ON BLIND CHANNEL IDENTIFICATION FOR DS/CDMA COMMUNICATIONS *

Urbashi Mitra
Dept. of Electrical Engineering
The Ohio State University
Columbus, OH 43210, USA
ubli@ee.eng.ohio-state.edu

Dirk Slock
Mobile Communications Group
EURECOM Institute
F-06904 Sophia Antipolis FRANCE
slock@eurecom.fr

Abstract - Blind channel identification techniques based on second order statistics are investigated for multi-user systems employing Direct-Sequence/Code-Division Multiple-Access (DS/CDMA). The methods exploit knowledge of the spreading code of the user of interest via matched filtering, as well as properties of spreading codes. Two schemes are presented: one for deterministic short sequence DS/CDMA and a method appropriate for randomized long sequence DS/CDMA. The matched filter based techniques offer reduced complexity *vis á vis* methods which operate on the received signal directly. In fact, it can be shown that as the number of matched filters increases, the proposed technique approaches a previously proposed method.

1 INTRODUCTION

In this work, we explore blind channel identification schemes for multi-user systems employing Direct-Sequence/Code-Division Multiple-Access (DS/CDMA) based on second order statistics. The search for blind identification methods is motivated by the need for equalizers to have channel knowledge and by the desire to have training signal independent schemes. The classical DS/CDMA receiver for a multipath channel is the RAKE receiver which is simply a receiver matched to the global channel response. This global channel response is the convolution of the spreading code of the desired user and its channel impulse response. We shall explore methods suitable both for deterministic short spreading codes as well as randomized long codes which have been proposed for IS-95.

Due to the spreading operation, DS/CDMA signals lend themselves naturally to a category of blind channel identification techniques that have become recently popular (see *e.g.* [6]). While it is possible to estimate the global channel response (see *e.g.* [3, 10]), we shall focus on schemes which estimate the pure channel only (*e.g.* [9]). Exploiting the knowledge of pulse shape for narrowband systems has been considered in, for example, [2, 7]. The basis of our methods will be the direct exploitation of the spreading sequence of the user of interest via matched filtering. For delay spreads that are much smaller than the symbol period, a moderate number of matched filter outputs form a set of statistics suitable for performing blind channel identification This allows for an often significant reduction in computational complexity.

*This research was supported in part by the NSF under an International Post-doctoral Fellowship and under Grant NCR 96-24375.

This paper is organized as follows. Section 2 presents the signal model and the assumptions employed. In the sequel, Section 3, the subspace identification scheme for deterministic short codes will be described. Section 4 presents the technique for blind channel identification for systems with randomized deterministic codes. Numerical results are presented in Section 5 and concluding remarks are provided in Section 6.

2 BASIC SIGNAL MODEL

We shall presume a coherent system where the active users transmit DS/CDMA signals with binary valued spreading signals and rectangular pulse shapes. The signal is transmitted over a multi-path channel. Coarse synchronization (accuracy to within half a chip) is assumed for the user of interest. Possible synchronization schemes include [8, 9, 10].

For brevity, we directly present the baseband representation of the received signal for a K user system, after chip-matched filtering and sampling at the chip-rate:

$$x_k(n) = \sum_{k=1}^{K} \sum_{i=0}^{N+M-2} A_k (s_k * h_k)(i - \tau_k) b_k \left(\left\lceil \frac{n-i}{N} \right\rceil \right) + v(n)$$

where $b_k(i)$ is the transmitted symbol for user k at time i and $\lceil \cdot \rceil$ is the ceiling operation which rounds up to the next largest integer. The received amplitude for user k is A_k. The sequence $s_k(n)$ is the spreading waveform for user k; this sequence is of length N for the case of deterministic spreading codes. For the case of randomized codes, the processing gain is also assumed to be N, however the code changes from symbol to symbol. Note that symbol duration is T and thus $T = NT_c$, where T_c is the chip duration. The sampled channel impulse response for user k is denoted by $h_k(n)$ and is of length M for each user's channel. This assumption is merely for facility of description and the techniques explored herein are not predicated on this assumption. The convolution operator is denoted by $*$. We shall assume that the channel is of finite length and that the multi-path spread is much less than a symbol interval. The delay for user k is denoted by τ_k. The additive Gaussian noise is given by $v(n)$ and is an independent sequence with variance σ_v^2. It is assumed that the data signal and the noise process are mutually independent.

