PILOT-AIDED CARRIER SYNCHRONIZATION USING AN APPROXIMATE DCT-BASED PHASE NOISE MODEL

Jabran Bhatti, Marc Moeneclaey

Jabran.Bhatti@telin.ugent.be
Ghent University, Dept. of Telecommunications and Information Processing,
DIGCOM Research Group
Sint-Pietersnieuwstraat 41, B-9000 Gent, Belgium

ABSTRACT

This contribution deals with phase noise estimation from pilot symbols. The phase noise process is approximated by an expansion of DCT basis functions containing only a few terms. We propose an algorithm that estimates the DCT coefficients without requiring detailed knowledge about the phase noise statistics. We demonstrate that the resulting (linearized) mean-square 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. Performance can be optimized by a proper selection of the symbol block length and of the number of DCT coefficients to be estimated. For large block sizes, considerable performance improvement is found as compared to the case where only the time-average of the carrier phase is estimated.

1. INTRODUCTION

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 lowpass processes with a bandwidth that is limited by the Doppler frequency [1, 2, 3].

Several methods for phase noise estimation exist :

- Phase noise can be estimated by means of a feedback algorithm that operates according to the principle of the PLL. As feedback algorithms give rise to rather long acquistion periods, they are not well suited to systems with burst transmission [4, 5].
- Phase noise is approximated as piecewise constant over the observation interval. In each subinterval over which the phase is assumed to be constant, a conventional feedforward algorithm is used to estimate the local time-average of the phase [4, 5, 6].
- Recently, a factor graph approach for the estimation of Markov-type phase noise has been presented in [7], but the algorithm appears rather cumbersome and assumes detailed knowledge about the phase noise statistics at the receiver.

In this contribution, we apply the basis expansion model to the problem of phase noise estimation; the considered basis functions are those 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. The phase noise estimation algorithm, based on the estimation of only a few DCT coefficients, is derived in section 3. Section 4 contains the performance analysis of the proposed algorithm in terms of the mean-square error (MSE) of the phase estimate; we consider the Cramer-Rao lower bound corresponding to the actual observation model and the performance analysis of the linearized observation model. Analysis results are confirmed by computer simulations in section 5, which considers both the mean-square phase estimation error and the associated bit error rate (BER) degradation. Conclusions are drawn in section 6.

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:

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

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 circular symmetric complex valued Gaussian random variables with $E[|w(k)|^2] =$

N_0, and $\theta(k)$ is the sum of a static phase offset θ_0 and a zero-mean phase noise process with KxK correlation matrix $\mathbf{R}_\theta$.

The symbol sequence $\{a(k)\}$ contains known pilot symbols at positions $k_i,\ i = 0, ..., K_P - 1$, 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, i.e., 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, i.e., $\hat{\theta}(k) = \theta(k)$. For uncoded transmission, the detection algorithm reduces to symbol-by-symbol detection:

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

with A denoting the symbol constellation.

3. PHASE ESTIMATION ALGORITHM

The phase $\theta(k)$ can be represented as a weighed sum of K basis functions over the interval $(0, K - 1)$.

$$\theta(k) = \sum_{n=0}^{K-1} x_n \psi_n(k),\ k = 0, ..., K - 1 \qquad (2)$$

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

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

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

$$\psi_n(k) = \begin{cases} \sqrt{\frac{1}{K}} & n = 0 \\ \sqrt{\frac{2}{K}} cos\left(\frac{\pi n}{K}\left(k + \frac{1}{2}\right)\right) & n > 0 \end{cases} \qquad (4)$$

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

In the following we will produce from the observation $\{r(k_i)\}$ an estimate $\{\hat{x}_n, n = 0, ..., N - 1\}$ of the coefficients $\{x_n, n = 0, ..., N - 1\}$, using the phase model (3) with equality. The final estimate $\hat{\theta}(k)$ will be obtained by computing the inverse DCT of $\{\hat{x}_n\}$:

$$\hat{\theta}(k) = \sum_{n=0}^{N-1} \hat{x}_n \psi_n(k) \text{ for } k = 0, ..., K - 1 \qquad (5)$$

However, as (3) is not an exact model of the true phase $\theta(k)$, the phase estimate will be 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 (3) holds with equality, we obtain:

$$\mathbf{r}_P = \mathbf{D}(\mathbf{x})\mathbf{a}_P + \mathbf{w}_P \qquad (6)$$

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

$$(\mathbf{D}(\mathbf{x}))_i = e^{j(\mathbf{\Psi_P x})_i} \qquad (7)$$

and $(\mathbf{\Psi_P})_{i,n} = \psi_n(k_i), (\mathbf{x})_n = x_n,\ n = 0, ..., N - 1$ with $N \leq K_P$.

