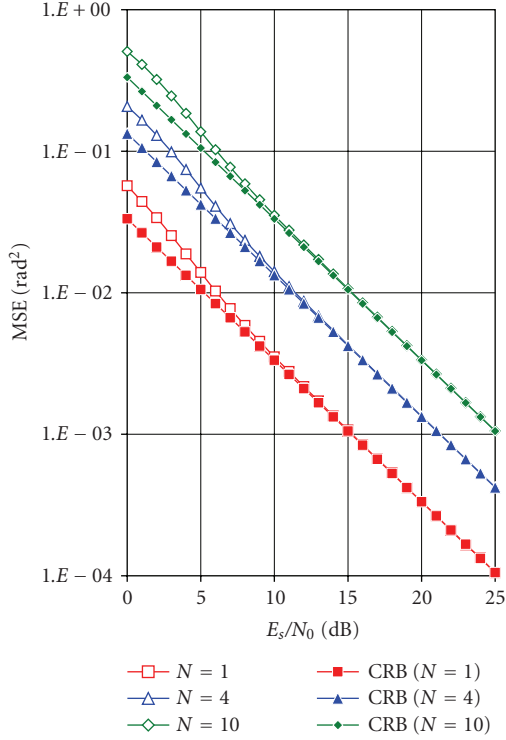


FIGURE 3: MSE in the absence of phase noise compared to the corresponding CRB.  $K = 105$ ,  $K_p = 15$ .

high values of  $E_s/N_0$ . For small  $E_s/N_0$ , the MSE exceeds the CRB, which is in agreement with the fact that the linearized observation model from (19) is no longer accurate in the low-SNR region. Furthermore, it is confirmed that the contribution from the additive noise to the MSE is proportional to the number of estimated coefficients  $N$ .

(ii) Figure 4 shows the MSE as a function of  $E_s/N_0$  for  $N = 1, 4$  and  $10$ , but this time in the presence of Wiener phase noise with  $\sigma_\Delta^2 = 0.0027 \text{ rad}^2$  (which corresponds to “strong” phase noise, with  $\sigma_\Delta = 3^\circ$ ). We observe an MSE floor in the high- $E_s/N_0$  region, which can be reduced by increasing the number  $N$  of estimated coefficients. Figure 4 also confirms that for low  $E_s/N_0$ , the MSE increases when  $N$  increases. This high- $E_s/N_0$  and low- $E_s/N_0$  behaviors indicate that for given  $K$ ,  $K_p$ , and  $E_s/N_0$ , the MSE can be minimized by proper selection of  $N$ .

(iii) Figure 5 shows the bit error rate (BER) as a function of  $E_b/N_0$  ( $E_b$  is the energy per transmitted bit,  $E_s = 2(1-\eta)E_b$  for QPSK) for  $N = 1, 4$ , and  $10$ . The reference BER curve corresponds to a system with perfect synchronization and no pilot symbols ( $\eta = 0$ ). We observe that for low  $E_b/N_0$ , it is sufficient to estimate only the time-average of the phase (i.e.,  $N = 1$ ). Estimating a higher number of DCT coefficients can lead to a worse BER performance for low  $E_b/N_0$  because the MSE of the phase estimate due to additive noise increases with  $N$ . At high  $E_b/N_0$ , a BER floor occurs which decreases with increasing  $N$ , so in this region it becomes beneficial to estimate more than just one DCT coefficient. Hence, the