For notational clarity, we begin by providing an expression for the observation vector under the assumption of a synchronized system ($\tau_k = \tau \quad \forall \quad k$). The extension to asynchronous interferers is straightforward once the relevant matrices have been defined. We can express an observation vector containing a whole data symbols (and two partial bits), resulting in a vector of length $P = aN + M - 1$ samples as,

$$\mathbf{x}(n) = \sum_{k=1}^{K} A_k \mathbf{S}_k \mathbf{H}_k \mathbf{b}(n) + \mathbf{v}(n), \qquad (1)$$

The channel matrix for user k of dimension $(a+2)M \times a+2$, is given by

$$\mathbf{H}_k = \mathbf{h}_k \otimes \mathbf{I}_{a+2},$$

where $\otimes$ denotes the Kronecker product operator and $\mathbf{I}_{a+2}$ is the identity matrix of dimension $a+2$. We shall use the notation $\mathcal{S}(\mathbf{x}, M)$ to denote the Sylvester matrix of width M for vector $\mathbf{x}$. This is equivalent to the convolution matrix for the convolution of vector $\mathbf{x}$ with another vector of length M. We define $\mathbf{S}^1_{k,M}$ to be the first N rows of $\mathcal{S}(\mathbf{s}_k, M)$ and $\mathbf{S}^2_{k,M}$ to be the last $M-1$ rows of $\mathcal{S}(\mathbf{s}_k, M)$. Then, the spreading code matrix for user k, of dimension $aN + M - 1 \times (a+2)M$, is given by,

$$\mathbf{S}_k = \begin{pmatrix} \mathbf{S}^2_{k,M} & \mathbf{S}^1_{k,M} & & & & \mathbf{0} \\ & \mathbf{S}^2_{k,M} & \mathbf{S}^1_{k,M} & & & \\ & & \mathbf{S}^2_{k,M} & \mathbf{S}^1_{k,M} & & \\ & \mathbf{0} & & \ddots & \ddots & \\ & & & & \mathbf{S}^2_{k,M} & \tilde{\mathbf{S}}^1_{k,M} \end{pmatrix}$$

The matrix $\tilde{\mathbf{S}}^1_{k,M}$ is composed of the $M-1$ first rows of $\mathbf{S}^1_{k,M}$.

In order to consider asynchronous systems, we can denote the global channel matrix as $\mathbf{G}_k = \mathbf{S}_k H_k$. Let $\tau_k = d_k T_c$ be the delay of interfering user k with respect to user 1, we assume that $\tau_k \geq 0$ and that $d_k \in \mathcal{Z}^+$. The global channel matrices for the interfering users are constructed by forming the matrix $\mathbf{G}^s_k$ as above, however this matrix is of dimension $(a+1)N + M - 1 \times (a+3)$. The channel matrix $\mathbf{G}_k$ is formed by removing the first $N - d_i$ rows, the last d_i rows from $\mathbf{G}^s_k$, and the last column. Thus the delays of the interfering users are incorporated into the signal description.

The two identification schemes presented in this work, manipulate matched filter outputs rather than the received signal $\mathbf{x}(n)$ directly. Given L matched filters per received symbol, the matched filter observation vector, $\mathbf{y}(n)$ which is of length aL, is given by

$$\begin{aligned} \mathbf{y}(n) &= \mathbf{S}^m_k(L)\mathbf{x}(n) \qquad (2) \\ &= \mathbf{S}^m_k(L)\left(\sum_{k=1}^{K} \mathbf{S}_k \mathbf{H}_k \mathbf{b}_k(n)\right) + \mathbf{S}^m_k(L)\mathbf{v}(n), \end{aligned}$$

where the matched filtering matrix $\mathbf{S}^m_k(L)$ is of dimension $(a-1)L + M \times aN + M - 1$,

$$\mathbf{S}^{m\,T}_k(L) = \begin{pmatrix} \mathbf{S}^1_{k,L} & \mathbf{0} & & & & \\ \mathbf{S}^2_{k,L} & \mathbf{S}^1_{k,L} & & & & \\ \mathbf{0} & \mathbf{S}^2_{k,L} & \mathbf{S}^1_{k,L} & & & \\ & & \ddots & \ddots & & \\ & & & & \mathbf{S}^1_{k,L} & \mathbf{0} \\ & & & & \tilde{\mathbf{S}}^2_{k,L} & \tilde{\mathbf{S}}^1_{k,L} \end{pmatrix}, \qquad (3)$$

where $\mathbf{S}^1_{k,L}$ and $\mathbf{S}^2_{k,L}$ are the first N and last $L-1$ rows of $\mathcal{S}(\mathbf{s}_k, L)$ respectively. Further, $\tilde{\mathbf{S}}^2_{k,L}$ and $\tilde{\mathbf{S}}^1_{k,L}$ contain the first $N - L - 1$ rows of $\mathbf{S}^2_{k,L}$ and $\mathbf{S}^1_{k,L}$. Thus the total number of matched filters applied will be $(a-1)L + M-$ this will be the dimension of the matched filter vector.

