A Semi-Blind Channel Estimation Technique based on Second-Order Blind Method for CDMA Systems

Samson Lasaulce, Philippe Loubaton, Eric Moulines

A Semi-Blind Channel Estimation Technique based on Second-Order Blind Method for CDMA Systems

Samson LASAULCE*, Philippe LOUBATON , Éric MOULINES

SP EDICS: 3-CEQU, 3-ACCS

Abstract

This paper aims at studying a semi-blind channel estimation scheme based on the subspace method or a carefully weighted linear prediction approach. The corresponding (composite) semi-blind cost functions result from a linear combination of the training-based cost function and a blind cost function. For each blind method, we show how to calculate the asymptotic estimation error. Therefore, by minimizing this error, we can properly tune the K-dimensional regularizing vector introduced in the composite semi-blind criterion (for K active users in the uplink). The asymptotic estimation error minimization is a K-variable minimization problem, which is a complex issue to deal with. We explicitly show under what conditions this problem boils down to K single-variable minimization problems. Our discussion is not limited to theoretical analyses. Simulation results performed in a realistic context (UMTS-TDD mode) are provided. In particular, we conclude about the potential of the proposed approach in real communication systems.

S. Lasaulce is with France Telecom R D, 38-40 rue du General Leclerc, 92794, Issy-les-Moulineaux, France;

S. Lasaulce is the corresponding author; e-mail: samson.lasaulce@francetelecom.com; tel: +33 1 45 29 58 24
P. Loubaton is with the Université of Marne-la-Vallée, 5 boulevard Descartes, 77454 Marne-la-Vallée, France
É. Moulines is with the École Nationale Supérieure des Télécommunications, 46 rue Barrault, 75634 Paris, France

I. Introduction

Traditional equalization techniques are based on training. The transmitter sends a sequence of known symbols (training sequence) which is used by the receiver to estimate the channel. Most current mobile telecommunication systems (*e.g.* GSM, UMTS) use a training sequence, which is usually designed to allow the receiver to estimate the channel with a desired accuracy. However, in certain adverse scenarios (low signal-to-noise ratios, high-level of interference), the training sequence alone does not suffice to obtain reliable estimates of the channel.

A sensible idea to improve accuracy of the channel estimates consists in taking into account the information that comes not only from the trained part (observations generated by the training sequence) but also from the blind part (observations generated by the information symbols), giving rise to the so-called semi-blind estimation technique (see [6] and the references therein). Semi-blind estimation techniques are usually obtained by combining both training-based and blind criteria. These methods generally improve channel estimation reliability (with respect to the purely trained case) and avoid the pitfalls of blind methods (and in particular, the lack of consistency of certain blind estimators).

The most efficient semi-blind approach is based on Maximum Likelihood estimation, which in practice can be implemented by the Expectation Maximization algorithm (see [7] for single-user systems and [17], [22] for multiuser systems). However, for synchronized CDMA communications, the computational complexity of this algorithm grows exponentially with the number of users and the size of the channel (in symbol duration) and is quite difficult (except when the number of users or channel length is very limited). Another kind of approach, based on deterministic or Gaussian Maximum Likelihood methods, have been proposed by Slock et al. [3], [6]. These methods lead to the minimization of a composite criterion defined as the sum of the classical

training-based least-squares criterion and of the criterion associated with the blind deterministic or Gaussian Maximum Likelihood method. Minimizing such a composite criterion was also proposed in [9] and [5]. In the latter cases, the blind criterion was derived from the subspace method originally introduced in [23]. Minimizing a composite criterion can be used not only to estimate the channel, but also to identify an equalizer. In particular, [14], [30] and [24] propose to mix a classical Least-Squares criterion with the Constant Modulus Algorithm. Finally, [25] proposes to adapt the blind algorithm of [28] to the semi-blind context.

Here, we consider the strategy consisting in linearly combining the training-based criterion with the blind criterion, which is not a new approach. To our knowledge, the two main contributions to this concept are [9] and [5]. In [9], it is proposed to define the semi-blind cost function as the weighted sum of the training-based and the (second-order) blind cost functions. However, no method for tuning the balance between the training-based and the blind criteria is proposed. Based on an asymptotic analysis, [5] proposes to tune the introduced weight (called the regularizing constant) by minimizing the channel estimation error. The expression of the asymptotic estimation error of the semi-blind subspace is evaluated for SIMO (single input multiple outputs) systems.

In this paper, we extend the approach of [5] to the context of synchronized CDMA systems for uplink channel estimation. In this context, we have to simultaneously estimate several channels corresponding to the different links between the mobile stations and the base station. As there are several channels to estimate in the uplink, there are also several regularizing constants to be tuned in the semi-blind cost function. This involves the challenging task of minimizing a multi-variable function, which is *not necessarily convex*. Indeed, in the uplink the asymptotic estimation error associated with the studied semi-blind schemes is a K-variable function when there are K active users. Additionally, this minimization involves a non-negligible additional

computational cost. We show how to deal with this problem and provide sufficient conditions under which the multi-variable minimization problem can be reduced to several single-variable minimization problems.

Unlike [5] where only the subspace method is considered, we discuss here the choice of the best second-order blind estimator to be used. In [5] the linear prediction approach is not considered whereas this approach might be more efficient than the subspace method if channel lengths are unknown, which is the case in the real life. In particular, we will show the importance of properly weighting the linear prediction. The corresponding weighting, which has been introduced in [10] and used in [11] in a pure blind context for a simple downlink CDMA system, will reveal a very special interest to the semi-blind estimators studied in this paper. Our discussion is not limited to theoretical analyses. We also provide simulations results performed in a realistic context (unknown channel lengths, limited number of samples, realistic signal-to-noise ratios, etc). In particular, robustness of the considered semi-blind schemes and relevancy of the proposed way of tuning the regularizing constants can be assessed.

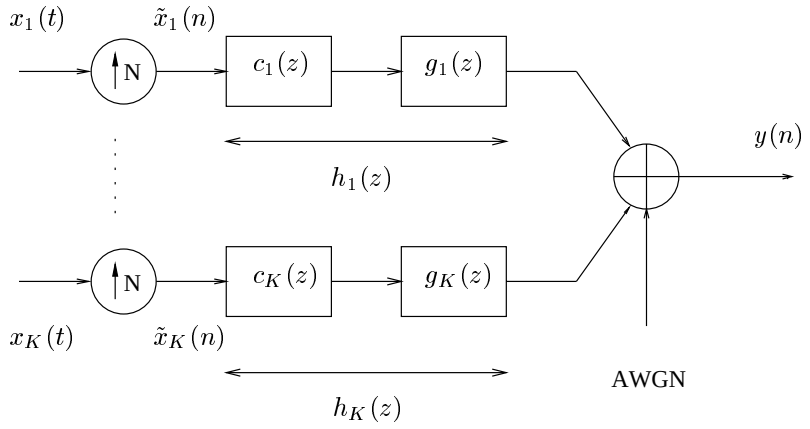
This paper is structured as follows. In section II, the classical discrete-time equivalent model of [27] describing a synchronized uplink CDMA system is reviewed. Section III briefly summarizes the main results pertaining to the blind subspace and linear prediction methods. More details on the weighted linear prediction are provided. Section IV aims at deriving the generic semi-blind cost function. In section V, we derive the closed-form expressions of the asymptotic channel estimation error in order to be able to tune the balance between the training-based cost function and the blind cost functions considered in this paper. Finally, section VI is devoted to simulations, which have been performed in the context of the UMTS-TDD[1] mode.