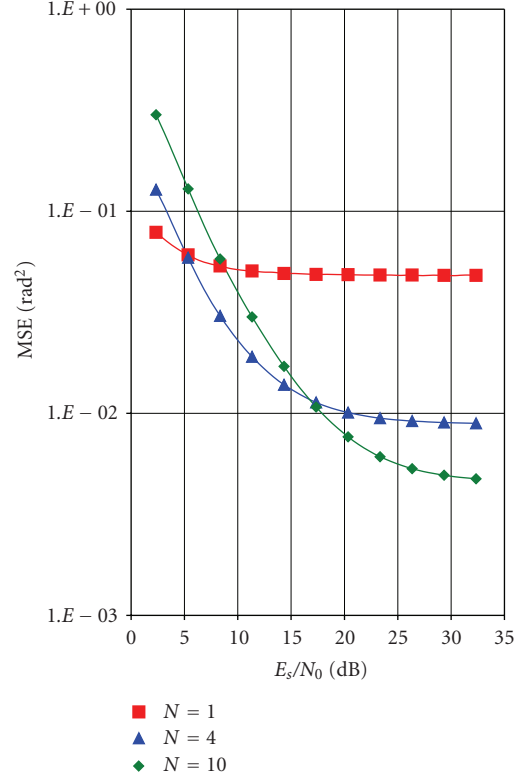


FIGURE 4: MSE when Wiener phase noise with  $\sigma_\Delta = 3^\circ$  is present  $K = 105$ ,  $K_p = 15$ .

optimal number of estimated coefficients  $N_{\text{opt}}$  will depend on the operating  $E_b/N_0$ .

(iv) Figure 6 compares the BER degradations at  $\text{BER}_{\text{ref}} = 10^{-4}$  resulting from Wiener phase noise and first-order phase noise; the value of  $\sigma_\Delta^2$  is the same for both phase noise processes, such that the Wiener phase noise spectrum and first-order phase noise spectrum are the same for large  $f$ . (The BER degradation caused by some impairment is characterized by the increase (in dB) of  $E_b/N_0$  (as compared to the case of no impairment) needed to maintain the BER at a specified reference level.) As the 3 dB frequency  $\alpha/(2\pi T)$  of the first-order phase noise is less than  $BT$ , the frequency contents of the Wiener phase noise and the first-order phase noise outside the estimator bandwidth are essentially the same, and the corresponding BER curves are nearly coincident; this is in agreement with the analysis from Section 4, where we showed that the low-frequency components of the phase noise practically do not contribute to the phase error. It is also confirmed that there is an optimum value of  $N$  that minimizes the BER degradation; this optimum  $N$  increases with  $\sigma_\Delta$ .

Next, we study the influence of the pilot symbol positions in the symbol sequence, assuming Wiener phase noise with  $\sigma_\Delta = 3^\circ$ . The following scenarios are considered (see Figure 7), with  $K_p = 15$ .

- (i) The pilot symbols are inserted according to (14) (SCEN1).

- (ii) All pilot symbols are located in the middle of the sequence (SCEN2).
- (iii)  $\lfloor K_p/2 \rfloor$  pilot symbols are inserted at the beginning of the sequence, the remaining  $\lfloor K_p/2 \rfloor$  pilot symbols are placed at the end (SCEN3).
- (iv) The  $K_p$  pilot symbols are placed equidistantly at positions  $\{0, K/K_p, \dots, (K_p - 1)K/K_p\}$  (SCEN4).
- (v) We divide the total number of 15 pilot symbols into 3 clusters of 5 consecutive pilot symbols each. The 3 clusters are centered at the positions (14) that correspond to  $K_p = 3$  (SCEN5).
- (vi) We divide the total number of 15 pilot symbols into 5 clusters of 3 consecutive pilot symbols each. The 5 clusters are centered at the positions (14) that correspond to  $K_p = 5$  (SCEN6).

Figure 8 shows the BER for each scenario with  $N = 4$ . We observe that SCEN2 and SCEN3 lead to essentially the same BER performance, that turns out to be very poor. The BER resulting from SCEN5 is slightly better, but still poor. Much better BER performance is obtained for SCEN1, SCEN4, and SCEN6, with SCEN1 yielding the best performance. The poor performance resulting from SCEN2, SCEN3, and SCEN5 comes from the poor conditioning of the  $15 \times 4$  matrix  $\Psi_P$ , yielding very large values when computing the inverse of  $\Psi_P^T \Psi_P$ . As the DCT basis functions  $\psi_0(k), \dots, \psi_3(k)$  change only slowly with  $k$ , SCEN2 yields a matrix  $\Psi_P$  with nearly identical rows, so it behaves like a matrix of rank 1. Similarly, the matrices  $\Psi_P$  that correspond to SCEN3 and SCEN5 behave like matrices of ranks 2 and 3, respectively. Hence, when the pilot symbols are placed in a number of clusters that are less than the number ( $N$ ) of DCT coefficients to be estimated, poor performance results. For SCEN1, SCEN4, and SCEN6, the number of pilot symbol clusters exceeds  $N$ ; the corresponding matrices  $\Psi_P$  are full-rank (rank = 4), and good performance results. Note that SCEN1 and SCEN4 can cope with values of  $N$  up to  $K_p$ , whereas SCEN6 cannot handle values of  $N$  in excess of 5.