We next describe two identification schemes tailored to DS/CDMA signals. The first method is appropriate for short deterministic codes and the second method is designed for long randomized codes.

3 DETERMINISTIC CODES

We combine the identification method of [6] with the pulse shaping techniques of [2, 7]. We note that for the non-matched filtered case, a similar combination of techniques was developed in [9]. We form the autocorrelation matrix of the received matched filter vector and note the following subspace representation,

$$\mathbf{R}^m_y = \mathbf{E}\left\{\mathbf{y}^m \mathbf{y}^{m\,H}\right\} = \mathbf{P}_s \Lambda_s \mathbf{P}^H_s + \mathbf{P}_n \Lambda_n \mathbf{P}^H_n,$$

The columns of $\mathbf{P}_s$ span the signal subspace and the columns of $\mathbf{P}_n$ span the noise subspace. Λ_s and Λ_n are diagonal matrices containing the corresponding eigenvalues of $\mathbf{R}^m_y$. We shall denote the column vectors (eigenvectors of $\mathbf{R}^m_y$) of $\mathbf{P}_n$ as $\mathbf{e}^m_i$. The identification technique exploits the fact that $\mathbf{G}_k \perp \mathbf{P}_n$ [6]. Thus an estimate for the channel vector, $\hat{\mathbf{h}}_k$ can be found,

$$\hat{\mathbf{h}}_k = \arg\min \|\mathbf{P}_n \mathbf{G}_k\|^2_F. \qquad (4)$$

To ensure the existence of a noise subspace, we require $(a-1)L + M > K(a+2)$. This is because the signal space is of dimension $K(a+2)$. Thus, the number of matched filters must be strictly greater than the number of users, and potentially even greater than K. Note that this in contrast to the single user case [5] where $L = M$ is sufficient.

We can form the effective global channel matrix $\mathbf{F}_k$ which is of dimension $aL \times a+2$,

$$\begin{aligned} \mathbf{F}_k &= \mathbf{S}^m_k(L)\mathbf{S}_k\mathbf{H}_k \\ &= \begin{pmatrix} \mathbf{f}_1 & \mathbf{f}_2 & \mathbf{f}_3 & \mathbf{0} & & \mathbf{0} \\ \mathbf{0} & \mathbf{f}_1 & \mathbf{f}_2 & \mathbf{f}_3 & & \mathbf{0} \\ & & \ddots & & \ddots & \vdots \\ \mathbf{0} & & \mathbf{0} & \mathbf{f}_1 & \mathbf{f}_2 & \mathbf{f}_3 \end{pmatrix}, \end{aligned}$$

where each vector $\mathbf{f}_j$ ($j = 1, 2, 3$) is of length L. Thus the matched filtering operation has transformed our N channel system into an L channel system where the effective channel is described by a coefficient vector of length three.

The noise process is given by $\mathbf{S}^m_k(L)\mathbf{v}(n)$ and has covariance matrix $\mathbf{R}^m_v = \sigma^2_n \mathbf{S}^m_k(L)\mathbf{S}^{m\,T}_k(L)$. Thus we shall pre-whiten the matched filter output $\mathbf{y}(n)$ by $\mathbf{W} = (\mathbf{R}^m_v)^{-\frac{1}{2}}$ as suggested in [6]. We denote the whitened observation vector as, $\mathbf{y}^m(n) = \mathbf{W}\mathbf{y}(n)$.

