

Optimization of pilot-aided DCT-based phase noise estimation

Jabran Bhatti

Marc Moeneclaey

Department of Telecommunications and Information Processing

Ghent University, Belgium

`jabran.bhatti@telin.ugent.be`

`mm@telin.ugent.be`

Abstract

The presented work addresses the issue of phase noise estimation for pilot-aided burst-mode transmission in digital communication systems. We propose to estimate the phase noise from a truncated discrete-cosine transform (DCT) expansion model. The key idea is to reconstruct the low-pass phase noise process via only a small number N of DCT coefficients of the phase expansion. An evident question that arises is how to choose N . Based on a few valid approximations, we derive an analytical expression of the bit-error rate (BER) degradation in the presence of residual phase noise, which allows us to determine the value of N that yields the minimum BER degradation.

1 Introduction

Imperfections and non-idealities in the transmitter and receiver local oscillators of a communication system cause the phase of the received signal to show a time-dependent behaviour. This phenomenon, called phase noise, can result in severe performance degradation. In the literature, several ways to tackle this problem have been proposed [1–5]. These methods, however, are either unsuited for burst-mode transmission, sensitive to additive noise or computationally complex. In [6], we present a low-complexity phase noise estimation method using a discrete-cosine transform (DCT) expansion model. The key idea is to reconstruct the low-pass phase noise process via the estimates of only a small number N of DCT coefficients of the phase expansion. We start by describing the system under investigation in section 2, including the considered low-pass phase noise model. In section 3 we present a slightly altered version of the phase noise estimation algorithm from [6], making it more robust against large phase fluctuations and enabling good performance for long bursts. In [6], it has been shown that the number of estimated DCT coefficients N has a large influence on the phase estimation error and the bit-error rate (BER) performance of the considered system. Based on an approximate expression of the BER degradation, a practical technique to determine the value $N = N_{min}$ that yields near-optimum BER performance is presented in section 4. Performance of the presented techniques is assessed by means of computer simulations in section 5.

2 System description and phase noise model

In our system, we consider the transmission of uncoded 4-QAM symbols in bursts of length K over an additive white Gaussian noise (AWGN) channel perturbed by phase noise. In order to facilitate the estimation procedure, K_P pilot symbols are included. Assuming perfect timing and frequency synchronization, the scaled observations at the matched filter output are given by

$$r_k = a_k e^{j\theta_k} + w_k, \quad (1)$$

where $k = 0, \dots, K-1$ refers to the k -th symbol interval with duration T , the additive noise $\{w_k\}$ is a sequence of i.i.d. zero-mean (ZM) circularly symmetric complex-valued Gaussian (CSCG) random variables (RVs) with $E[|w_k|^2] = N_0/E_s$ and $\{a_k\}$ is the transmitted symbol sequence (with symbol energy $E[|a_k|^2] = 1$), which also contains the pilot symbols at positions $\{k_i, i = 0, \dots, K_P - 1\}$. The time-varying phase θ_k consists of a static phase offset θ_{stat} and a ZM phase noise process. We consider Wiener phase noise which is described by the following system equation:

$$\theta_{k+1} = \theta_k + \Delta_k, \text{ for } k = 0, \dots, K-2,$$

where θ_0 is uniformly distributed in $[-\pi, \pi]$ and $\{\Delta_k\}$ is a sequence of ZM Gaussian RVs with variance σ_Δ^2 . The Wiener phase noise model can be characterized by the power spectral density $S_\theta(e^{j2\pi fT})$:

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

where the approximation in (2) is valid when $fT \ll 1/2$. Due to its low-pass character, a valid approximation would be to represent θ_k by a weighed sum of a limited number $N (\ll K)$ of suitable basis functions $\psi_{k,n}$:

$$\theta_k \approx \sum_{n=0}^{N-1} x_n \psi_{k,n}. \quad (3)$$

Note that in expression (3), the higher-order expansion coefficients are neglected, resulting in a modeling error. We use the basis functions from the discrete-cosine transform (DCT), which are particularly well-suited to represent low-pass processes [7]. The orthonormal DCT basis functions are defined as

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

From the observations (1) at the pilot symbol positions, we will compute an estimate $\{\hat{x}_n, n = 0, \dots, N-1\}$ of the DCT coefficients $\{x_n, n = 0, \dots, N-1\}$, using the model (3) with equality. The resulting phase estimate is obtained via the inverse DCT:

$$\hat{\theta}_k = \sum_{n=0}^{N-1} \hat{x}_n \psi_{k,n}.$$

This phase estimate is used to derotate the received signal before data detection. The detector is designed under the assumption of perfect carrier synchronization, i.e., $\hat{\theta}_k = \theta_k$. For uncoded transmission, the detection algorithm reduces to symbol-by-symbol detection.