In the following, we investigate the influence of the number of pilot symbols on the MSE and the BER. The constant-energy pilot symbols are inserted into the data sequence according to (14). For (14) to hold, the block length  $K$  should be an odd multiple of the number of pilot symbols  $K_p$ . We assume a total block length  $K = 105$  and simulate the BER and MSE for  $K_p = 7, 15$ , and 35. Figure 9 shows the BER degradation at  $\text{BER} = 10^{-4}$  with respect to the reference system, for a fixed ratio  $\eta = K_p/K = 20\%$  and various values of the block length  $K$ . The BER degradation  $-10 \log(1 - \eta)$  due to the insertion of pilot symbols (which amounts to 0.97 dB for  $\eta = 0.2$ ) is included. The following observation can be made.

- (i) For given block size  $K$ , there is an optimum number  $N_{\text{opt}}$  of DCT coefficients to be estimated that minimizes the BER degradation. This is consistent with the observation that the MSE of the phase estimate can be minimized by a suitable choice of  $N$ .
- (ii) For very small  $K$ ,  $N_{\text{opt}} = 1$ . The optimum value  $N_{\text{opt}}$  increases with increasing  $K$  because more DCT coefficients

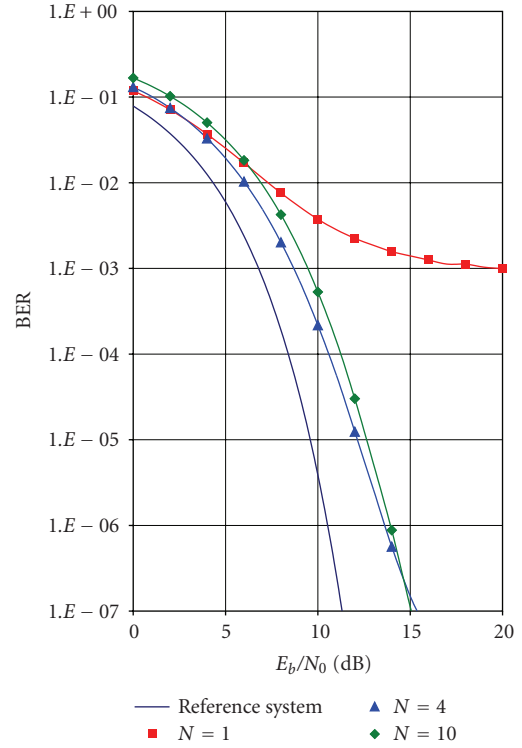


FIGURE 5: BER when Wiener phase noise with  $\sigma_\Delta = 3^\circ$  is present.  $K = 105$ ,  $K_p = 15$ .