Maximum likelihood estimation of $\mathbf{x}$ from $\mathbf{r}_P$ results in

$$\hat{x}_{ML} = arg \min_{\mathbf{x}} |\mathbf{r}_P - \mathbf{D}(\mathbf{x})\mathbf{a_P}|^2 \qquad (8)$$

As $\mathbf{x}$ enters the observation $\mathbf{r}_P$ in a non-linear way, the ML estimate is not easily obtained. Therefore, we resort to a suboptimum ad-hoc estimation of **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 static phase offset θ_0 is close to $-\pi$ or π. In order to reduce the probability of phase wrapping, we first rotate the observation $\mathbf{r}$ over an angle θ_{avg} that is close to the time-average of $\theta(k)$, then we estimate the DCT coefficients of the fluctuation $\theta(k) - \theta_{avg}$ and finally, we compute the phase estimate $\hat{\theta}(k)$. We select

$$\theta_{avg} = arg\left(\sum_{i=0}^{K_p-1} r(k_i)\right) \qquad (9)$$

and construct $\mathbf{r}'$ with

$$(\mathbf{r}')_i = arg(r(k_i)a^*(k_i)exp(-j\theta_{avg})) \qquad (10)$$

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

$$\hat{\mathbf{x}}' = (\mathbf{\Psi_P}^T \mathbf{\Psi_P})^{-1} \mathbf{\Psi_P}^T \mathbf{r}' \qquad (11)$$

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

$$\hat{\boldsymbol{\theta}} = \theta_{avg}\mathbf{1} + \mathbf{\Psi_K}\hat{\mathbf{x}}' \qquad (12)$$

with $(\hat{\boldsymbol{\theta}})_k = \hat{\theta}(k), (\mathbf{1})_k = 1, (\mathbf{\Psi_K})_{k,n} = \psi_n(k), k = 0, ..., K-1; n = 0, ..., N-1$. In order to avoid matrix inversion in (11), we select the positions k_i of the K_P pilot symbols such that $\mathbf{\Psi_P}^T\mathbf{\Psi_P}$ is diagonal. In other words, the functions $\psi_n(k_i)$ must form N orthogonal functions of length K_P. Let us consider the DCT basis functions $\phi_n(i)$ of length K_P. By selecting

$$k_i = \frac{K(2i+1) - K_P}{2K_P} \tag{13}$$

we obtain

$$\psi_n(k_i) = \sqrt{\frac{K_P}{K}}\phi_n(i) \text{ for } n = 0, ..., K_P \tag{14}$$

so that $\mathbf{\Psi_P}^T\mathbf{\Psi_P} = \frac{K_P}{K}\mathbf{I}_N$, with $\mathbf{I}_N$ denoting the NxN identity matrix. When (13) holds, (11) and (12) reduce to

$$\hat{\mathbf{x}}' = \frac{K}{K_P}\mathbf{\Psi_P}^T\mathbf{r}' \tag{15}$$

$$\hat{\boldsymbol{\theta}} = \theta_{avg}\mathbf{1} + \frac{K}{K_P}\mathbf{\Psi_K}\mathbf{\Psi_P}^T\hat{\mathbf{r}}' \tag{16}$$

Note from (16) that the estimation algorithm does not need specific knowledge about the phase noise process. In order that k_i from (13) be integer, K must be an odd multiple of K_P. However, when K is not an odd multiple of K_P, rounding the right-hand side of (13) to the nearest integer gives rise to pilot symbol positions that still yield an essentially diagonal matrix $\mathbf{\Psi_P}^T\mathbf{\Psi_P}$.