Let $\mathbf{f} = \left[\mathbf{f}^H_1, \mathbf{f}^H_2, \mathbf{f}^H_3\right]^H$. It can be shown that there exist matrices $\mathbf{E}^m$ and $Q(\rho)$ such that

$$\mathbf{e}^{m\,H}\mathbf{F} = \mathbf{f}^H \mathbf{E}^m \quad \text{and} \quad \mathbf{f} = \mathbf{Q}(\rho)\mathbf{h}_k \qquad (5)$$

where the auto-correlation function, ρ, is given by

$$\rho(n) = \sum_{k=0}^{N-1} s(k)s(k-n).$$

The matrix $\mathbf{E}_i^m$ is defined as in (5) where $\mathbf{e}_i^m$ is an eigenvector of the signal subspace of the whitened matched filter observation vector $\mathbf{y}^m(n)$. The desired channel vector is found by minimization of the cost function below which is equivalent to the minimization of (4),

$$q_{MF}(\mathbf{h}) = \mathbf{h}^H \underbrace{\left(\sum_{i=0}^{aL-K(a+2)-1} \mathbf{Q}(\rho)^H \mathbf{E}_i^m \mathbf{E}_i^{m\,H} \mathbf{Q}(\rho) \right)}_{\mathbf{Q}^*} \mathbf{h}.$$

Thus the estimated channel is the eigenvector of $\mathbf{Q}^*$ corresponding to its minimum eigenvalue.

A complexity reduction is induced through the reduced dimension of $\mathbf{y}(n)$ versus the original received vector $\mathbf{x}(n)$. Whereas the length of $\mathbf{x}(n)$ is directly a function of the spreading code length (which can be as large as 128), the length of $\mathbf{y}(n)$ is independent of N. Thus our algorithm has reduced complexity *vis á vis* counterpart algorithms [10, 9]. There is some loss in theoretic performance if one compares the pulse shaping [9] and matched filter schemes - this is due to the fact that while the matched filter outputs are suitable for estimation, they are not sufficient in this asynchronous case. It has been empirically observed that performance improves as the dimension of the noise subspace increases [1]. Thus, as the number of matched filters increases, the channel estimation error will decrease, the limit being the performance achievable by the pulse shaping performance [9]. The matched filter scheme allows the system designer to trade off between complexity and performance.

4 RANDOMIZED CODES

In the current proposed DS/CDMA standard, IS-95, long randomized spreading codes are suggested. These spreading codes have periods that are much longer than a symbol duration. Thus, in this section, we develop an identification scheme appropriate for such a system. In particular, we model the spreading codes of all active users as random sequences with chip values that are equal to $\frac{\pm 1}{\sqrt{N}}$ with equal probability. Thus, the autocorrelation function of the desired user's sequence is a Dirac delta function. Similarly, the cross-correlation function between two distinct users is approximately the zero function. To begin, we consider an observation vector $\mathbf{x}(n)$ of length $N + M - 1$, that is a vector corresponding to a single whole symbol and two partially received symbols.

We form $\mathbf{y}(n)$ as in 2 with $L = M$. With our assumptions on the spreading sequences, which can be considered to be time-varying,

$$\mathbf{R}_y = \mathbf{E}\left\{\mathbf{y}(n)\mathbf{y}(n)^H\right\} = \mathbf{h}_k\mathbf{h}_k^H + \mathbf{R}_I + \sigma^2\mathbf{R}_v^m$$

Where $\mathbf{R}_v^m = \mathbf{S}_k^m(L)\mathbf{S}_k^{m\,T}(L)$. Note that this matrix is Toeplitz. The contribution due to the desired user is simply a rank one matrix. The interferers' contribution at the output of the matched filter for the desired user is cyclostationary with period equal to the chip period and hence stationary when sampled at the chip rate. Therefore this interference matrix, $\mathbf{R}_I$, is Toeplitz and Hermitian. Using statistical arguments, one can argue that as the number of asynchronous interferers increases and as the processing gain increases, this Toeplitz matrix actually converges to a diagonal matrix.

We define $\mathbf{h}_k^- = [h_k(1), \cdots, h_k(M-1)]^T]$ and $\mathbf{h}_k^+ = [h_k(2), \cdots, h_k(M)]^T]$. The Toeplitz contribution of the channel noise and interference can be removed by considering the following operation:

$$\begin{aligned}\mathbf{R}_h &= \mathbf{R}_y(2:M, 2:M) - \mathbf{R}_y(1:M-1, 1:M-1) \\ &= \mathbf{h}_k^+\mathbf{h}_k^{+\,H} - \mathbf{h}_k^-\mathbf{h}_k^{-\,H},\end{aligned}$$

where the matrix notation $\mathbf{B}(i:j, i:j)$ corresponds to the submatrix of $\mathbf{B}$ formed by the appropriately truncated i through j columns and i through j rows. An eigendecomposition of $\mathbf{R}_h$ is given by

$$\mathbf{R}_h = \sum_{i=1}^{M-1} \lambda_i \mathbf{v}_i \mathbf{v}_i^H. \quad (6)$$

Ordering the eigenvalues such that $\lambda_1 \geq \lambda_2, \cdots, \geq \lambda_{M-1}$, only λ_1 and λ_{M-1} should be non-zero with $\lambda_1 > 0$ and $\lambda_{M-1} < 0$. In practice, from the sampled average correlation matrix, we form an estimate of $\mathbf{R}_h$ and compute $\mathbf{v}_1$ and $\mathbf{v}_{M-1}$.

