PARAMETER ESTIMATION IN A SPACE TIME BIT-INTERLEAVED CODED MODULATION SCHEME FOR DS-CDMA

M. Guenach, F. Simoens and M. Moeneclaey

Department of Telecommunications and Information Processing
Ghent University, St.-Pietersnieuwstraat 41, B-9000 GENT, Belgium
E-mail: {guenach,fsimoens,mm}@telin.UGent.be

ABSTRACT

This paper deals with space-time detection and channel parameter estimation in a bit-interleaved coded modulation (BICM) scheme for asynchronous Direct-Sequence Code Division Multiple Access (DS-CDMA) transmission over frequency selective channels. Using a pilot sequence to obtain sufficiently accurate channel estimates for this system requires an unacceptably large number of pilot symbols. In this contribution, we consider several code-aided estimation schemes which can be incorporated in an iterative Space-Time (ST) turbo detection scheme. We point out that the EM algorithm, which has recently been addressed as a convenient tool for code-aided estimation, has some serious drawbacks because of the large number of parameters to be estimated. These can be resolved using the Space Alternating Generalized Expectation Maximization (SAGE) algorithm. We show through computer simulations that the proposed low complexity ST receiver with SAGE estimation considerably outperforms conventional estimation schemes.

1. INTRODUCTION

Direct-Sequence code-division multiple-access systems have the ability to accommodate multiple users in multi-path fading environments. Recently developed coding and detection schemes allow a reliable transmission at very high data rates. However, in order to exploit the channel capacity and diversity, the receiver requires accurate parameter estimates.

In recent years, a lot of effort has been put in developing powerful *channel* estimation algorithms. Most of them are based on the Expectation-Maximization (EM) algorithm and have been shown to have excellent performance in a wide variety of scenarios. Estimation algorithms using *pilot symbols* were investigated in [1, 2] for synchronous and asynchronous systems, respectively. A flat Rayleigh fading channel was assumed in [1] with an EM-based estimator. In [2], a different approach is taken: based on an extension of the EM algorithm, the SAGE algorithm [3] was applied for the estimation of channel parameters of a static multipath channel. Both papers report excellent performance, though the estimator based on the SAGE algorithm generally has faster convergence and is less computationally demanding. In a multi-antenna context, we mention [4], also applying the SAGE algorithm in a data-aided (DA) context.

On the other hand, the literature on *code-aided* estimation for DS-CDMA and in particular *frequency offset* estimation is quite scarce. A general framework for code-aided EM estimation of channel and frequency offset for coded signals was proposed in [5], but only applied to a simple SISO system in AWGN channel.

In this paper, we focus on a coded asynchronous DS-CDMA configuration with multiple receive antennas. We propose a detector that iterates between space-time detection and estimation. Parameter estimation is developed starting from the ML criterion which is solved by the EM algorithm in an iterative manner. The latter turns out to be still complex, and therefore an efficient estimator by means of the SAGE algorithm is proposed. Computer simulations verify the gain achieved by exploiting information from the coded symbols compared to data-aided estimation.

2. SYSTEM MODEL

At the transmitter side, a block of convolutionally encoded bits are interleaved and grouped into sub-blocks of q bits. The resulting block of coded bits is mapped to a sequence of M_d symbols (denoted by vector $\mathbf{d}$), belonging to a 2^q-point complex constellation Ω. Multiplexing with M_t pilot symbols $\mathbf{p}$ yields the sequence $[d[-M_t], .., d[M_d-1]]$. The complex symbols $d[m]$ are shaped by a normalized spreading waveform $u(t)$ that contains a spreading code $a(i)$ of length N_c: $u_k(t) = (1/\sqrt{N_c})\sum_{i=0}^{N_c-1} a_k(i)\pi(t-iT_c)$. We will denote by T_d, T_c and N_c, the symbol period, the chip period and the spreading factor, respectively. Note that $T_d = N_cT_c$. We consider a system where the receiver is equipped with an array of n_R antennas. The resulting signal propagates through a multi-path fading channel, with L paths. The channel impulse response, seen by the p-th receive an-