General Notations

[1]UMTS-TDD stands for Universal Mobile Telecommunication System - Time Division Duplex.

The notations s, $\underline{v}$ and $\mathbf{M}$ stand for scalar, vector and matrix respectively. The notations $(.)^*$, $(.)^T$, $(.)^H$ and $(.)^\#$ stand for conjugate operator, transpose operator, Hermitian operator and Moore-Penrose pseudo-inverse. For a given matrix $\mathbf{M}$, $\mathbf{M} \overset{d}{=} r \times c$ denotes the dimensions of this matrix (r is the number of rows, c is the number of columns). For a given matrix $\mathbf{M}$, $\hat{\mathbf{M}}$ denotes the estimated version of $\mathbf{M}$. At last, the Kronecker product of two matrices $\mathbf{M}$ and $\mathbf{N}$ is denoted by $\mathbf{M} \otimes \mathbf{N}$.

II. Signal Model



We consider a synchronized[2] CDMA system in the uplink with K active users. Let N be the spreading factor. Figure II represents the corresponding classical multirate (chip rate) discrete-time equivalent model (see [27] for more details). The following notations and assumptions are used:

$(A1)$ for each $k \in \{1, \cdots, K\}$, $\{x_k(t)\}_{t\in\mathbb{Z}}$ is a sequence of independent and identically distributed QPSK symbols. $\tilde{x}_k(n)$ is the corresponding up-sampled sequence. For $1 \leq k \leq K$, $1 \leq k' \leq K$, the sequences $\{x_k(t)\}_{t\in\mathbb{Z}}$ and $\{x_{k'}(t)\}_{t\in\mathbb{Z}}$ are independent.

[2]User synchronization is not an easy task but there are certain systems such as the UMTS TDD system for which this assumption is well verified. See for instance references [1] and [2].

$(A2)$ For each $k \in \{1, \cdots, K\}$, denote by $(c_k(n))_{n=0}^{N-1}$ the spreading code of user k, and by $c_k(z) = \sum_{n=0}^{N-1} c_k(n) z^{-n}$ the corresponding degree $(N-1)$ polynomial vector .

$(A3)$ The polynomials $\{g_k(z)\}_{k=1,...,K}$ are the z-transforms associated with the discrete-time equivalent of **unknown** channels sampled at the chip-rate. These channels are assumed to be causal. Since the users are synchronized, for $k \in \{1, \ldots, K\}$, $\lim_{|z|\to\infty} g_k(z) \neq 0$. The degree of $\{g_k(z)\}$ is assumed to be unknown: only an upper bound is available. For sake of simplicity, the assumed degrees of the polynomials $\{g_k(z)\}_{k=1,...,K}$ are supposed to be equal and multiple of the spreading factor. We denote by LN this degree.

We denote by $h_k(z)$ the transfer function defined by $h_k(z) = c_k(z)g_k(z)$. Then, the received signal $y(n)$ (sampled at the chip rate) may be expressed as

$$y(n) = \sum_{k=1}^{K} [h_k(z)]\tilde{x}_k(n) + v(n) \tag{1}$$

where $v(n)$ is an additive white complex circular Gaussian noise. It is often more convenient to represent the received signal by the stationary N–dimensional signal $\underline{y}(t) = (y(tN)...y(tN+N-1))^T$. It is easily seen [27] that

$$\begin{aligned} \underline{y}(t) &= [\mathbf{H}(z)]\ \underline{x}(t) + \underline{v}(t) & (2)\\ &= \sum_{l=0}^{L} \mathbf{H}(l)\ \underline{x}(t-l) + \underline{v}(t) & (3) \end{aligned}$$

where $\underline{x}(n) = (x_1(n) \ldots x_K(n))^T$ and $\underline{v}(n) = (v(nN) \ldots v(nN+N-1))^T$. $\mathbf{H}(z) = [\underline{h}_1(z) \ldots \underline{h}_K(z)]$ is a $N \times K$ polynomial matrix of degree L, with $\underline{h}_k(z) = (h_k^{(0)}(z) \ldots h_k^{(N-1)}(z))^T$ where $h_k^{(0)}(z) \ldots h_k^{(N-1)}(z)$ are the polyphase components of $h_k(z) = \sum_{l=0}^{LN+N-1} h_k(l) z^{-l}$.

It is useful to note that the relation $h_k(z) = c_k(z)g_k(z)$ may be expressed in matrix/vector form as

$$\underline{h}_k = \mathbf{C}_k \underline{g}_k \tag{4}$$

with $\underline{g}_k = (g_k(0), \ldots, g_k(LN))^T$ and

$$\mathbf{C}_k = \begin{pmatrix} c_k(0) & & \mathbf{0} \\ \vdots & \ddots & \\ \vdots & & c_k(0) \\ c_k(N-1) & & \vdots \\ & \ddots & \vdots \\ \mathbf{0} & & c_k(N-1) \end{pmatrix} \stackrel{d}{=} N(L+1) \times (NL+1). \tag{5}$$

At last, it is assumed that data are transmitted by time-slots during which the channel can be assumed to be constant. Each time-slot comprises a training sequence located in the middle of the slot, referred to as the midamble. In this paper, we address the problem of estimating the vectors $(\underline{g}_k)_{k=1,\ldots,K}$ from the midambles transmitted by each user in the current time-slot and from the observations of the current time-slot corresponding to the unknown symbols.

III. Second-Order Blind Channel Estimation

The ultimate goal of this section is obtain the blind cost functions for the subspace method, the conventional and weighted linear prediction approaches. After giving the observation model adapted to second-order blind estimation, we summarily present the subspace and linear prediction methods. Next, we study in more details a weighted version of the linear prediction because the proposed weighting not only allows the estimation performance of the semi-blind estimator to be improved (section III.C) but also makes easier its implementation (section V).