3 Phase noise estimation

Considering (3) with equality, we can write

$$\boldsymbol{\theta} = \boldsymbol{\Psi}_K \mathbf{x}, \quad (4)$$

where $\boldsymbol{\theta} = (\theta_0, \dots, \theta_{K-1})^T$, $\mathbf{x} = (x_0, \dots, x_{N-1})^T$ and the elements of the DCT matrix are defined as $(\Psi_{\mathbf{K}})_{k,n} = \psi_{k,n}$ for $k = 0, \dots, K-1$ and $n = 0, \dots, N-1$. Combining (1) and (4), the maximum-likelihood (ML) estimate of \mathbf{x} based on the observations at pilot symbol positions is obtained as

$$\hat{\mathbf{x}} = \arg \max_{\mathbf{x}} \sum_{i=0}^{K_P-1} |r_{k_i} a_{k_i}^*| \cos(\arg(r_{k_i} a_{k_i}^*) - (\Psi_{\mathbf{K}} \mathbf{x})_{k_i}). \quad (5)$$

As the function to be maximized in (5) depends on \mathbf{x} in a highly non-linear way, obtaining a closed form expression of $\hat{\mathbf{x}}$ is infeasible. Instead, as an alternative to (5), we propose to compute the least-squares estimate $\hat{\mathbf{x}}$ that minimizes $\sum_i |arg(r_{k_i} a_{k_i}^*) - (\Psi_{\mathbf{K}} \mathbf{x})_{k_i}|^2$ in [6]. However, as the codomain of the function $arg(\cdot)$ is limited to an interval $[-\pi, \pi]$, this approach is susceptible to phase wrapping. In [6], we avoid phase wrapping by rotating the observations over the estimated average phase $\hat{\theta}_{stat}$. For small phase variations, this approach succeeds in restricting the angle of $r_{k_i} a_{k_i}^* e^{-j\hat{\theta}_{stat}}$ to $[-\pi, \pi]$. For long bursts, however, the phase fluctuations become larger and compensating for the average phase no longer guarantees a small phase wrapping probability. Here, we modify the technique from [6] by unwrapping the $arg(\cdot)$ function. The unwrapped angles $\{\angle y_k, k = 0, \dots, K-1\}$ of the elements of a given sequence $\{y_k, k = 0, \dots, K-1\}$ are determined by the following recursive rule:

$$\angle y_k = \angle y_{k-1} + \text{mod}(\arg(y_k) - \arg(y_{k-1}), 2\pi),$$

where $\angle y_0 = \arg(y_0)$ and $\text{mod}(y, 2\pi)$ returns the modulus after division of y by 2π . In the remainder of this paper, $\arg(y)$ denotes the function that returns the unwrapped angle of y . Least-squares estimation results in the following estimate of \mathbf{x} :

$$\hat{\mathbf{x}} = (\Psi_{\mathbf{P}}^T \Psi_{\mathbf{P}})^{-1} \Psi_{\mathbf{P}}^T \mathbf{z},$$

where $(\mathbf{z})_i = \arg(r_{k_i} a_{k_i}^*)$ and $(\Psi_{\mathbf{P}})_{i,n} = \psi_{k_i,n}$ for $i = 0, \dots, K_P-1$. In order that $\Psi_{\mathbf{P}}^T \Psi_{\mathbf{P}}$ be invertible, we need $K_P \geq N$. It is particularly convenient to arrange the pilot symbols equidistantly according to

$$k_i = \frac{K(2i+1) - K_P}{2K_P}, \quad (6)$$

for $i = 0, \dots, K_P-1$, as this yields a diagonal matrix $\Psi_{\mathbf{P}}^T \Psi_{\mathbf{P}}$: $\Psi_{\mathbf{P}}^T \Psi_{\mathbf{P}} = K_P/K \mathbf{I}_N$ [6], where \mathbf{I}_N is the $N \times N$ identity matrix. Note that, for correct phase wrapping to occur, the distance between two consecutive pilot symbols should not be too large. As the pilot spacing corresponding to (6) is given by K/K_P , this implies that for a given burst length K , a sufficient number of pilot symbols should be included. In the remainder of this paper, we assume the pilot symbols are inserted according to (6) and correct phase wrapping takes place. The resulting DCT coefficient and phase estimates are then given by

$$\hat{\mathbf{x}} = \frac{K}{K_P} \Psi_{\mathbf{P}}^T \mathbf{z} \quad (7)$$

$$\hat{\boldsymbol{\theta}} = \frac{K}{K_P} \Psi_{\mathbf{K}} \Psi_{\mathbf{P}}^T \mathbf{z}. \quad (8)$$