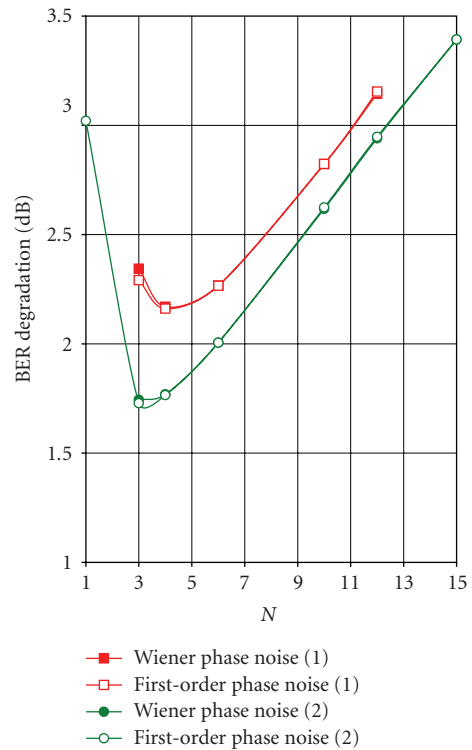


FIGURE 6: BER degradation as a function of the number of estimated coefficients  $N$  for Wiener phase noise and first-order phase noise with  $\alpha = 0.015$ . (1)  $\sigma_\Delta^2 = 0.0027 \text{ rad}^2$ ; (2)  $\sigma_\Delta^2 = 0.0015 \text{ rad}^2$ .  $K = 105$ ,  $K_p = 15$ .

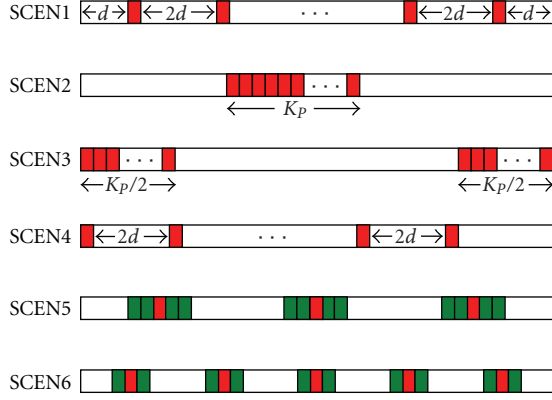
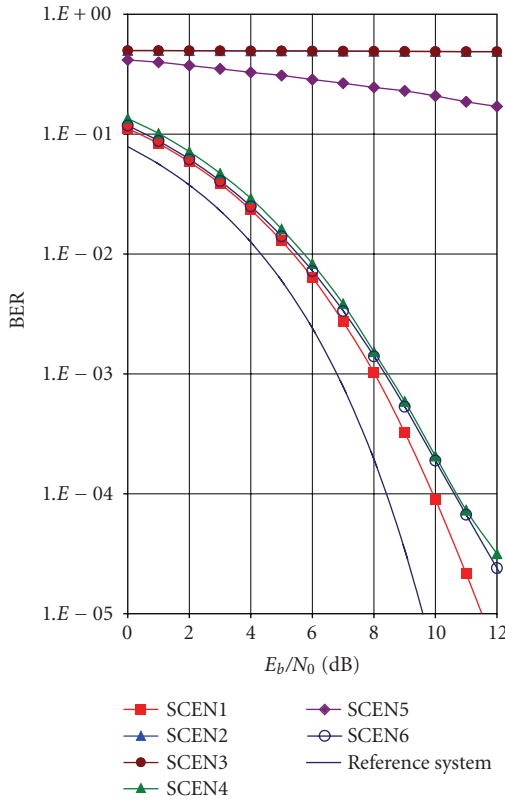
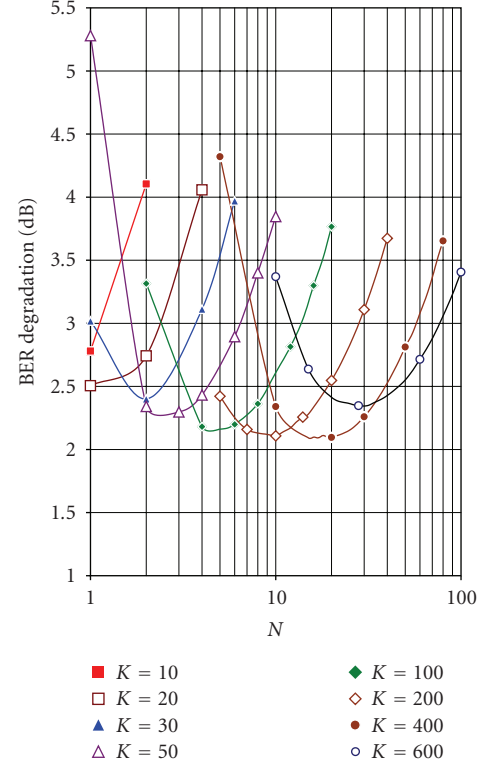


FIGURE 7: Pilot symbol insertion schemes.