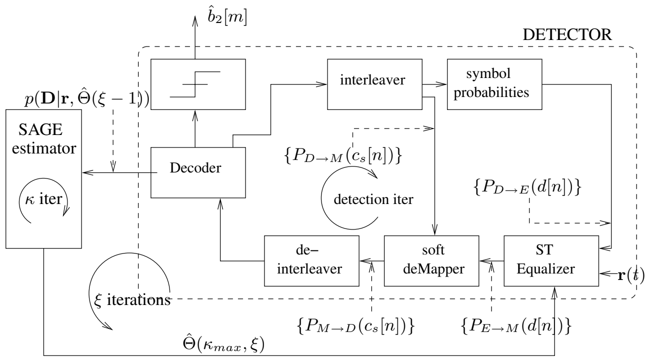

Fig. 1. Receiver structure.

tenna is given by

$$g^p(t) = \sum_{l=1}^{L} g_l^p \delta(t-\tau_l^p) \tag{1}$$

where g_l^p and τ_l^p are the complex gain and path delay of the l-th propagation path, respectively. We further assume that each path invokes a different frequency offset F_l^p caused by an oscillator mismatch and/or Doppler shift. Because of the latter we consider a model with different frequency offsets for the different paths. Θ will represent the parameter vector defined as $\Theta = \left[\Theta^1, \ldots, \Theta^{n_R}\right]$ where $\Theta^p = [g_1^p, \tau_1^p, F_1^p, \ldots, g_L^p, \tau_L^p, F_L^p]$. The data frames are corrupted by a vector of independent additive white Gaussian noise $\mathbf{n}(t)$ with power spectral density $2N_0$. Hence, the equivalent baseband signal on the different antennas at the receiver is given by the $(n_R \times 1)$ vector:

$$\mathbf{r}(t) \quad = \quad \mathbf{s}(t, \mathbf{d}, \mathbf{p}, \Theta) + \mathbf{n}(t) \tag{2}$$

where the p-th component of $\mathbf{s}(t, \mathbf{d}, \mathbf{P}, \Theta)$ can be written as

$$s^p(t, \mathbf{d}, \mathbf{p}, \Theta^p) = \sum_{m=-M_t}^{M_d-1} d\,[m] \sum_{l=1}^{L} g_l^p u(t - mT_d - \tau_l^p) e^{j2\pi F_l^p t}.$$

The ST turbo-detector, is depicted in Fig. 1. Space-time equalization is performed using a MMSE parallel interference cancellation (PIC) [6]. The detection process is performed in an iterative fashion. Each detection iteration consists of equalization [6], soft demodulation, and soft decoding. For further detail we refer to the system outline in [6,7].

3. CHANNEL PARAMETER ESTIMATION

In this section we will apply the three estimation algorithms (ML,EM,SAGE) to the problem at hand: channel parameter estimation in the presence of unknown data symbols (which we will refer to as nuisance parameters) for DS-CDMA. Our aim is to estimate a parameter vector Θ from an observation $\mathbf{r}$ in the presence of a so-called nuisance parameter $\mathbf{d}$, corresponding to the unknown transmitted data-symbols. We will illustrate the benefit of using the SAGE algorithm in terms of complexity and convergence speed.

3.1. ML estimation

In principle, an estimate of Θ can be obtained by maximizing the likelihood function, averaged over the (uniformly distributed) nuisance parameter $\mathbf{d}$:

$$\hat{\Theta} \quad = \quad \arg\max_{\Theta} \mathrm{E}_{\mathbf{d}}\left[p(\mathbf{r}|\mathbf{d}, \Theta)\right] \tag{3}$$

Unfortunately, straightforward application of the ML estimation procedure has two complexity-related problems: first of all, averaging in (3) is performed over all possible codewords, making ML estimation in coded systems intractable (even for uncoded systems). Secondly, (3) involves a high-dimensional optimization problem ($3 \times n_R \times L$ parameters), which is in practice very hard to solve.