Note that the estimation procedure involves neglecting the $K-N$ higher-order DCT coefficients. As a result, this modeling error introduces an estimation error floor. When N is too small, the truncated DCT phase model becomes inadequate and the resulting error floor is intolerably high. On the other hand, when N is chosen too large, sensitivity of the phase estimate to additive noise increases, leading to a large estimation error. An efficient method to determine the value of N is presented in the next section.

4 Optimization of the BER degradation

In the following, we derive an approximate expression for the BER degradation resulting from the residual phase noise after rotating the observations (1) over the estimated phase. Denoting E_b as the energy per information bit, 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* BER_{ref} . The average BER for perfect synchronization with 4-QAM symbols is given by:

$$BER_{4-QAM}(E_b/N_0) = Q\left(\sqrt{\frac{2E_b}{N_0}}\right), \quad (9)$$

where the Q -function is defined as

$$Q(v) = \frac{1}{\sqrt{2\pi}} \int_v^{+\infty} e^{-\frac{u^2}{2}} du.$$

In the presence of an impairment (such as e.g., residual phase noise), the BER increases as compared to $BER_{4-QAM}(E_b/N_0)$. The corresponding average BER is denoted as $BER_{avg}(E_b/N_0)$. Let us denote as $(E_b/N_0)_{ref}$ and $(E_b/N_0)_{req}$, the values of E_b/N_0 which yield

$$\begin{aligned} BER_{4-QAM}((E_b/N_0)_{ref}) &= BER_{ref} \\ BER_{avg}((E_b/N_0)_{req}) &= BER_{ref}. \end{aligned}$$

The BER degradation - expressed in dB - is then defined as

$$deg = 10 \log_{10} \left(\frac{(E_b/N_0)_{req}}{(E_b/N_0)_{ref}} \right).$$

Suppose we have obtained a reasonably good phase estimate $\hat{\theta}$ from (8). Then, the variance $\sigma_{\phi_k}^2$ of the residual phase noise $\phi_k = \theta_k - \hat{\theta}_k$ at position k can be safely assumed to be small. Although $\sigma_{\phi_k}^2$ does depend on the position k , for long bursts ($K \gg 1$), we can approximate $\sigma_{\phi_k}^2$ by its average over all positions: $\sigma_{\phi_k}^2 \approx \sigma_{\phi}^2$. Furthermore, we can assume the residual phase noise has a ZM Gaussian distribution with small variance $\sigma_{\phi}^2 \ll 1$. Note that the accuracy of the phase estimate depends on E_s/N_0 and hence the residual phase noise variance is a function of the signal-to-noise ratio: $\sigma_{\phi}^2 = \sigma_{\phi}^2(E_s/N_0)$.

The derotated observations $r'_k = r_k e^{-j\hat{\theta}_k}$ are given by

$$r'_k = a_k e^{j\phi_k} + w'_k, \quad (10)$$

where $\{w'_k\}$ is a sequence of ZM CSCG RVs with $E[|w'_k|^2] = N_0/E_s$. For small σ_{ϕ} , linearization of the exponential function in (10) yields

$$\begin{aligned} r'_k &\approx a_k(1 + j\phi_k) + w'_k \\ &\approx a_k + n_k, \end{aligned} \quad (11)$$

*For transmission of uncoded data symbols, the reference value is typically chosen as $BER_{ref} = 10^{-4}$.

where $\{n_k\}$ is a sequence of ZM CSCG RVs with $E[|n_k|^2] = N_0/E_s + \sigma_\phi^2(E_s/N_0)$. Expression (11) indicates that under the made assumptions, the derotated observations can be viewed as the observations resulting from transmission over a perfectly synchronized AWGN channel where the additive noise n_k has a larger variance than w_k from (1). The BER is easily obtained as

$$BER_{avg} \approx Q \left(\sqrt{\left(\frac{N_0}{E_s} + \sigma_\phi^2(E_s/N_0) \right)^{-1}} \right). \quad (12)$$