4. 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 (10) in order to obtain tractable results.
Linearization of the argument function yields

$$r'(i) = arg(r(k_i)a^*(k_i)e^{-j\theta_{avg}}) \tag{17}$$
$$= \theta(k_i) - \theta_{avg} + n_P(i), i = 0, ..., K_P - 1 \tag{18}$$

where $\{n_P(i)\}$ is a sequence of i.i.d. zero-mean Gaussian random variables with variance $\frac{N_0}{2E_s}$. Note that (18) incorporates the true phase $\theta(k_i)$ instead of the approximate model (3), so that our performance analysis will take the modeling error into account. Substituting (18) into (16) and using $\mathbf{\Psi}_P^T\mathbf{\Psi}_P = \frac{K_P}{K}\mathbf{I}_N$ yields

$$\hat{\boldsymbol{\theta}} = \frac{K}{K_P}\mathbf{\Psi}_K\mathbf{\Psi}_P^T(\boldsymbol{\theta_P} + \mathbf{n}_P) \tag{19}$$

with $(\mathbf{n}_P)_i = n_P(i)$ and $(\boldsymbol{\theta_P})_i = \theta(k_i)$. If the model (3) were exact, we would have $\boldsymbol{\theta} = \mathbf{\Psi_K}\mathbf{x}$ and $\boldsymbol{\theta_P} = \mathbf{\Psi_P}\mathbf{x}$ yielding

$$\hat{\boldsymbol{\theta}} = \boldsymbol{\theta} + \frac{K}{K_P}\mathbf{\Psi}_K\mathbf{\Psi}_P^T\mathbf{n}_P \tag{20}$$

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

$$MSE = \frac{1}{K}E\left[|\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}|^2\right] \tag{21}$$

Substituting (19) into (21) yields

$$MSE = \frac{N_0}{2E_s}\frac{N}{K_P} + MSE_\infty \tag{22}$$

The first term in (22) denotes the contribution from the additive noise, whereas the second term in (22) constitutes a MSE floor, caused by the phase noise modeling error. Note that the noise contribution to the MSE is proportional to N (because N parameters need to be estimated), whereas the MSE floor decreases with increasing N (because the modeling error is reduced when more DCT coefficients are taken into account). Hence, there is an optimum value of N that minimizes the MSE; this optimum value depends on $\frac{E_s}{N_0}, K, K_P$ and the phase noise statistics.
From the nonlinear observation model (6), which assumes that (3) holds with equality, we compute the Cramer-Rao lower bound on the MSE (21) resulting from any unbiased estimate $\hat{\mathbf{x}}$ of the DCT coefficients of $\theta(k)$:

$$MSE \geq \frac{1}{K}tr\left(\mathbf{J}^{-1}\right) \tag{23}$$

In (23), $\mathbf{J}$ denotes the Fisher information matrix related to the estimation of $\mathbf{x}$ from (6), which is found to be :

$$(\mathbf{J})_{n,n'} = \frac{2E_s}{N_0}\frac{K}{K_P}\delta_{n-n'} \tag{24}$$

Combining (23) with (24) yields the following performance bound:

$$MSE \geq \frac{N_0}{2E_s}\frac{N}{K_P} \tag{25}$$

Comparison of (25) and (22) indicates that our ad hoc algorithm (16) yields the minimum possible (over all unbiased estimates) noise contribution to the MSE (assuming that the linearization of the observation model is valid).

5. NUMERICAL RESULTS

In this section we assess the performance of the proposed technique in terms of the MSE of the phase estimate and

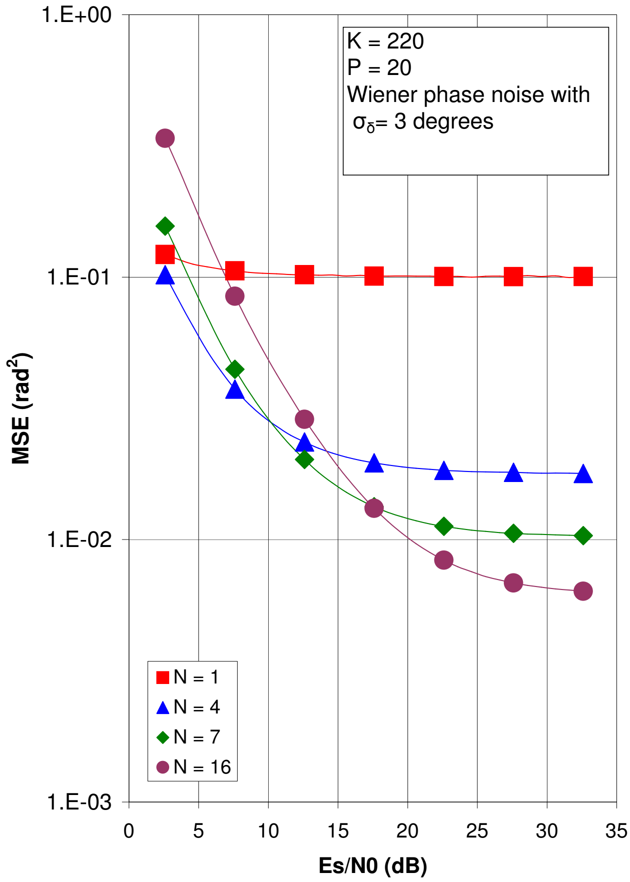

Figure 1: *MSE for Wiener phase noise with* $\sigma_\delta \approx 3°$.