3.2. The EM algorithm

In this section we introduce the EM algorithm as an iterative solution for the estimation problem outlined above. It requires us to define the so-called *complete data* $\mathbf{z}$. Suppose we have somehow obtained an estimate of Θ: $\hat{\Theta}(\xi)$, with ξ denoting the iteration index. We now iterate between the so-called E-step and the M-step. In the E-step, we take the expectation of the log-likelihood function (LLF) of the complete data, given the observed data and the current estimate of Θ:

$$Q\left(\Theta \middle| \hat{\Theta}(\xi)\right) \quad = \quad \mathrm{E}_{\mathbf{z}}\left[\log p\left(\mathbf{z}|\,\Theta\right) \middle| \mathbf{r}; \hat{\Theta}(\xi)\right]. \tag{4}$$

In the M-step, we then maximize the average LLF with respect to Θ:

$$\hat{\Theta}(\xi+1) = \arg\max_{\Theta} Q\left(\Theta \middle| \hat{\Theta}(\xi)\right). \tag{5}$$

In order to achieve convergence to the ML estimate, a fairly good initial estimate of Θ is required.

3.2.1. Conventional EM-algorithm

In correspondence with [5], we select as complete data $\mathbf{z} = [\mathbf{d}, \mathbf{r}]$. It was shown in [5] that the averaging in (4) may be performed based solely on the marginal a posteriori probabilities (APP) of the coded symbols: $P\left(d_k[m] \middle| \mathbf{r}; \hat{\Theta}(\xi)\right)$. These probabilities can be obtained from the detector (in an iterative way). Thus, the complexity of the ML estimation due to the presence of the code has been reduced to an acceptable level. However, the maximization in (5) is still a high-dimensional problem ($3 \times n_R \times L$ parameters).

3.2.2. *Signal Decomposition for EM estimation*

The complexity of the EM algorithm w.r.t. the high dimensional maximization can be resolved by selecting a *different complete data*, based on signal decomposition [8]. We first focus on a single receive antenna. The useful signal w.r.t. the estimation of Θ_l^p at the p-th receive antenna is given by

$$s_l^p\left(t, \mathbf{d}, \mathbf{p}, \Theta_l^p\right) = g_l^p \sum_{m=-M_t}^{M_d-1} d\left[m\right] u\left(t - nT_d - \tau_l^p\right) e^{j2\pi F_l^p t}.$$

If we now define $x_l^p\left(t\right) = s_l^p\left(t, \mathbf{d}, \mathbf{p}, \Theta_l^p\right) + \mathrm{n}_l^p\left(t\right)$, where $\mathrm{n}_l^p\left(t\right)$ is the noise term obtained by decomposing $\mathrm{n}^p\left(t\right)$ [8], then the signal at the p-th receive antenna can be written as the sum of L *independent* contributions: $r^p\left(t\right) = \sum_{l=1}^{L} x_l^p\left(t\right)$.

We denote by $\mathbf{x}_l^p$ the projection of $x_l^p\left(t\right)$, and define $\mathbf{x}^p = \left[\mathbf{x}_1^p, \ldots, \mathbf{x}_L^p\right]$ and $\mathbf{x} = \left[\mathbf{x}^1, \ldots, \mathbf{x}^{n_R}\right]$. In contrast to the previous paragraph, we define the complete data as $\mathbf{z} = \left[\mathbf{x}, \mathbf{d}\right]$. Since the components $\mathbf{x}_l^p$ in $\mathbf{x}$ are mutually independent, the LLF can be decomposed. Some straightforward calculation yields the following E-step:

$$Q\left(\Theta\middle|\,\hat{\Theta}\left(\xi\right)\right) = \sum_{p=1}^{n_R}\sum_{l=1}^{L} Q'\left(\Theta_l^p\middle|\,\hat{\Theta}\left(\xi\right)\right) \quad (6)$$