A. Observation model

Under certain assumptions ([29], [21]), each vector $\underline{g}_k$ may be estimated up to a scalar constant from the sole knowledge of the *observations* generated by the T *unknown symbols* of the slot. Let M be an integer, which is usually called smoothing factor or regression order. For $i \in \{M-1, M\}$,

we define the $N(i+1)$-dimensional regression vector as:

$$\underline{Y}_i(t) \triangleq \left[\underline{y}^T(t)\dots\underline{y}^T(t-i)\right]^T \quad (6)$$

$$= \mathcal{T}_i(\mathbf{H})\underline{X}(t) + \underline{V}(t) \quad (7)$$

where $\mathcal{T}_i(\mathbf{H})$ is the filtering matrix associated with the matrix $\mathbf{H}(z)$, defined by

$$\mathcal{T}_i(\mathbf{H}) = \begin{pmatrix} \mathbf{H}(0) & \dots & \mathbf{H}(L) & & \mathbf{0} \\ & \ddots & & \ddots & \\ \mathbf{0} & & \mathbf{H}(0) & \dots & \mathbf{H}(L) \end{pmatrix} \stackrel{d}{=} N(i+1) \times K(L+i+1)$$

and we defined $\underline{X}(t) \triangleq \left[\underline{x}^T(t)\dots\underline{x}^T(t-L-i)\right]^T$ and $\underline{V}(t) \triangleq \left[\underline{v}^T(t)\dots\underline{v}^T(t-i)\right]^T$.

B. Subspace method and conventional linear prediction in a nutshell

In this subsection, we briefly review the main steps of the subspace and linear prediction algorithms. These steps are described in the table III-B. Readers who are not very familiar with the usual notations are invited to look at the Appendix of this paper. The subspace method description corresponds to the adaptation of the original algorithm of [23] to the context of CDMA systems (see *e.g.* [4], [20], [26] and [29]). As for the linear prediction, it corresponds to the blind algorithm described in [10] and extended to downlink CDMA systems in [11].

The most important thing to note in this table is that for both algorithms, the vector impulse response $\underline{g}_k$ of user k belongs to the null-space of the matrix $\mathbf{\Delta}_k$, where $\mathbf{\Delta}_k = \mathbf{\Delta}_{k,sub} = \mathbf{\Delta}_k(\boldsymbol{\pi})$ and $\mathbf{\Delta}_k = \mathbf{\Delta}_{k,lin} = \mathbf{\Delta}_k(\mathbf{A},\mathbf{D})$ in the subspace case and the linear prediction case respectively.

In practice, the second-order statistics of the observations $(\mathcal{R}_{M-1}, \mathcal{R}_M, \{\mathbf{R}(\tau), \tau \in [0,M]\})$ cannot be perfectly recovered since the number of observations is finite (T samples per time-slot). This means that the matrices $\mathbf{\Delta}_k(\boldsymbol{\pi})$ and $\mathbf{\Delta}_k(\mathbf{A},\mathbf{D})$ have to be replaced with their consistent estimates $\mathbf{\Delta}_k(\hat{\boldsymbol{\pi}})$ and $\mathbf{\Delta}_k(\hat{\mathbf{A}},\hat{\mathbf{D}})$ respectively. Therefore, the corresponding blind cost functions

write:

$$\hat{\underline{g}}_{k,sub} = \arg\min_{\underline{f}_k} \underline{f}_k^H \underbrace{\boldsymbol{\Delta}_k^H(\hat{\boldsymbol{\pi}})\boldsymbol{\Delta}_k(\hat{\boldsymbol{\pi}})}_{\hat{\mathbf{Q}}_{k,sub}} \underline{f}_k \quad \text{and} \quad \hat{\underline{g}}_{k,lin} = \arg\min_{\underline{f}_k} \underline{f}_k^H \underbrace{\boldsymbol{\Delta}_k^H(\hat{\mathbf{A}},\hat{\mathbf{D}})\boldsymbol{\Delta}_k(\hat{\mathbf{A}},\hat{\mathbf{D}})}_{\hat{\mathbf{Q}}_{k,lin}} \underline{f}_k. \quad (8)$$

Subspace method	**Linear prediction**
• Used second-order statistics $\mathcal{R}_M \triangleq E\left[\underline{Y}_M(t)\underline{Y}_M^H(t)\right]$	• Used second-order statistics $\mathcal{R}_{M-1} \triangleq E\left[\underline{Y}_{M-1}(t)\underline{Y}_{M-1}^H(t)\right]$ $\forall\tau \in [0,M], \mathbf{R}(\tau) \triangleq E\left[\underline{y}(t+\tau)\underline{y}^H(t)\right]$
• Noise subspace equation $\boldsymbol{\pi}\mathcal{T}_M(\mathbf{H}) = \mathbf{0}$ with $\mathcal{T}_M(\mathbf{H})\mathcal{T}_M^H(\mathbf{H}) = \mathcal{R}_M - \sigma^2\mathbf{I}$	• Yule-Walker equations $[\mathbf{A}(1)\dots\mathbf{A}(M)] = -[\mathbf{R}(1)\dots\mathbf{R}(M)]\,(\mathcal{R}_{M-1} - \sigma^2\mathbf{I})^{\#}$ $\mathbf{D} = \mathbf{R}(0) + \sum_{\tau=1}^{M}\mathbf{A}(\tau)\mathbf{R}^H(\tau)$
• Code-structured subspace equations $\boldsymbol{\pi}\mathcal{T}_M(\mathbf{H}) = \mathbf{0}$ $\Leftrightarrow \forall k \in [1,K], \mathcal{D}(\boldsymbol{\pi})\underline{h}_k = \underline{0}$ $\Leftrightarrow \forall k \in [1,K], \underbrace{\mathcal{D}(\boldsymbol{\pi})\mathbf{C}_k}_{\boldsymbol{\Delta}_k(\boldsymbol{\pi})}\underline{g}_k = \underline{0}$	• Innovation covariance matrix $\mathbf{D} = \mathbf{H}(0)\mathbf{H}^H(0)$ and $\boldsymbol{\pi}_{\mathbf{D}^\perp}$ defined by $\boldsymbol{\pi}_{\mathbf{D}^\perp}\mathbf{D} = \mathbf{0}$ • Irreductibility of $\mathbf{H}(z)$ and Bezout identity $\forall z \neq 0, \text{Rank}\,(\mathbf{H}(z)) = K$ $\Rightarrow \exists\ \mathbf{A}(z)$ such that $\mathbf{A}(z)\mathbf{H}(z) = \mathbf{H}(0)$ • Code-structured linear prediction equations $\forall k \in [1,K], \underbrace{\text{Diag}(\boldsymbol{\pi}_{\mathbf{D}^\perp},\mathbf{I},\dots,\mathbf{I})\mathcal{S}(\mathbf{A})\mathbf{C}_k}_{\boldsymbol{\Delta}_k(\mathbf{A},\mathbf{D})}\underline{g}_k = \underline{0}$