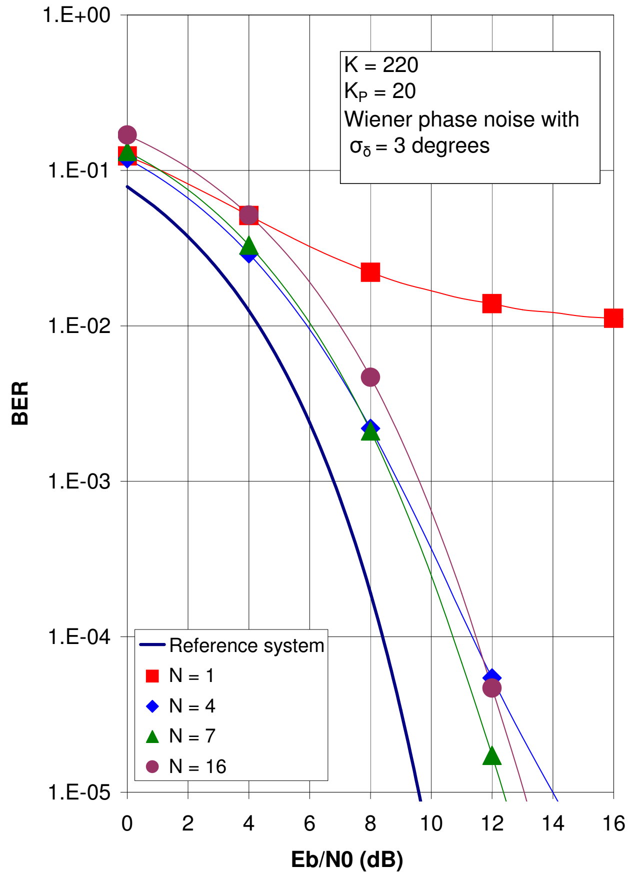

Figure 2: *BER performance of the algorithm in the presence of Wiener phase noise with* $\sigma_\delta \approx 3°$.

the resulting BER degradation by means of computer simulations.

First, we assume transmission of a total block of length $K = 220$ symbols, consisting of 200 uncoded QPSK data symbols and $K_P = 20$ constant-energy pilot symbols that are inserted into the data sequence according to (13).

We consider the presence of Wiener phase noise:

$$\theta(k) = \theta(k-1) + \delta(k) \qquad (26)$$

where $\{\delta(k)\}$ is a sequence of i.i.d. zero-mean Gaussian increments with variance σ_δ^2. From (26) it follows that the variance of the Wiener phase noise increases linearly with the time index k.

Figure 1 shows the MSE of the phase estimate, assuming that the increment of the Wiener phase noise over a symbol interval has variance $\sigma_\delta^2 = 0.0027\, rad^2$ (which corresponds to $\sigma_\delta \approx 3°$). From figure 1 it is obvious that there is a MSE floor in the high-SNR region which can be reduced by increasing the number N of estimated coefficients. We also observe that for low SNR the MSE curve is approximately inversely proportional to $\frac{E_s}{N_0}$, which agrees with (22).

Figure 2 shows the bit-error rate (BER) which is affected by the residual phase noise. The reference curve corresponds to a system with perfect synchronisation and no pilot symbols. We observe that for (very) low SNR values, it is sufficient to estimate only one DCT coefficient. In the high-SNR region, a BER floor occurs which decreases with increasing N, so it becomes beneficial to estimate more than just one DCT coefficient.