We introduce the following subspace fitting problem

$$\min_{\mathbf{h},\mathbf{T}} ||\mathbf{H} - \mathbf{V}\,\mathbf{T}||_F^2,$$

where $\mathbf{H} = \left[\mathbf{h}_k^- \mathbf{h}_k^+\right]$ and $\mathbf{V} = [\mathbf{v}_1 \ \mathbf{v}_{M-1}]$. Minimization with respect to $\mathbf{T}$ yields

$$\min_{||\mathbf{h}||=1} \text{ trace } \mathbf{H}^H \mathbf{P}_V^\perp \mathbf{H} = \min_{||\mathbf{h}||=1} \mathbf{h}^H \mathbf{Q} \mathbf{h}.$$

The matrix $\mathbf{P}_V^\perp$ is a null-matrix for $\mathbf{V}$. And thus the solution is the eigenvector of $\mathbf{Q}$ corresonding to its minimum eigenvalue, where $\mathbf{Q}$ is defined as

$$\mathbf{Q} = \begin{bmatrix} \mathbf{P}_V^\perp & \begin{matrix} 0 \\ \vdots \end{matrix} \\ 0 \ \cdots & 0 \end{bmatrix} + \begin{bmatrix} 0 & \cdots \ 0 \\ \begin{matrix} \vdots \\ 0 \end{matrix} & \mathbf{P}_V^\perp \end{bmatrix}.$$

In practice, two modifications are employed. The first lies in the recognition that $\mathbf{S}_k^m(L)\mathbf{S}_k \neq I$. Thus the matched filtering is augmented by the the linear transformation $\mathbf{W} = \left(\mathbf{R}_v^k\right)^{-1}$. That is ,$\mathbf{y}^T(n) = \mathbf{W}\mathbf{S}_k^m(L)\mathbf{x}(n)$. The second change is to incorporate longer data lengths. If a complete and two partial data symbols are present in the vector $\mathbf{x}(n)$ then $\mathbf{R}_h$ in (6) will contain $2a$ non-zero eigenvalues. This data *smoothing* will result in improved performance.

The recognition of the Toeplitz interference matrix was also made in [4]. However instead of working with the Toeplitz displacement of $\mathbf{R}_y^T = \mathbf{E}\left\{\mathbf{y}^T(n)\mathbf{y}^{T\,H}(n)\right\}$, the difference of the covariance matrices before and after code

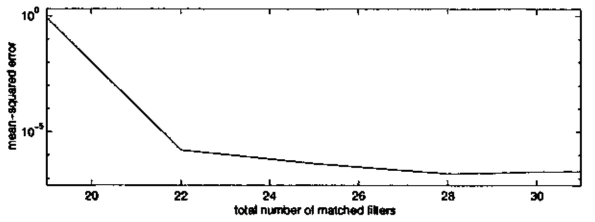

Figure 1: Channel estimation error versus total number of matched filters for the deterministic code case.

matched filtering is manipulated. Under ideal spreading code assumptions, it can be shown that this difference is in principle $\mathbf{h}_k\mathbf{h}_k^H$. While the work in [4] focused on equalization, there is the allusion to a corresponding identification scheme. We shall see that the proposed algorithm herein outperforms that of [4].

5 NUMERICAL RESULTS

Figure 1 shows how performance of the deterministic short code scheme is affected by the number of matched filters. In this experiment, length 31 Gold codes were used in an eight user, asynchronous system with SNR = 20 dB. The channel length was 3 and 400 Monte Carlo runs were conducted. As can be seen, increasing the number of filters improves performance. The poor performance exhibited for 19 filters is due to the lack of a noise subspace.