C. Weighted linear prediction approach

For single-user systems, the linear prediction approach is known to have poor statistical performance. Unlike the subspace method, the asymptotic estimation error of the linear prediction estimate is non-zero in absence of noise ($\sigma^2 = 0$). Motivated by this observation, [10] proposed to

use a weighted linear prediction estimate and addressed the *blind* identification of *unstructured* MIMO FIR transfer function. The idea of [10] was recently adapted to the context of blind channel identification of *downlink CDMA* channel [11], where it was shown that the use of a simple weighted matrix can produce significant performance improvements. More specifically, the corresponding asymptotic estimation error decreases toward 0 when $\sigma^2 \to 0$ if the assumed channel length does not exceed the true one by more than one symbol duration. In this case, the blind subspace and weighted linear prediction methods have nearly the same performance. However, if the assumed degree of $\mathbf{H}(z)$ (say $\hat{L}$) is greater than the true degree of $\mathbf{H}(z)$ by more than one $(L+1)$, the subspace method is no longer consistent whereas the weighted linear prediction scheme still provides satisfying performance. The results of [11] can be easily adapted to our context (*uplink CDMA*) by estimating each vector impulse response as the solution of the following minimization problem:

$$\underline{g}_{k,wlin} = \arg\min_{\underline{f}_k} \underline{f}_k^H \underbrace{\boldsymbol{\Delta}_k(\hat{\mathbf{A}},\hat{\mathbf{D}})^H \mathbf{W}_k(\mathbf{D},\sigma^2)\boldsymbol{\Delta}_k(\hat{\mathbf{A}},\hat{\mathbf{D}})}_{\hat{\mathbf{Q}}_{k,wlin}} \underline{f}_k \tag{9}$$

where the admissible weighting matrix $\mathbf{W}_k(\mathbf{D},\sigma^2)$ is defined by

$$\forall k = 1,\ \ldots, K,\ \ \mathbf{W}_k(\mathbf{D},\sigma^2) = \mathbf{I}_{M+L+1} \otimes (\boldsymbol{\pi}_{\mathbf{D}^\perp} + \sigma^2 \mathbf{D}^{\#}). \tag{10}$$

More insights on the construction of the weighting matrix will be given in section V.

IV. Semi-Blind Channel Estimation

The main purpose of this section is to derive a generic expression of the composite semi-blind cost function, which is obtained by linearly combining the training-based cost function and the blind cost functions (corresponding to the different channels). In this section, we first introduce the classical training sequence based estimates of the channels $(\underline{g}_k)_{k=1,\ldots,K}$ and next make use of the results of section III to form a generic semi-blind cost function.

A. Training Sequence Based Estimation

In the sequel, we denote by "$m+(NL+1)$" the number of chips of each user's midamble, and $\underline{Y}_{tr}$ the m–dimensional vector obtained by stacking the m useful observation samples generated by the midambles. It is easily seen that

$$\underline{Y}_{tr} \quad = \quad \mathbf{S}\underline{g} + \underline{V}_{tr} \tag{11}$$

where $\underline{g} = [\underline{g}_1^T ... \underline{g}_K^T]^T$ and where $\mathbf{S}$ is a $m \times K(NL+1)$ matrix straightforwardly constructed from the midambles. The Maximum Likelihood estimate of $\underline{g}$ based on the observation $\underline{Y}_{tr}$ is given by

$$\hat{\underline{g}}_{tr} = \arg\min_{\underline{f}} \parallel \underline{Y}_{tr} - \mathbf{S}\underline{f} \parallel^2 \tag{12}$$

i.e.

$$\hat{\underline{g}}_{tr} = \mathbf{R}_{SS}^{-1}\underline{R}_{SY}, \tag{13}$$

where we introduced the matrix $\mathbf{R}_{SS} = m^{-1}\mathbf{S}^H\mathbf{S}$ and the vector $\underline{R}_{SY} = m^{-1}\mathbf{S}^H\underline{Y}_{tr}$.

B. Semi-Blind Estimation

In order to introduce our results, we first remark that the previous blind estimates are obtained by minimizing for each $k = 1, \ldots, K$ a generic quadratic form $\underline{f}_k^H \hat{\mathbf{Q}}_k \underline{f}_k$. The matrix $\hat{\mathbf{Q}}_k$ respectively equals $\hat{\mathbf{Q}}_{k,sub}$, $\hat{\mathbf{Q}}_{k,lin}$ or $\hat{\mathbf{Q}}_{k,wlin}$ in the context of the subspace method, the linear prediction approach or the weighted linear prediction approach. For the sake of clarity, a few definitions are in order. We define the matrix $\hat{\mathbf{Q}}$ by

$$\hat{\mathbf{Q}} = \mathrm{Diag}(\hat{\mathbf{Q}}_1, \ldots, \hat{\mathbf{Q}}_K) \tag{14}$$

and for each K–dimensional vector $\underline{\alpha}$ the matrix $\hat{\mathbf{Q}}(\underline{\alpha})$ by

$$\hat{\mathbf{Q}}(\underline{\alpha}) = \mathrm{Diag}\left(\alpha_1\hat{\mathbf{Q}}_1, \ldots, \alpha_K\hat{\mathbf{Q}}_K\right). \tag{15}$$

Moreover, $\mathbf{W}(\mathbf{D},\sigma^2)$ is defined by

$$\mathbf{W}(\mathbf{D},\sigma^2) = \mathrm{Diag}\left(\mathbf{W}_1(\mathbf{D},\sigma^2),\ldots,\mathbf{W}_K(\mathbf{D},\sigma^2)\right) = \mathbf{I}_{K(M+L+1)} \otimes (\boldsymbol{\pi}_{\mathbf{D}^\perp} + \sigma^2 \mathbf{D}^{\#}). \tag{16}$$

Finally, we define in the same way the block diagonal matrix $\boldsymbol{\Delta}(\underline{\alpha}) = \mathrm{Diag}(\alpha_1\boldsymbol{\Delta}_1,\ldots,\alpha_K\boldsymbol{\Delta}_K)$, $\hat{\boldsymbol{\Delta}}$; each $\boldsymbol{\Delta}_k$ equals $\boldsymbol{\Delta}_k(\boldsymbol{\pi})$ in the subspace method or $\boldsymbol{\Delta}_k(\mathbf{A},\mathbf{D})$ in the linear prediction approaches.