When employing K_P pilot symbols, the relationship between the energy per information bit E_b and the energy per transmitted symbol E_s for 4-QAM is given by

$$KE_s = 2(K - K_P)E_b. \quad (13)$$

The variance of the residual phase noise is given by

$$\begin{aligned} \sigma_\phi^2(E_s/N_0) &= \frac{1}{K} \sum_{k=0}^{K-1} E \left[\left(\hat{\theta}_k - \theta_k \right)^2 \right] \\ &= \frac{1}{K} \sum_{n=0}^{N-1} E \left[(\hat{x}_n - x_n)^2 \right] + \frac{1}{K} \sum_{n=N}^{K-1} E \left[x_n^2 \right], \end{aligned} \quad (14)$$

where the second equality follows from the orthonormality of the DCT basis functions. We note that the first term in (14) is lower bounded by the Cramer-Rao lower bound relating to the estimation of the N lower order DCT coefficients. From (7), we easily obtain [6]:

$$\frac{1}{K} \sum_{n=0}^{N-1} E \left[(\hat{x}_n - x_n)^2 \right] \geq \frac{N_0}{2E_s} \frac{N}{K_P}. \quad (15)$$

The second term in (14) can be computed as follows:

$$\frac{1}{K} \sum_{n=N}^{K-1} E \left[x_n^2 \right] = \frac{1}{K} \int_{1/T} S_\theta \left(e^{j2\pi fT} \right) \sum_{n=N}^{K-1} |\Psi_n \left(e^{j2\pi fT} \right)|^2 df,$$

where $\Psi_n \left(e^{j2\pi fT} \right)$ is the discrete-time Fourier transform of the basis function $\psi_{k,n}$. For sufficiently large K , $\Psi_n \left(e^{j2\pi fT} \right)$ is centered around the frequencies $-n/2KT$ and $n/2KT$. Hence, we can apply the following approximation:

$$\begin{aligned} \frac{1}{K} \sum_{n=N}^{K-1} E \left[x_n^2 \right] &\approx \frac{1}{K} \sum_{n=N}^{K-1} S_\theta \left(e^{j2\pi n/2K} \right) \int_{1/T} |\Psi_n \left(e^{j2\pi fT} \right)|^2 df \\ &= \frac{1}{K} \sum_{n=N}^{K-1} S_\theta \left(e^{j2\pi n/2K} \right) \\ &\approx 2T \int_{\frac{N}{2KT}}^{\frac{1}{2T}} S_\theta \left(e^{j2\pi fT} \right) df. \end{aligned}$$

Keeping in mind that $S_\theta \left(e^{j2\pi fT} \right)$ is given by (2) yields

$$\frac{1}{K} \sum_{n=N}^{K-1} E \left[x_n^2 \right] \approx \frac{\sigma_\Delta^2}{\pi^2} \left(\frac{K}{N} - 1 \right). \quad (16)$$

Combining expressions (14), (15) and (16) we obtain the following approximate expression for the residual phase noise variance:

$$\sigma_\phi^2(E_s/N_0) \approx \frac{N_0}{2E_s} \frac{N}{K_P} + \frac{\sigma_\Delta^2}{\pi^2} \left(\frac{K}{N} - 1 \right). \quad (17)$$

Using (13) and (17) and equating the arguments of the Q-functions in (9) and (12) at the reference BER value BER_{ref} yields:

$$\frac{1}{2(E_b/N_0)_{ref}} = \frac{1}{2(E_b/N_0)_{req}} \frac{K}{K - K_P} \left(1 + \frac{N}{2K_P} \right) + \frac{\sigma_\Delta^2}{\pi^2} \left(\frac{K}{N} - 1 \right).$$

Hence, an approximate expression for the BER degradation is given by

$$\begin{aligned} deg = 10 \log_{10} \left(\frac{K}{K - K_P} \right) + 10 \log_{10} \left(1 + \frac{N}{2K_P} \right) \\ - 10 \log_{10} \left(1 - 2(E_b/N_0)_{ref} \frac{\sigma_\Delta^2}{\pi^2} \left(\frac{K}{N} - 1 \right) \right). \end{aligned} \quad (18)$$

For small values of the BER degradation, expression (18) can be further simplified by linearizing the $\log_{10}(\cdot)$ function:

$$deg = \frac{10}{\ln(10)} \left(\frac{K_P}{K} + \frac{N}{2K_P} + 2(E_b/N_0)_{ref} \frac{\sigma_\Delta^2}{\pi^2} \left(\frac{K}{N} - 1 \right) \right). \quad (19)$$

Expression (19) indicates that, for given K and K_P an optimal value of N exists that optimizes the BER performance. For fixed K and K_P , the minimum value of the BER degradation (19) is achieved when the number of estimated DCT coefficients equals N_{min} :

$$N_{min} = \text{round} \left(\sqrt{4 \frac{\sigma_\Delta^2}{\pi^2} K_P K (E_b/N_0)_{ref}} \right), \quad (20)$$

where $\text{round}(y)$ selects the integer value closest to y .