FIGURE 8: BER for different pilot symbol placement scenarios.  $K = 105$ ,  $K_p = 15$ ,  $N = 4$ .

are needed to model the phase fluctuations when  $K$  gets larger. Keeping  $N = 1$  yields very large degradations when  $K$  increases.

(iii) The BER degradation that corresponds to  $N = N_{\text{opt}}$  exhibits a (broad) minimum as a function of  $K$ . As long as the fluctuation of  $\theta(k)$  about its time-average is small, so that linearization of the argument function in (11) applies, the degradation decreases with increasing  $K$  because the number  $K_p$  of noisy observations of the phase noise increases when the ratio  $K_p/K$  is fixed. However, for too large  $K$ , the fluctuation of the Wiener phase noise is so large that

FIGURE 9: BER degradation for  $\text{BER} = 10^{-4}$  as function of  $N$  for various  $K$  and fixed pilot symbol ratio  $\eta = K_p/K = 20\%$  and  $\sigma_\Delta = 3^\circ$ .

linearization is no longer valid (for Wiener phase noise, we need  $K\sigma_\Delta^2 \ll 1$  for the linearization to be accurate) and the resulting degradation increases with increasing  $K$ .

For the considered scenario, the minimum degradation occurs at  $(K_{\text{opt}}, N_{\text{opt}}) \approx (400, 20)$  and amounts to about 2.1 dB. When the actual block size  $K$  exceeds  $K_{\text{opt}}$ , the degradation can be limited by dividing the block in subblocks of at most  $K_{\text{opt}}$  symbols, and estimating the phase trajectory for each subblock separately.

Figure 10 shows the BER degradation when (1)  $\eta = 20\%$  and  $\sigma_\Delta = 3^\circ$  and (2)  $\eta = 10\%$  and  $\sigma_\Delta = 2^\circ$ , for the following phase noise estimation algorithms.

- (i) The proposed DCT-based algorithm with pilot symbol placement according to SCEN1 (14) and selection of the optimum  $N$ .
- (ii) Estimation of only the time-average of the phase noise, with the pilot symbols arranged according to SCEN3.
- (iii) The method from Luise et al. [11], with the pilot symbols arranged according to SCEN3. The phase noise over the total symbol block is approximated as a linear interpolation between the average phase values over the first and the second pilot symbol clusters.

We observe that estimating only the time-average or the linear trend of the phase noise yields poor BER performance, except for small  $K$ . For  $K = 10$ , the DCT-based algorithm



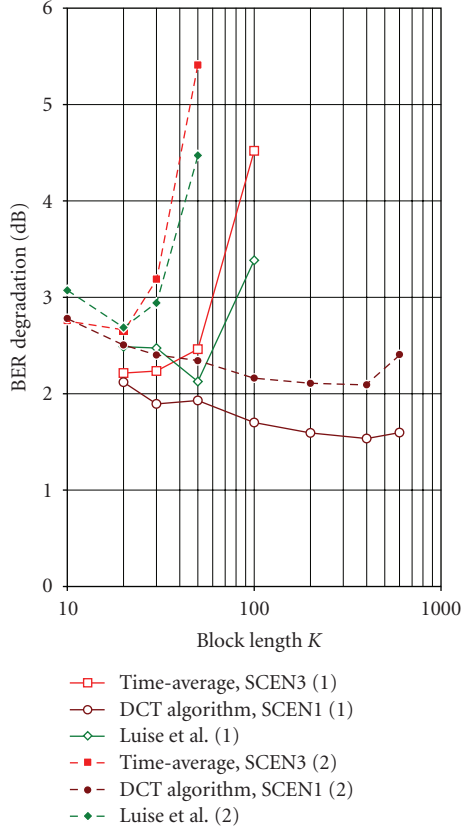


FIGURE 10: Comparison of BER degradation for  $\text{BER} = 10^{-4}$  as function of  $K$ . (1)  $\eta = 20\%$  and  $\sigma_{\Delta} = 3^{\circ}$ ; (2)  $\eta = 10\%$  and  $\sigma_{\Delta} = 2^{\circ}$ .

also estimates the time-average only (because  $N = 1$  is optimum for  $K = 10$ ); we observe that SCEN3 (with pilot symbols at positions 0 and 9) performs slightly better than the DCT-based algorithm (with pilot symbols at positions 2 and 7) for  $K = 10$ . However, when the block length is increased, the DCT algorithm that estimates multiple DCT coefficients outperforms both SCEN3 and Luise et al. and leads to a BER degradation that decreases with increasing  $K$  until an optimal value for  $K$  is reached.