For every matrix $\hat{\mathbf{Q}}_k$, we study the semi-blind estimates of $\underline{g}_1,\ldots,\underline{g}_K$ obtained by minimizing the composite cost function

$$\phi(\underline{f},\underline{\alpha}) = \|\underline{Y}_{tr} - \mathbf{S}\underline{f}\|^2 + T\left(\sum_{k=1}^{K} \alpha_k \underline{f}_k^H \hat{\mathbf{Q}}_k \underline{f}_k\right) = \|\underline{Y}_{tr} - \mathbf{S}\underline{f}\|^2 + T\underline{f}^H \hat{\mathbf{Q}}(\underline{\alpha})\underline{f} \tag{17}$$

where $\underline{f} = (\underline{f}_1^T \ldots \underline{f}_K^T)^T$ and the components of the vector $\underline{\alpha} = (\alpha_1,\ldots,\alpha_K)^T$ are positive constants weighting the contribution of each cost function in the global cost function (17). The corresponding estimate is given by

$$\underline{\hat{g}} = \arg\min_{\underline{f}} \| \underline{Y}_{tr} - \mathbf{S}\underline{f} \|^2 + T\underline{f}^H \hat{\mathbf{Q}}(\underline{\alpha})\underline{f} \tag{18}$$

that is

$$\underline{\hat{g}} = \left[\mathbf{R}_{SS} + \rho\hat{\mathbf{Q}}(\underline{\alpha})\right]^{-1} \underline{R}_{SY}, \tag{19}$$

where ρ is defined by $\rho = T/m$, which is the ratio of the time-slot size to the midamble size. One of the crucial point here is to derive the "best" value of $\underline{\alpha}$. In the single-user context of [5] (there a single parameter α to adjust), it has been noticed that the choice of α dramatically influenced the performance of the estimate.

It is therefore of great practical importance to adjust in a relevant way the vector $\underline{\alpha}$. In the same spirit as in [5], we propose choosing the value of the vector $\underline{\alpha}$ in order to minimize the asymptotic estimation error of the semi-blind estimate. By asymptotic, we mean that both the

midamble size (m) and the time-slot size (T) converge toward the infinity, but in such a way that $\rho = T/m$ remains constant. To this end, we calculate in section V the closed-form expressions of the estimation error for each semi-blind approach under consideration in this paper.

V. Asymptotic Estimation Error

In this section we generalize the calculations of [5] (single-user system, subspace method) to the context under consideration (uplink CDMA, subspace and linear prediction methods). Proposition 1, which is given below, provides the general expression of the asymptotic estimation error for the semi-blind schemes studied in this paper.

Proposition 1: The asymptotic covariance matrix of the estimation error ($\delta\underline{g} = \underline{g} - \underline{\hat{g}}$) *is given by:*

$$\mathrm{Cov}(\delta\underline{g}) = \lim_{T\to\infty, T/m=\rho} \left\{ TE(\delta\underline{g}\delta\underline{g}^H) \right\} = \mathbf{M}^{-1}(\underline{\alpha}) \left[\rho\sigma^2 \mathbf{R}_{SS}^{(\infty)} + \rho^2 \mathbf{\Delta}^H(\underline{\alpha}) \mathbf{W}\mathbf{\Sigma}\mathbf{W}\mathbf{\Delta}(\underline{\alpha}) \right] \mathbf{M}^{-1}(\underline{\alpha}) \tag{20}$$

where $\mathbf{\Sigma} \triangleq \mathrm{Cov}(\delta\mathbf{\Delta}\ \underline{g})$, $\delta\mathbf{\Delta} \triangleq \mathbf{\Delta} - \hat{\mathbf{\Delta}}$ and $\mathbf{M}(\underline{\alpha}) \triangleq \mathbf{R}_{SS}^{(\infty)} + \rho\mathbf{Q}(\underline{\alpha})$.

$\mathbf{W} = \mathbf{I}$ *in the subspace and the non-weighted linear prediction cases, and* $\mathbf{W} = \mathbf{W}(\mathbf{D}, \sigma^2)$ *(see equation (16)) in the context of the weighted linear prediction approach. The matrix* $\mathbf{R}_{SS}^{(\infty)}$ *represents* $\lim_{m\to+\infty} m^{-1}\mathbf{R}_{SS}$. This result can be proved along the lines of [5]. The proof is therefore omitted.

In the sequel, $\mathbf{\Gamma}(\underline{\alpha})$ denotes the asymptotic covariance matrix of $\delta\underline{g}$ defined in (20). Our approach consists in selecting the value of $\underline{\alpha}$ minimizing $\gamma(\underline{\alpha}) \triangleq \mathrm{Trace}\,(\mathbf{\Gamma}(\underline{\alpha}))$. However, this minimization is not easy for at least four reasons.

- The optimum value of $\underline{\alpha}$ cannot be explicitly found from equation (20).
- The multi-variable function $\gamma(\underline{\alpha}) = \gamma(\alpha_1, ..., \alpha_K)$ is not necessarily convex.
- The exhaustive search for the best value of $\underline{\alpha}$ is in general not implementable. Indeed, if we

want to test n_α values for each regularizing constant $\{\alpha_k, k \in [1,K]\}$, there are $N_{\underline{\alpha}} = K^{n_\alpha}$ vectors to be tested for the vector $\underline{\alpha}$.

- It can be shown that the matrix $\mathbf{\Sigma}$ defined in (20) depends on the true channels ($\underline{g} = \left[\underline{g}_1 \dots \underline{g}_K\right]^T$) and is thus difficult to estimate consistently.

It turns out that it is possible to overcome this problem by making reasonable assumptions. For this purpose, we provide in the following two theorems sufficient conditions under which the tough multi-dimensional minimization problem emphasized above can be reduced to several one-dimensional minimization problems.

Theorem 1: Assume that $\mathbf{R}_{SS}^{(\infty)} = \lim_{m\to+\infty} m^{-1}\mathbf{R}_{SS}$ *is block diagonal and* $\mathcal{T}_M(\mathbf{H})$ *has full column rank. Then, the estimation error covariance matrix for the semi-blind subspace case is given by:*

$$\mathbf{\Gamma}_{sub}(\underline{\alpha}) = \sigma^2 \mathrm{Diag}\left(\mathbf{\Gamma}_{1,sub}^{(1)}(\alpha_1), \dots, \mathbf{\Gamma}_{K,sub}^{(1)}(\alpha_K)\right) + \mathbf{O}(\sigma^4) \tag{21}$$

where $\forall k \in [1,K]$ *the matrix* $\mathbf{\Gamma}_{k,sub}^{(1)}(\alpha_k)$ *depends only on* α_k *and* $\boldsymbol{\pi}$.

A sketchy proof of this theorem is provided in the Appendix and a more detailed proof can be found in [16]. From this theorem, we see that, under three additional assumptions, the estimation error covariance matrix is block diagonal and the function $\gamma(\underline{\alpha})$ writes $\gamma(\underline{\alpha}) = \sum_{k=1}^{K} \gamma_k(\alpha_k)$, which is easy to minimize with respect to the different regularizing constants $\alpha_1, ..., \alpha_K$.

In the linear prediction case, we can get the same kind of result if the weighting matrix $\mathbf{W}$ is properly chosen as the following theorem shows.