with $Q'\left(\Theta_l^p\middle|\,\hat{\Theta}\left(\xi\right)\right) = \mathrm{E}_{\mathbf{x},\mathbf{d}}\left[\log p\left(\mathbf{x}_l^p\left|\Theta_l^p, \mathbf{d}\right.\right)\middle|\mathbf{r}; \hat{\Theta}\left(\xi\right)\right]$. Thanks to the decomposition, the M-step is now transformed into $n_R \times L$ parallel 3-dimensional maximization problems:

$$\hat{\Theta}_l^p\left(\xi+1\right) = \arg\max_{\Theta_l^p} Q'\left(\Theta_l^p\middle|\,\hat{\Theta}\left(\xi\right)\right) \quad (7)$$

The update equations, obtained by a tedious but straightforward derivation have the following form:

$$\left(\hat{\tau}_l^p(\xi), \hat{F}_l^p(\xi)\right) = \max_{\tau,F} \psi_l^p\left(\tau, F, \tilde{\mathbf{d}}, \hat{\Theta}^p(\xi-1)\right) \quad (8)$$

$$\hat{g}_l^p(\xi) = \frac{\psi_l^p\left(\widehat{\tau}_l^p(\xi), \widehat{F}_l^p(\xi), \tilde{\mathbf{d}}, \hat{\Theta}^p(\xi-1)\right)}{M_d + M_t} \quad (9)$$

where $\tilde{\mathbf{d}}$ denotes the soft-symbols computed from the APP probabilities of the coded symbols $P\left(d\left[m\right]\middle|\mathbf{r}; \hat{\Theta}(\xi-1)\right)$, provided by the detector. Because of page limitations the expressions for $\psi_l^p\left(.\right)$ are omitted. We see that by selecting a different complete data, the EM algorithm is transformed from one large $3 \times n_R \times L$-dimensional problem into $n_R \times L$ parallel 3-dimensional problems. Hence, the estimation problem is partially decoupled and finally tractable. However, we still encounter a 2-D maximization process (8), and because of the large complete data set, the convergence rate is unsatisfactory [3]. In a final step, we apply the SAGE algorithm to encompass these impediments.

Algorithm 1 SAGE estimation

input: $\hat{\Theta}\left(0, \xi\right) \leftarrow \hat{\Theta}\left(\xi-1\right) \equiv \hat{\Theta}\left(\kappa_{max}, \xi-1\right)$
input: $P\left(d_k\left[m\right]\middle|\mathbf{r}; \hat{\Theta}\left(\xi-1\right)\right) \rightarrow \tilde{\mathbf{d}}$
for $\kappa = 1$ to κ_{max} **do**
  $n \leftarrow (\kappa \mod L+1)$
  for $p = 1$ to n_R **do**
    E-step: compute $\psi_n^p\left(\tau, F, \tilde{\mathbf{d}}, \hat{\Theta}^p(\kappa-1, \xi)\right)$
    M-step: update $\hat{\Theta}_\kappa^p\left(\kappa, \xi\right)$ ((10) - (12)).
  end for
end for
return $\hat{\Theta}\left(\kappa_{max}, \xi\right)$

3.3. The SAGE algorithm

The Space Alternating Generalized Expectation Maximization (SAGE) algorithm can be seen as an extension of the EM algorithm [3]: rather than updating all these parameters at once in (5), we break up the problem in low-dimensional subproblems by conditioning on a subset of the parameters. We then use the EM algorithm to solve the subproblems.