## 6. Complexity Analysis

In order to assess the complexity of the proposed algorithm, we determine the number of complex multiplications required per symbol interval. The calculation of the second term in (18) requires the highest number of computations. This term can be evaluated in the following ways.

(1) In a first approach,  $(K/K_p)\Psi_K\Psi_P^T\mathbf{r}'$  is calculated via two matrix multiplications: first  $\Psi_P^T$  (dimension  $N \times K_p$ ) and  $\mathbf{r}'$  (dimension  $K_p \times 1$ ) are multiplied and then  $(K/K_p)\Psi_K$  (dimension  $K \times N$ ) and  $\Psi_P^T\mathbf{r}'$  (dimension  $N \times 1$ ) are multiplied. The resulting complexity is of the order  $O(NK_p + KN) \approx O(KN)$ , with the approximation holding for  $K \gg K_p$ . Hence, the complexity per symbol interval amounts to  $O(N)$ .

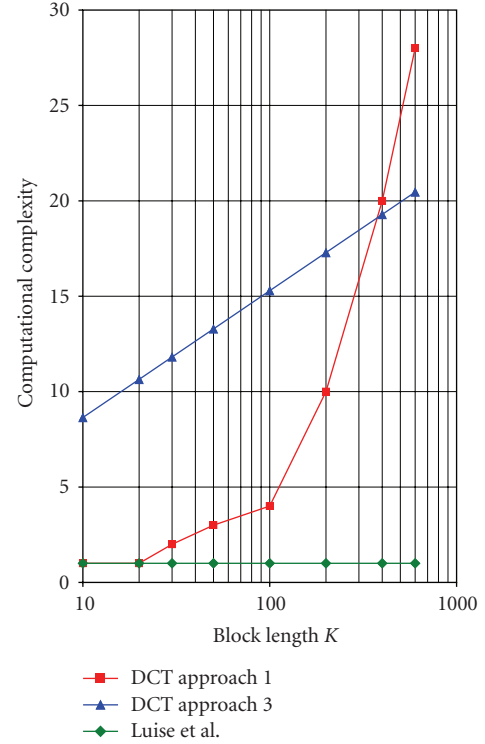


FIGURE 11: Complexity comparison for the proposed algorithm (approaches 1 and 3) and for Luise et al. algorithm.

(2) In a second approach,  $(K/K_p)\Psi_K\Psi_P^T\mathbf{r}'$  is calculated via a single-matrix multiplication:  $(K/K_p)\Psi_K\Psi_P^T$  (dimension  $K \times K_p$ ) and  $\mathbf{r}'$  (dimension  $K_p \times 1$ ) are multiplied. Taking into account that  $(K/K_p)\Psi_K\Psi_P^T$  can be computed offline, the resulting complexity per symbol is  $O(K_p)$ . As  $N \leq K_p$ , the first approach is to be preferred over the second approach.

(3) The third approach exploits the fact that  $\Psi_K$  and  $\Psi_P$  are submatrices of  $K \times K$  and  $K_p \times K_p$  DCT transform matrices, respectively. Hence, the two matrix multiplications from the first approach can be replaced by an inverse DCT transform (size  $K_p$ ) followed by a DCT transform (size  $K$ ). As  $K \gg K_p$ , the complexity of the size- $K$  DCT dominates. The DCT of a vector  $\{s(0), s(1), \dots, s(K-1)\}$  of length  $K$  can be obtained by calculating the discrete Fourier transform (DFT) of its even expansion  $\{s(K-1), \dots, s(1), s(0), s(0), s(1), \dots, s(K-1)\}$  (note that the even expansion has length  $2K$ ). As the FFT algorithm used for calculating the DFT of length  $M$  has a computational complexity  $O(M \log_2(M))$ , the complexity of the size- $K$  DCT is  $O(2K \log_2(2K))$ , yielding a complexity per symbol interval of  $O(\log_2(4K^2))$ .