Theorem 2: Assume that $\mathbf{R}_{SS}^{(\infty)} = \lim_{m\to+\infty} m^{-1}\mathbf{R}_{SS}$ *is block diagonal and* $\mathcal{T}_M(\mathbf{H})$ *has full column rank. Then, the estimation error covariance matrix for the semi-blind linear prediction case verifies the following properties:*

1. *if* $\mathbf{W} = \mathbf{I}$ *then*

$$\lim_{\sigma \to 0} \mathbf{\Gamma}_{lin}(\underline{\alpha}) = \mathbf{\Gamma}_{lin}^{(0)}(\underline{\alpha}) \neq \mathbf{0} \quad \text{and} \quad \mathbf{\Gamma}_{lin}(\underline{\alpha}) = \mathbf{\Gamma}_{lin}^{(0)}(\underline{\alpha}) + \sigma^2 \mathbf{\Gamma}_{lin}^{(1)}(\underline{\alpha}, \underline{g}) + \mathbf{O}(\sigma^4) \tag{22}$$

2. *if* $\mathbf{W} = \mathbf{I} \otimes (\boldsymbol{\pi}_{\mathbf{D}^\perp} + \sigma^2 \mathbf{D}^\#)$ *then*

$$\lim_{\sigma \to 0} \mathbf{\Gamma}_{wlin}(\underline{\alpha}) = \mathbf{0} \quad \text{and} \quad \mathbf{\Gamma}_{wlin}(\underline{\alpha}) = \sigma^2 \text{Diag}\left(\mathbf{\Gamma}_{1,wlin}^{(1)}(\alpha_1), \ldots, \mathbf{\Gamma}_{K,wlin}^{(1)}(\alpha_K)\right) + \mathbf{O}(\sigma^4) \tag{23}$$

where $\forall k \in [1, K]$ *the matrix* $\mathbf{\Gamma}_{k,wlin}^{(1)}(\alpha_k)$ *depends only on* α_k, $\mathbf{A}$ *and* $\mathbf{D}$.

A sketchy proof of this theorem is provided in the Appendix and a more detailed proof can be found in [16]. From Theorem 2, we first see that the estimation error of the non-weighted semi-blind linear prediction does not decrease toward 0 when $\sigma^2 \to 0$. We also see that the term in σ^2 directly depends on the true channels (via $\underline{h}$). But the most critical issue is that the covariance matrix for the semi-blind linear prediction is not block diagonal. On other hand, the proposed weighted semi-blind linear prediction estimator has exactly the same behavior as the semi-blind subspace estimator, which means that the three drawbacks of its non-weighted counterpart are eliminated. As a consequence, the proposed weighting not only improves the statistical performance of the semi-blind linear prediction but also considerably facilitates the implementation of the underlying algorithm. Because of these very attractive features, only the weighted version of the linear prediction will be considered in the simulation section. Before tackling the simulation part it is convenient to review the main steps to follow in order to implement the proposed semi-blind scheme.

Final algorithm implementation: main steps

We consider the semi-blind subspace example, the adaptation to the semi-blind linear prediction is straightforward.

- Estimate the matrices associated with the blind algorithm ($\hat{\mathcal{R}}_M$, SVD on $\hat{\mathcal{R}}_M$, $\hat{\boldsymbol{\pi}}$)
- Choose a finite set of values (say a n_α-element set) to tune the regularizing constants

- For each user k
  - Compute for each value of the chosen set, $\hat{\gamma_k}(\alpha_k) = \sigma^2 \text{Trace}(\hat{\mathbf{\Gamma}}^{(1)}_{k,sub})$ by using Theorem 1 (21)
  - Keep only the value of α_k that minimizes $\hat{\gamma_k}(\alpha_k)$
- Put the optimum regularizing constant in one vector $\underline{\hat{\alpha}}$
- Plug the latter in the semi-blind channel estimate given by (19).

VI. Simulations

A. Goals

Essentially, the objective of this section is threefold.

1. As the proposed way of tuning the regularizing vector is based on an asymptotic analysis, we want to know to what extent this assumption is realistic. We also evaluate the influence of tuning accuracy of the regularizing vector on the receiver performance.

2. Our second aim is to evaluate the potential of the proposed semi-blind approach in terms of performance for different midamble sizes and especially for training sequences shorter than those used in the TDD mode.

3. Eventually, we want to complete the (theoretical) discussion of section V on the best blind scheme to be selected for semi-blind estimation. To this end, we make a comparison between the semi-blind subspace and (weighted) linear prediction estimators in the realistic context considered here; in particular we want to study robustness of the proposed algorithms to channel overmodeling. This comparison is of great interest because the pure blind linear prediction is known to be more robust to flawed channel knowledge than its subspace counterpart.

B. Simulation setup

The chosen simulation context is the TDD mode of the UMTS system. In the uplink, the various users are synchronized, and the spreading factor is $N = 16$. The number of active users is, on each time slot, less than $K_{max} = 8$. The transmission is divided in slots of 2560 chips ended by a guard interval of 96 chips. In the uplink, each slot conveys two blocks of $T/2 = 61$ QPSK symbols separated by a 512-chip midamble (recall that $m = 512 - (\hat{L}N + 1)$). The shaping filter is a root raised-cosine filter which roll-off factor equals 0.22 and it is truncated at 7 chip durations. The most classical uplink receivers consist in estimating the discrete-time version (sampled at the chip rate) of each channel by using the midamble. Then, a joint detection algorithm [13], based on the trained estimate, allows for the recovery of the emitted symbols. In the following, we evaluate the performance of our estimation schemes by means of the bit error rate (BER) provided by the MMSE block linear joint detection (MMSE-BL-JD) algorithm of [13] based on our semi-blind estimates. The parameters of our simulation chain are the followings:

– uplink case (single antenna) and $K = 4$ active users per time-slot;

– fixed spreading factor: $N = 16$;

– no channel coding;

– propagation channel: the chosen propagation environment is the Vehicular A specified by the ITU (6 paths having each one a specified time delay and an average power). The propagation channel length is about 10 chips;

– physical channel: it is the convolution of the propagation channel by the shaping filter. The time duration of the physical channels is about 16 chips;

– the channel coefficients are drawn for each time slot according to the Jakes fading model revisited by Dent et al. [8]; the Doppler velocity is fixed at 120 km/h;

– the noise standard deviation is: $\sigma = \sqrt{\frac{N}{2}} 10^{-\frac{Eb/No|dB}{20}}$; it is assumed to be known, which is not restrictive since it can be shown that accurate noise variance estimation is possible. For instance, based on the (very realistic) continuous time-slot transmission assumption, [18] proposes to exploit the fact that channel lengths are generally overestimated. The idea is simply to estimate the energy of the weakest channel coefficients by making use of several channel estimates (provided by different slots), given the fact that these coefficients are pure noise;

– bit error rates are averaged over 2000 time-slots and over the 4 users;

– the matrix $\mathbf{R}_{SS}$ is assumed to be block diagonal. It can checked that for UMTS-TDD midambles and channel lengths considered here, the diagonal blocks contain 95% of the energy of this matrix;

– according to the results of section V on the error covariance matrices, the terms in σ^4 will be neglected;

– for each regularizing constant, $n_\alpha = 20$ test values are used over the interval $[0, 4]$ by means of an hyperbolic tangent scale.