5 Numerical results

In this section, we assess the performance of the proposed techniques from sections 3 and 4 using computer simulations. We first verify the accuracy of the procedure to select the optimal number of estimated DCT coefficients from section 4, by comparing the analytically obtained BER degradation (19) with the BER degradation obtained via simulations. We consider transmission of a burst of 4-QAM symbols of length $K = 105$, including $K_P = 15$ pilot symbols inserted according to (6). Figure 1 illustrates the analytically obtained BER degradation (19) and the exact BER degradation obtained via simulations at $BER_{ref} = 10^{-4}$ as a function of N . Two phase noise levels, $\sigma_\Delta = 1^\circ$ and $\sigma_\Delta = 3^\circ$ are considered. We observe that although the exact BER degradation deviates from (19), the behaviour as a function of N is very similar. Moreover, as the BER degradation shows a rather broad minimum for N , the value of N_{min} obtained through expression (20) is very accurate. Figure 2 shows the minimum BER degradation as a function of the burst length K , for a fixed pilot symbol ratio $K_P/K = 10\%$ and $\sigma_\Delta = 3^\circ$, where for each value of K , the number of estimated coefficients N is optimized using (20). The dashed line represents the BER degradation

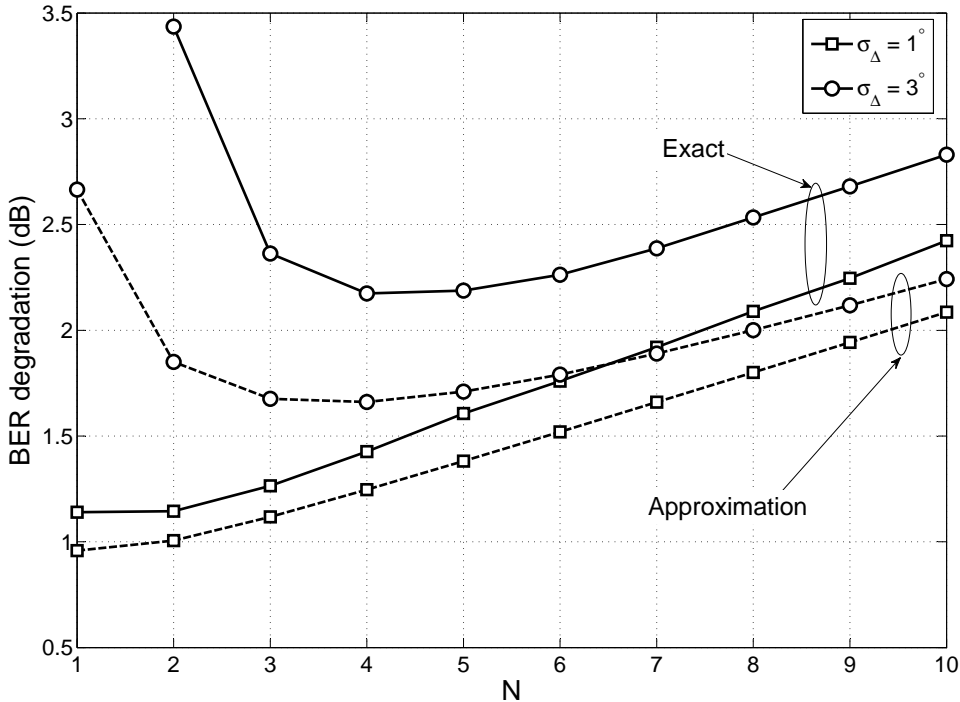


Figure 1: *Approximate versus exact BER degradation for $K = 105$ and $K_P = 15$.*

resulting from the phase noise estimation algorithm from [6], where the observations are rotated over the estimated average angle, prior to estimating the phase fluctuations. The solid line curve is the BER degradation resulting from the phase noise estimation algorithm presented here, which makes use of the unwrapped argument function. For small burst lengths ($K < 400$), we observe that both estimation strategies yield comparable BER degradation. However, as K becomes larger, the approach from [6] shows a severely deteriorated BER performance. The BER degradation becomes larger since, for longer bursts, phase wrapping may still occur, despite compensating for the average phase. As long as the pilot symbol spacing is sufficiently small, these issues do not occur when using the unwrapped argument function. In this case, the BER degradation decreases as K increases (although the decrease is less prominent as the burst size becomes larger), since the influence of the additive noise is reduced due to the larger number of observations that are available.