Figure 3 shows the BER-degradation at $BER = 10^{-4}$ with respect to the reference system, for a fixed ratio $\eta = \frac{K_P}{K} = 10\%$. The BER degradation is caused not only by the residual phase noise, but also by the loss of power efficiency due to the insertion of pilot symbols into the sequence of information symbols; the latter BER degradation (in dB) amounts to about $0.46\, dB$ for $\eta = 10\%$. The following observation can be made:

- For given block size K, there is an optimum number N_{opt} of DCT coefficients to be estimated, that minimizes the BER degradation. This is consistent with the observation that the MSE of the phase es-

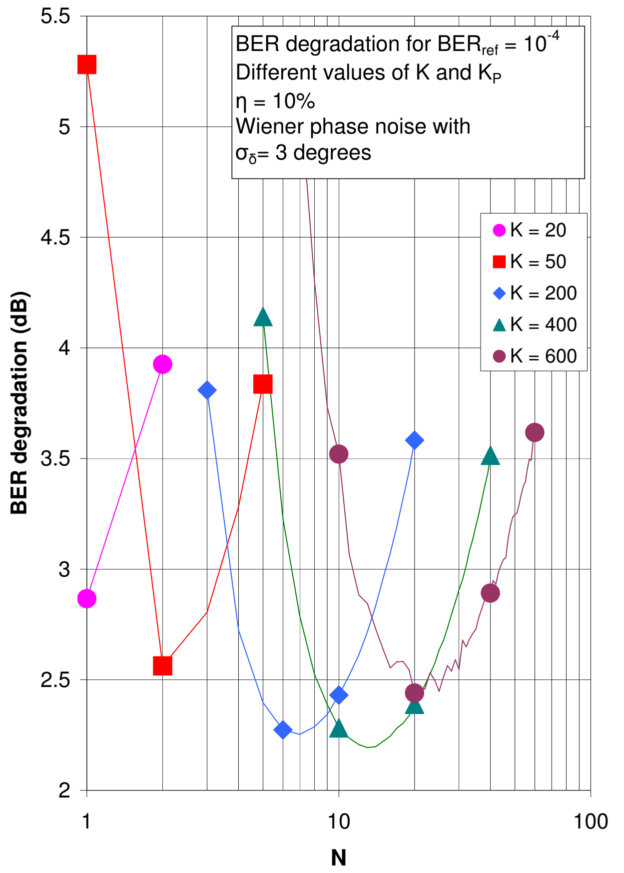

Figure 3: *BER degradation in the presence of Wiener phase noise with* $\sigma_\delta \approx 3°$.

timate can be minimized by a suitable choice of N.

- For very small K, $N_{opt} = 1$. The optimum value N_{opt} increases with increasing K, because more DCT coefficients are needed to model the phase fluctuations when K gets larger. Keeping $N = 1$ yields very large degradations when K increases.

- The BER degradation that corresponds to $N = N_{opt}$ 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 (10) applies, the degradation decreases with increasing K, because more noisy observations of the phase noise are available. However, for too large K the fluctuaction of the Wiener phase noise is so large that linearization is no longer valid and the resulting degradation increases with increasing K.

For the considered scenario, the minimum degradation occurs at $(K_{opt}, N_{opt}) \approx (400, 13)$ and amounts to about $2.3\,dB$. When the actual block size K exceeds K_{opt}, the degradation can be limited by dividing the block in subblocks of at most K_{opt} symbols, and estimating the phase for each subblock separately.

6. CONCLUSIONS AND REMARKS

In this contribution we have considered an ad hoc 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 a 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 numerical results illustrate that the MSE of the phase noise estimate and the associated BER degradation can be minimized by a suitable choice of K and N. For large K, substantial improvement is obtained as compared to the case where only the time-average of the phase is estimated (i.e. $N = 1$).

Acknowledgement

This work is supported by the FWO project G.0047.06 ”Advanced space-time processing techniques for communication through multi-antenna systems in realistic mobile channels”.

7. REFERENCES

[1] 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, pp. 1969–1986, 1998.

[2] 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, pp. 207–218, January 2006.

[3] O. Rousseaux, G. Leus, and M. Moonen, “Estimation and equalization of doubly selective channels using known symbol padding,” *IEEE Transactions on Signal Processing*, vol. 54, pp. 979–990, March 2006.

[4] H. Meyr, M. Moeneclaey, and S. Fechtel, *Digital Communication Receivers : Synchronization, Channel Estimation and Signal Processing*, Wiley Series in Telecommunications and Signal Processing, USA, 1998.

[5] U. Mengali and A.N. D'Andrea, *Synchronization Techniques for Digital Receivers*, Plenum Press, New York, USA, 1997.

[6] L. Benvenuti, L. Giugno, V. Lottici, and M. Luise, "Code-aware carrier phase noise compensation on turbo-coded spectrally-efficient high-order modulations," *8-th Intern. Work. on Signal Processing for Space Commun.*, pp. 177–184, September 2003.

[7] 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.