The complexity per symbol interval of the phase noise estimation method used by Benvenuti et al. [11] is about  $O(1)$ . Figure 11 shows the order of complexity as a function of the block length  $K$ , for the proposed algorithm (approaches 1 and 3) and for Luise et al. algorithm; the result related to the first approach in the proposed algorithm corresponds to taking for each  $K$  the value of  $N$  that is optimum for  $\sigma_{\Delta} = 3^{\circ}$ . Luise et al. algorithm has

a smaller complexity than the proposed algorithm, but the latter algorithm outperforms the former, especially when the phase noise is strong. For the proposed algorithm, we notice that matrix multiplication according to the first approach leads to the lowest computational complexity for  $K < 400$ . As  $K$  becomes larger than 400, calculation via FFT (third approach) is less complex. At the point  $(K_{\text{opt}}, N_{\text{opt}}) = (400, 20)$  yielding minimum BER degradation (see Figure 9), the first and third approaches give rise to the same complexity.

## 7. Conclusions and Remarks

In this contribution, we have considered an ad hoc feed-forward data-aided phase noise estimation algorithm that is based on the estimation of only a few ( $N$ ) coefficients of the DCT basis expansion of the time-varying phase. The algorithm does not require detailed knowledge about the phase noise statistics. Linearization of the observation model has indicated that the mean-square error of the resulting estimate consists of an additive noise contribution (that increases with  $N$ ) and an MSE floor caused by the phase noise modeling error (that decreases with  $N$ ). The noise contribution coincides with the Cramer-Rao lower bound.

These analytical findings have been confirmed by means of computer simulations. The influence of the position and number  $K_P$  of pilot symbols inserted into the symbol sequence has been investigated. Computer simulations were carried out for several pilot symbol configurations. Arranging the pilot symbols according to (14), such that the subsampled DCT basis functions remain orthogonal, reduces the BER degradation as compared to the case of a preamble/postamble or midamble pilot symbol arrangement with estimation of only the time-average; in addition, the configuration (14) allows to estimate up to  $K_P$  DCT coefficients with a reduced computational complexity. The BER degradation can be minimized by a suitable choice of block length  $K$ , the number  $K_P$  of pilot symbols, and the number  $N$  of DCT coefficients to be estimated.

The considered DCT-based phase estimation algorithm makes use of the energy associated with the pilot symbols only. Further research will involve the incorporation of the DCT-based algorithm in an iterative phase noise estimation algorithm that exploits soft decisions about the data symbols, so that the resulting algorithm benefits from the energy associated with the data symbols as well. The performance and complexity of such an iterative algorithm will be investigated and compared to other iterative algorithms (such as those from [12–16]).

## Acknowledgments

The authors wish to acknowledge the activity of the Network of Excellence in Wireless COMMunications (NEWCOM++) of the European Commission (Contract no. 216715) that motivated this work. This work is also supported by the FWO Project G.0047.06 Advanced space-time processing

techniques for communication through multiantenna systems in realistic mobile channels.