Acknowledgement

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.

References

- [1] H. Meyr, M. Moeneclaey, and S. Fechtel, *Digital Communication Receivers : Synchronization, Channel Estimation and Signal Processing*. USA: Wiley Series in

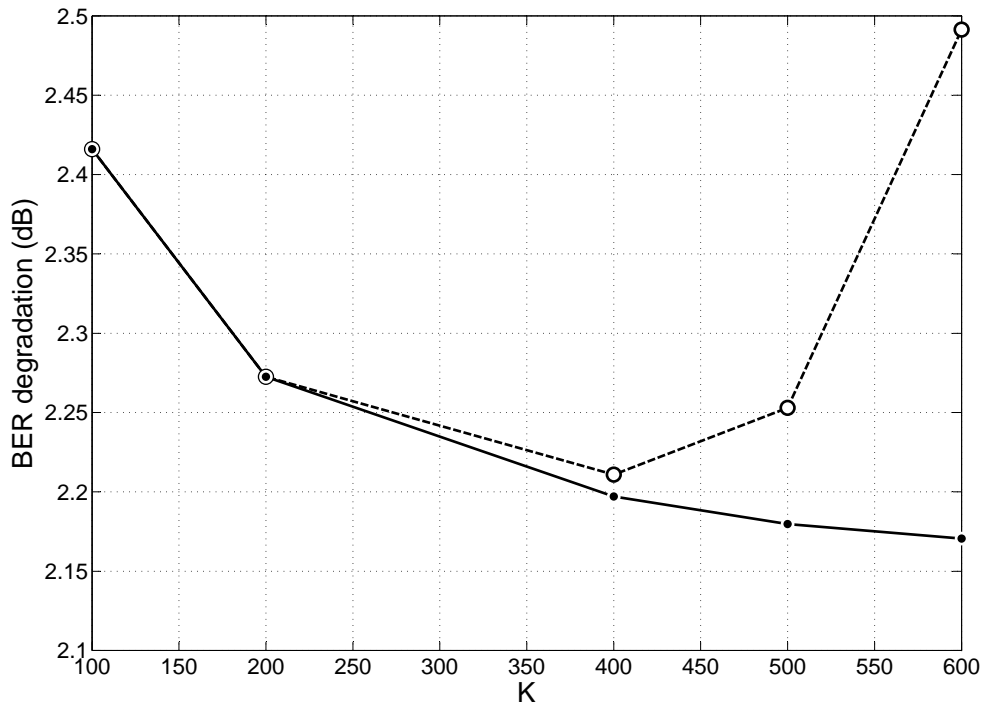


Figure 2: BER degradation at $BER_{ref} = 10^{-4}$ as a function of K .

Telecommunications and Signal Processing, 1998.

- [2] U. Mengali and A. D’Andrea, *Synchronization Techniques for Digital Receivers*. New York, USA: Plenum Press, 1997.
- [3] L. Benvenuti, L. Giugno, V. Lottici, and M. Luise, “Code-aware carrier phase noise compensation on turbo-coded spectrally-efficient high-order modulations,” in *Proc. of 8-th Int. Workshop on Sig. Proc. for Space Comm.*, (Catania, Italy), pp. 177–184, September 2003.
- [4] E. Panayirci, H. Cirpan, and M. Moeneclaey, “A sequential Monte-Carlo method for blind phase noise estimation and data detection,” in *Proceedings of European Signal Processing Conference (EUSIPCO)*, (Antalya, Turkey), September 2005.
- [5] 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, pp. 1748–1757, September 2005.
- [6] J. Bhatti and M. Moeneclaey, “Feedforward data-aided phase noise estimation from a DCT basis expansion,” *EURASIP JWCN, Special Issue on Synch. in Wireless Comm.*, January 2009.
- [7] T. Ahmed, N. Natarajan and K. Rao, “Discrete cosine transform,” *IEEE Trans. Computers*, vol. 23, pp. 90 – 93, January 1974.