Note: the "gain on Eb/No" (say G_{dB}) is defined for a given symbol detection strategy (namely an MMSE Block Linear Joint Detector) and a given bit error rate (BER target). For the MMSE-BL Joint Detector using the training-based channel estimate, denote by ρ_{TS} the Eb/No (in dB) needed to achieve the BER target. In the same way, for the MMSE-BL Joint Detector using the semi-blind channel estimate under consideration, denote by ρ_{SB} the Eb/No (in dB) needed to achieve the BER target. The gain on Eb/No is simply defined by $G_{dB} \triangleq \rho_{TS} - \rho_{SB}$.

C. Simulation results

Accuracy of the estimation of the parameters $(\alpha_k)_{k=1,\ldots,K}$

For each time-slot, the values of the regularizing constants are chosen in order to minimize

the trace of the estimated asymptotic covariance matrix of the channel estimation error. In the following experiments, we compare our semi-blind estimates when the values of parameters $(\alpha_k)_{k=1,\ldots,K}$ are given by our procedure, and when the values of $(\alpha_k)_{k=1,\ldots,K}$ are given by an oracle minimizing the true channel estimation error on each time slot ($\forall k = 1,\ldots,K,\ \hat{\alpha}_k = \arg\min_\alpha \|\underline{g}_k - \underline{\hat{g}}_k(\alpha_k)\|^2$). The various parameters of the simulation all correspond to the specifications of the TDD mode of the UMTS. In particular, $T = 122$ unknown symbols and the midamble length is 512. The assumed channel length is two symbol durations ($\hat{L} = 2$, $\hat{L}N = 32$ chip durations).

The following table provides some statistics on the regularizing constant associated with a given user. Notation " $< . >$ " stands for averaging over all the generated slots. "SS" and "LP" mean subspace and (weighted) linear prediction respectively. As mentioned above, the oracle under consideration is based on the knowledge of the true channel estimation error for every time-slot.

$\frac{E_b}{N_o}$	5 dB	7.5 dB	10 dB	12.5 dB
$< \alpha >\ \ (SS/\{SS, oracle\})$	0.793 / 0.629	0.783 / 0.599	0.772 / 0.610	0.763 / 0.659
$< \alpha >\ \ (LP/\{LP, oracle\})$	0.623 / 0.241	0.638 / 0.219	0.630 / 0.222	0.660 / 0.2594

It turns out that the regularizing constant is quite well tuned in the regularized semi-blind subspace case, even for low signal-to-noise ratios. This means that the assumption made on the σ^4-terms is justified. This does not seem to be the case in the regularized semi-blind weighted linear prediction case. This might be due to the limited number of samples. The linear prediction seems to need more samples to reach the "asymptotic regime".

Now, we study the impact of tuning accuracy of the regularizing constants on the receiver performance. *Figure 1* represents the gain on E_b/N_o provided by semi-blind schemes under consideration as a function of the targeted BER. We compare the regularized semi-blind subspace method with the regularized semi-blind subspace method equipped with the oracle; the corresponding curves are very close. In the weighted linear prediction case, for a BER target ranging from 1% to 10%, the loss due to the non-optimality of the tuning is more important (about 0.2 dB). Equally, note that for the specified parameters of the TDD mode and for a 1% BER target, our semi-blind methods provides **gains of 0.4 dB and 1 dB** in the weighted linear prediction and the subspace cases respectively.

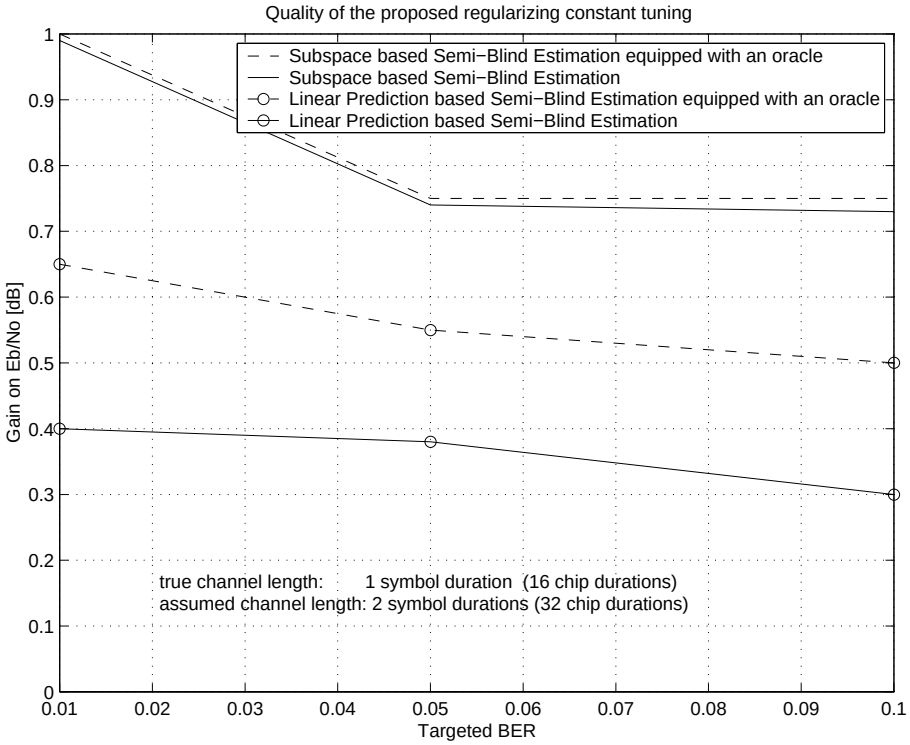


Fig. 1. Influence of the tuning accuracy of the regularizing constant on the gain on Eb/No

Influence of the midamble size

In this experiment, parameters are always the same as those used in the TDD mode, except for

the training sequence length. E_b/N_o is fixed to 12.5 dB (the results are similar for lower SNRs). The assumed channel length is taken to be equal to 32 chip durations.

Figure 2 represents the BER provided by semi-blind approaches versus m (recall that the size of the midamble equals $m + \hat{L}N$). For $m = 480$ chips, the BER associated with the trained channel estimator equals 1%. The BER is about 0.85% for the weighted linear prediction method and 0.65% for the subspace method. But for shortest training sequences, BERs are no longer that close. Indeed, for $m = 152$, the achieved BER equals 10% in the trained case and 1% in the regularized semi-blind subspace case. We also note that in order to achieve a BER of 1%, the trained estimate needs $m = 480$ chips while the semi-blind subspace estimate only requires $m = 152$. The use of this semi-blind estimate thus allows us to reduce the midamble size by 3. This roughly corresponds to a gain of **15% in terms of data rate**.