Figure 2 shows performance for the long randomized code algorithm (TD algorithm). For both experiments depicted, 8 user asynchronous systems are under consideration with SNR=20dB. In addition, the channel length is 5; 300 Monte Carlo runs were conducted. Figure 2(a) shows how performance varies as the number of received symbols increases for $N = 30$. Note that performance is not monotonic in the number of symbols. This is due to the fact that while increasing a increases desired signal energy, it also increases the amount of interference energy in the received signal. In Figure 2(b), we compare performance of the TD algorithm to that noted in [4], denoted the LZ algorithm. Performance is compared as the processing gain increases. Note that as N increases, the spreading code correlation functions become more ideal. It is clear that the TD algorithm outperforms the LZ algorithm.

6 CONCLUSIONS

Two blind identification schemes for multi-path channels for DS/CDMA signals have been considered. The method for deterministic short codes is based on determining noise and signal subspaces using the matched signal correlation matrix. It can be shown that explicit use of the spreading code enables the system designer to trade-off between complexity and performance. The randomized long code identification technique exploits the statistical properties of aperiodic spreading codes. The matched filtering operation renders the desired user's signal stationary at the symbol rate, while the multi-user interference remains stationary at the chip rate. Thus, its contribution can be eliminated via a Toeplitz displacement. The proposed scheme outperforms that of [4], which exploits similar properties of DS/CDMA signals.

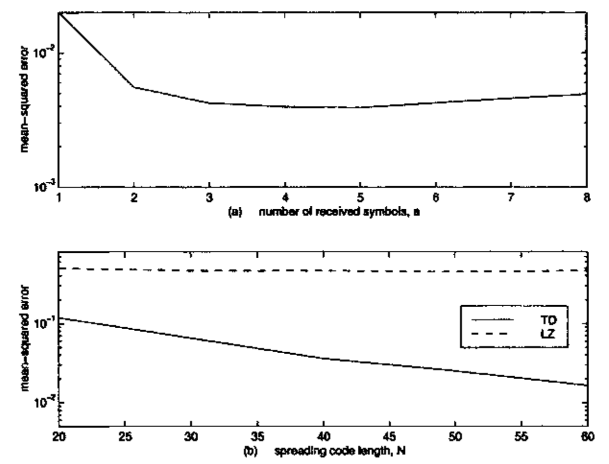

Figure 2: Channel estimation error (a) versus the number of received symbols a; and (b) versus the processing gain N with comparison to the LZ algorithm for the randomized code case.

REFERENCES

[1] E. Aktas and U. Mitra. Reduced complexity blind identification algorithm for multi-user DS/CDMA signals. In *Proceedings of ICC, Atlanta, GA*, June 1998.

[2] Z. Ding and Z. Mao, "Knowledge Based Identification of Fractionally Sampled Channels," in *ICASSP95*.

[3] H. Liu and G. Xu, "Closed-Form Blind Symbol Estimation in Digital Communications", *IEEE Trans. Signal Proc.*, pp. 2714-2723, vol. 43, no. 11, Nov. 1995.

[4] H. Liu and M. Zoltowski. Blind equalization in antenna array CDMA systems. *IEEE Trans. on Signal Proc.*, vol. 45, no. 1, pp. 161-172, Jan. 1997.

[5] U. Mitra, D. Slock, and C. Escudero. Blind equalization techniques for CDMA communications. In *Proceedings of the 31sh Annual Conference on Information Sciences and Systems, Johns Hopkins, MD*,pp. 696–701, March 1997.

[6] E. Moulines, P. Duhamel, J.-F. Cardoso, and S. Mayrargue, "Subspace methods for the blind identification of multichannel FIR filters," *IEEE Trans. on Signal Proc.*, vol. 43, no. 2, pp. 516–525, February 1995.

[7] S.V. Schell, D.L. Smith and W.A. Gardner, "Blind Channel Identification Using 2nd-Order Cyclostationary Statistics," in *Proc. Eusipco94*, Edinburgh, Scotland, 13–16 Sept. 1994, pp. 716–719.

[8] E. Ström, S. Parkvall, S. Miller, and B. Ottersten, "Propagation delay estimation in asynchronous direct-sequence code-division multiple access systems," *IEEE Trans. Comm.*, vol. 44, no. 1, pp. 84–93, Jan. 1996.

[9] M. Torlak and G. Xu, "Blind multi-user channel estimation in asynchronous CDMA systems," *IEEE Trans. Signal Proc.*, vol. 45, no. 1, pp. 137–147, Jan. 1997.

[10] M. K. Tsatsanis and G. B. Giannakis, "Blind estimation of direct sequence spread spectrum signals in multipath," *IEEE Trans. Signal Proc.*, vol. 45, no. 5, pp. 1241-1252, May 1997.