Without going into much detail, we will briefly outline the final estimator resulting from the SAGE algorithm. Strict application of the SAGE algorithm would require an update of the soft-symbols $\tilde{\mathbf{d}}$ at the beginning of each SAGE-iteration. However, since the detector-step (which computes $\tilde{\mathbf{d}}$) is the main computational burden, we will only update the soft-symbols after a fixed number of κ_{max} SAGE-iterations (κ-iterations). Assume we are within iteration ξ between estimator and detector. Instead of updating all the $3 \times n_R \times L$ parameters at once, the SAGE algorithm *iteratively* updates the parameters, one at a time. In the κ-th (SAGE-)iteration, we update $3 \times n_R$ parameters: $\Theta_n^p = \left(\tau_n^p, F_n^p, g_n^p\right)$, $p = 1 \ldots n_R$ with $n = (\kappa \mod L+1)$. We have available: $\tilde{\mathbf{d}}$ and $\hat{\Theta}^p\left(\kappa-1, \xi\right)$. The update-equations for parameters $\Theta^p = \left(\tau_n^p, F_n^p, g_n^p\right)$ are:

$$\hat{\tau}_n^p(\kappa, \xi) = \max_{\tau} \psi_n^p\left(\tau, \widehat{F}_n^p(\kappa-1, \xi), \tilde{\mathbf{d}}, \hat{\Theta}^p(\kappa-1, \xi)\right) \quad (10)$$

$$\hat{F}_n^p(\kappa, \xi) = \max_{F} \psi_n^p\left(\hat{\tau}_n^p(\kappa, \xi), F, \tilde{\mathbf{d}}, \hat{\Theta}^p(\kappa-1, \xi)\right) \quad (11)$$

$$\hat{g}_n^p(\kappa, \xi) = \frac{\psi_n^p\left(\widehat{\tau}_n^p(\kappa, \xi), \widehat{F}_n^p(\kappa, \xi), \tilde{\mathbf{d}}, \hat{\Theta}^p(\kappa-1, \xi)\right)}{M_d + M_t} \quad (12)$$

The remaining parameters are not updated in iteration κ: $\hat{\Theta}_l^p\left(\kappa, \xi\right) = \hat{\Theta}_l^p\left(\kappa-1, \xi\right)$, $l \neq n$. As explained above, the soft-symbols $\tilde{\mathbf{d}}$ are not updated within each κ-iteration for reasons concerning the complexity. The two main differences with the EM algorithm are that the 2-D maximization problem (8) is now decoupled into (10) and (11) and newly obtained parameter updates are used immediately for subsequent κ-iterations, whereas in the EM algorithm newly updated parameters are only applied in the next ξ-iteration. Hence, the complexity is reduced and, the convergence rate increased.

To summarize, we outline the SAGE estimator in Algorithm 1. Suppose we are operating at the start of iteration ξ between detection and estimation. We denote by $\hat{\Theta}(\kappa, \xi)$ the estimate of Θ obtained in the κ-th iteration within the SAGE algorithm, for a fixed ξ.

The algorithm accepts as input the previous estimate of Θ (i.e., $\hat{\Theta}(\xi - 1)$), as well as the APPs of the coded symbols $\tilde{\mathbf{d}}$ based on these estimates.

Note that in the E-step of Algorithm 1, the soft-symbols $\tilde{\mathbf{d}}$ remain unchanged for all κ-iterations. Furthermore we only perform one detection iteration (Fig. 1) for each soft-symbol update (for each ξ-iteration). This means that the estimation is integrated in the detection iterations and causes little overhead (no extra decoding iterations are required only the iterative calculation (κ-iterations) of (10)-(12)).

4. NUMERICAL RESULTS

In this section we will provide numerical results to evaluate the performance of the proposed iterative multiuser receiver. We have carried out computer simulations for a system, using a rate $R = 1/2$ convolutional code with 8-PSK signaling. Frames consist of 120 coded data symbols and 10 training symbols. We consider randomly chosen but fixed path delays and channel gains.

Fig. 2 shows plots for different estimation scenarios of the SAGE algorithm. The initial channel estimate is obtained by the DA SAGE algorithm [1, 4] (using 10 pilot symbols), whereas the initial frequency offset is set equal to zero. The actual frequency offset is chosen the same for all paths and is randomly distributed between $\pm 0.1 \frac{1}{M_d + M_t}$. Compared to the DA estimation ($\xi = 0$), a substantial performance gain is observed (up to $3dB$ for the joint estimation of all parameters).