## References

- [1] A. Demir, A. Mehrotra, and J. Roychowdhury, "Phase noise in oscillators: a unifying theory and numerical methods for characterization," *IEEE Transactions on Circuits and Systems I*, vol. 47, no. 5, pp. 655–674, 2000.
- [2] T. E. Parker, "Characteristics and sources of phase noise in stable oscillators," in *Proceedings of the 41st Annual Frequency Control Symposium*, pp. 99–110, Philadelphia, Pa, USA, May 1987.
- [3] G. B. Giannakis and C. Tepedelenlioglu, "Basis expansion models and diversity techniques for blind identification and equalization of time-varying channels," *Proceedings of the IEEE*, vol. 86, no. 10, pp. 1969–1986, 1998.
- [4] J. K. Tugnait and W. Luo, "Blind space-time multiuser channel estimation in time-varying DS-CDMA systems," *IEEE Transactions on Vehicular Technology*, vol. 55, no. 1, pp. 207–218, 2006.
- [5] O. Rousseaoux, G. Leus, and M. Moonen, "Estimation and equalization of doubly selective channels using known symbol padding," *IEEE Transactions on Signal Processing*, vol. 54, no. 3, pp. 979–990, 2006.
- [6] J.-C. Nallatamby, M. Prigent, E. Vauray, A. Laloue, M. Camiade, and J. Obregon, "Low phase noise operation of microwave oscillator circuits," *IEEE Transactions on Ultrasonics, Ferroelectrics and Frequency Control*, vol. 47, no. 2, pp. 411–420, 2000.
- [7] J. Mukherjee, "Optimizing MOSFET channel width for low phase noise in LC oscillators," in *Proceedings of the 50th Midwest Symposium on Circuits and Systems (MWSCAS '07)*, pp. 610–613, Montreal, Canada, August 2007.
- [8] D. Y. Jung and C. S. Park, "Power efficient Ka-band low phase noise VCO in 0.13  $\mu\text{m}$  CMOS," *Electronics Letters*, vol. 44, no. 10, pp. 628–630, 2008.
- [9] H. Meyr, M. Moeneclaey, and S. Fechtel, *Digital Communication Receivers: Synchronization, Channel Estimation, and Signal Processing*, Wiley Series in Telecommunications and Signal Processing, John Wiley & Sons, New York, NY, USA, 1998.
- [10] U. Mengali and A. N. D'Andrea, *Synchronization Techniques for Digital Receivers*, Plenum Press, New York, NY, USA, 1997.
- [11] L. Benvenuti, L. Giugno, V. Lottici, and M. Luise, "Code-aware carrier phase noise compensation on turbo-coded spectrally-efficient high-order modulations," in *Proceedings of the 8th International Workshop on Signal Processing for Space Communications (SPSC '03)*, pp. 177–184, Catania, Italy, September 2003.
- [12] G. Colavolpe, A. Barbieri, and G. Caire, "Algorithms for iterative decoding in the presence of strong phase noise," *IEEE Journal on Selected Areas in Communications*, vol. 23, no. 9, pp. 1748–1757, 2005.
- [13] J. Dauwels and H.-A. Loeliger, "Phase estimation by message passing," in *Proceedings of the IEEE International Conference on Communications (ICC '04)*, vol. 1, pp. 523–527, Paris, France, June 2004.
- [14] E. Panayirci, H. Cirpan, and M. Moeneclaey, "A sequential Monte Carlo method for blind phase noise estimation and data detection," in *Proceedings of the 13th European Signal Processing Conference (EUSIPCO '05)*, Antalya, Turkey, September 2005.

- [15] E. Panayırçı, H. A. Çırpan, M. Moeneclaey, and N. Noels, "Blind-phase noise estimation in OFDM systems by sequential Monte Carlo method," *European Transactions on Telecommunications*, vol. 17, no. 6, pp. 685–693, 2006.
- [16] S. Godtmann, N. Hadaschik, A. Pollok, G. Ascheid, and H. Meyr, "Iterative code-aided phase noise synchronization based on the LMMSE criterion," in *Proceedings of the 8th IEEE Signal Processing Advances in Wireless Communications (SPAWC '07)*, pp. 1–5, Helsinki, Finland, June 2007.
- [17] D. Petrovic, W. Rave, and G. Fettweis, "Effects of phase noise on OFDM systems with and without PLL: characterization and compensation," *IEEE Transactions on Communications*, vol. 55, no. 8, pp. 1607–1616, 2007.
- [18] ETSI, "Digital video broadcasting (dvb), second generation framing structure, channel coding and modulation systems for broadcasting, interactive services, news gathering and other broadband satellite applications".
- [19] V. S. Abhayawardhana and I. J. Wassell, "Common phase error correction with feedback for OFDM in wireless communication," in *Proceedings of the IEEE Global Telecommunications Conference (GLOBECOM '02)*, vol. 1, pp. 651–655, Taipei, Taiwan, November 2002.