5. CONCLUSION

We have investigated a DS-CDMA receiver with BICM, performing iterative space-time detection and channel parameter estimation. The estimator operates by accepting soft information from the detector. We showed how the low-complexity SAGE algorithm can be applied for this problem. The computational overhead of the estimator is minimized by embedding the estimation stages in the detection stages so that a form of *joint* detection and estimation is performed.

The performance of the proposed algorithm is evaluated in terms of BER. It turns out that the SAGE algorithm, exploiting information from all the data symbols significantly outperforms the standard iterative DA SAGE algorithm.

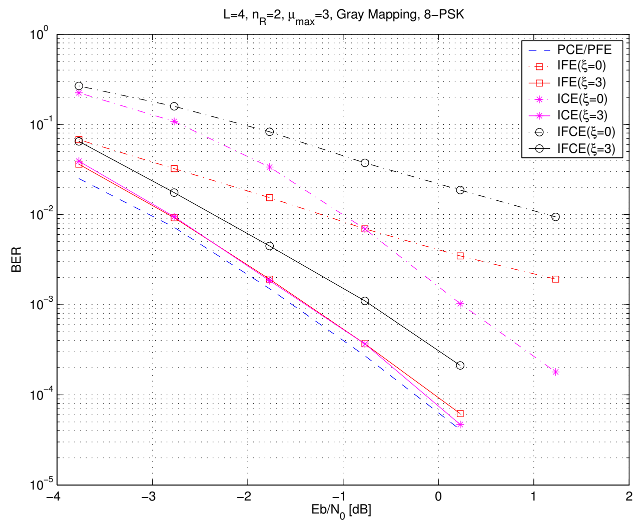

Fig. 2. BER-performance for different estimation scenarios. (PCE/PFE: perfect synchronization, IFE: estimation of frequency offset only (channel known), ICE: estimation of channel only (frequency known), IFCE: both frequency and channel estimation)

6. REFERENCES

[1] A. Kocian and B.H. Fleury. "EM-based joint data detection and channel estimation of DS-CDMA signals". *IEEE Trans. Comm.*, 51(10):1709–1720, October 2003.

[2] A.A. D'Amico, U. Mengali and M. Morelli. "Channel estimation for the Uplink of a DS-CDMA system". *IEEE Trans. on Wireless Comm.*, 2(6):1132–1137, November 2003.

[3] J.A. Fessler and A.O. Hero. "Space alternating generalized expectation maximization algorithm". *IEEE Trans. Comm.*, 42(10):pp.2664–2677, October 1994.

[4] B.H. Fleury, M. Tschudin, R. Heddergott, D. Dahlhaus and K.I. Pedersen. "Channel parameter estimation in mobile radio environments using the SAGE algorithm". *IEEE J. Sel. A. Comm.*, 17(3):pp.434–450, March 1999.

[5] N. Noels, C. Herzet, A. Dejonghe, V. Lottici, H. Steendam, M. Moeneclaey. M. Luise and L. Vandendorpe. "Turbo-synchronization: an EM algorithm interpretation". In *IEEE ICC'03*, 2003.

[6] X. Wang and H. V. Poor. "Iterative (turbo) soft interference cancellation and decoding for coded CDMA". *IEEE Trans. Comm.*, 47(7):pp.1046–1061, July 1999.

[7] M. Guenach, H. Wymeersch and M. Moeneclaey. "On DS-CDMA uplink channel parameter estimation for bit-interleaved-coded modulation". In *Proc. SPWC*, London, UK, June 2004.

[8] M. Feder and E. Weinstein. "Parameter estimation of superimposed signals using the EM algorithm". *IEEE Trans. on Acoustics, Speech and Signal Processing*, 36:477–489, April 1988.