In order to confirm the results presented in *figure 2*, we compare in *figure 3* the performance of the proposed semi-blind estimates with the trained one versus E_b/N_0 in the case where $m = 152$. The duration of the channel is always 32 chip durations and $T = 122$. It is seen that the semi-blind approaches outperform quite significantly the performance provided by the classical trained estimate.

Influence of the assumed channel length

Robustness of channel estimator to channel overmodeling is of course an important issue. The impact of channel overmodeling on receiver performance is evaluated in *figure 4*. For a 1% BER target, it depicts the gains provided by the semi-blind approaches on E_b/N_o versus the assumed channel length in symbol duration. In dashdot line, we have represented the gain that could be achieved if channels were known from the symbol detector. We notice that the semi-blind approaches provide gains on E_b/N_o between **0.8 dB** (weighted linear prediction approach) and

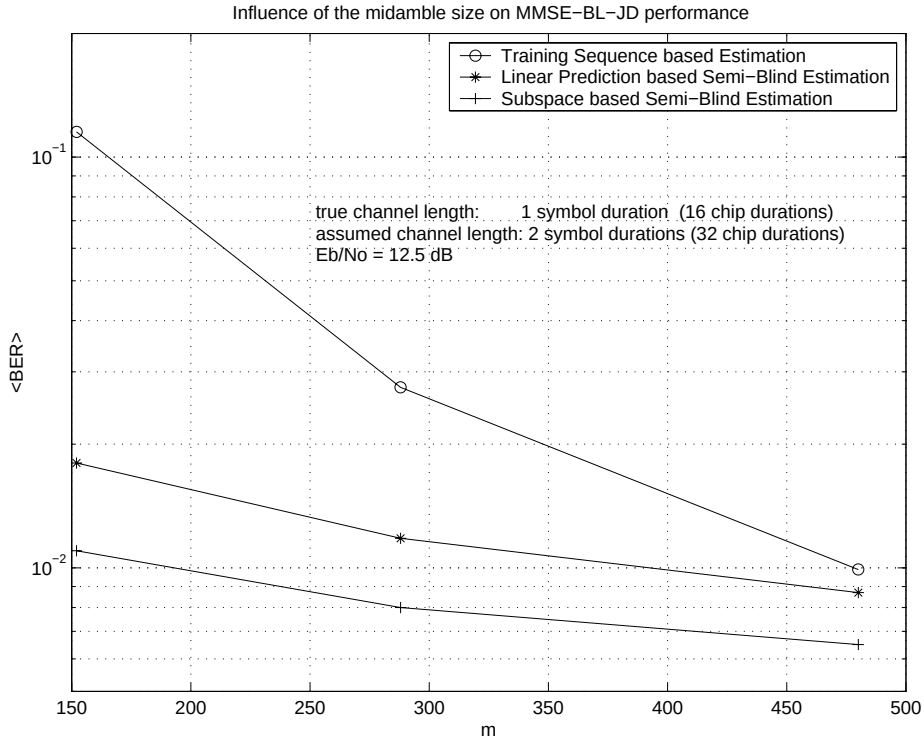


Fig. 2. Influence of the midamble size on performance

1.2 dB (subspace method) for an assumed channel length of 48 chip durations, which roughly corresponds to the maximum channel length in the context of UMTS (see [12]).

VII. Conclusions

Semi-blind channel estimation methods for synchronized uplink CDMA systems have been considered in this paper. The central idea of this approach was to minimize composite criteria and to tune the underlying regularizing constants by evaluating the corresponding asymptotic covariances of the estimation error. We provided these covariance matrices for the subspace and a weighted linear prediction based semi-blind schemes.

We have identified sufficient conditions under which the considered multi-variable minimiza-

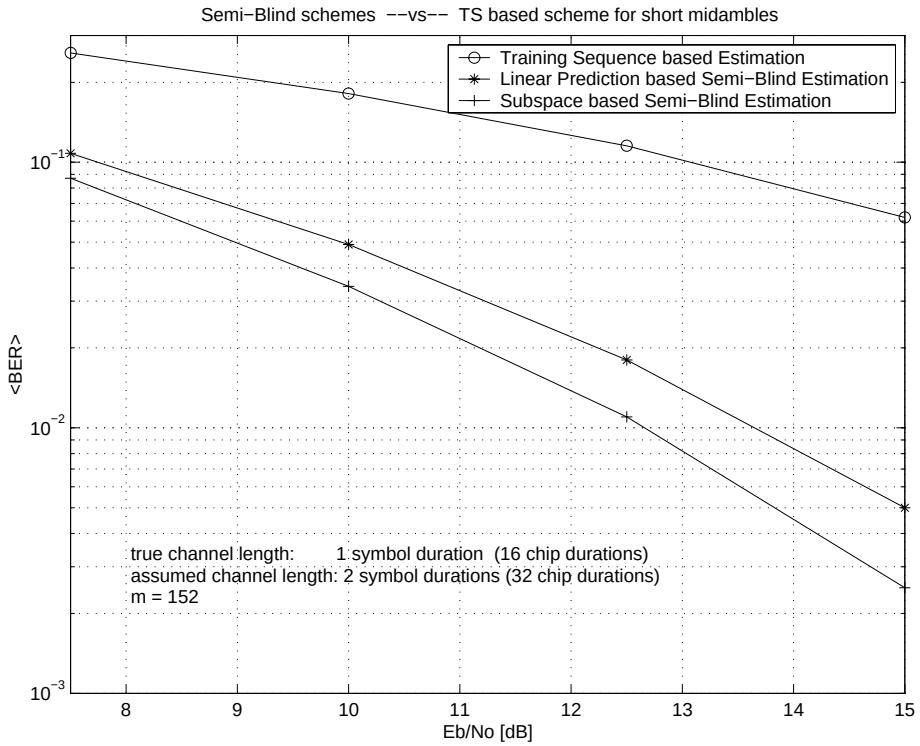


Fig. 3. MMSE-BL-JD performance with trained channel estimate versus MMSE-BL-JD with semi-blind channel estimates when training sequence is short

tions can be split into several one-variable minimizations, which considerably facilitates the research of the optimum values of the regularizing constants. Moreover, it has been seen how beneficial it can be to weight the linear prediction criterion: the proposed weighting both improves the statistical performance of the semi-blind estimation and makes easier the estimation of the covariance matrix of the error, which is required to tune the regularizing constants.

The potential of the presented approaches has been evaluated in a realistic context. In this respect, at least two points are worth being highlighted:

- Semi-blind approaches perform quite well but they are especially useful when training sequences are short (5% to 10% of the time-slot duration typically). In this case, they work dramatically better than the pure trained approach. In the context of the UMTS-TDD mode, significant

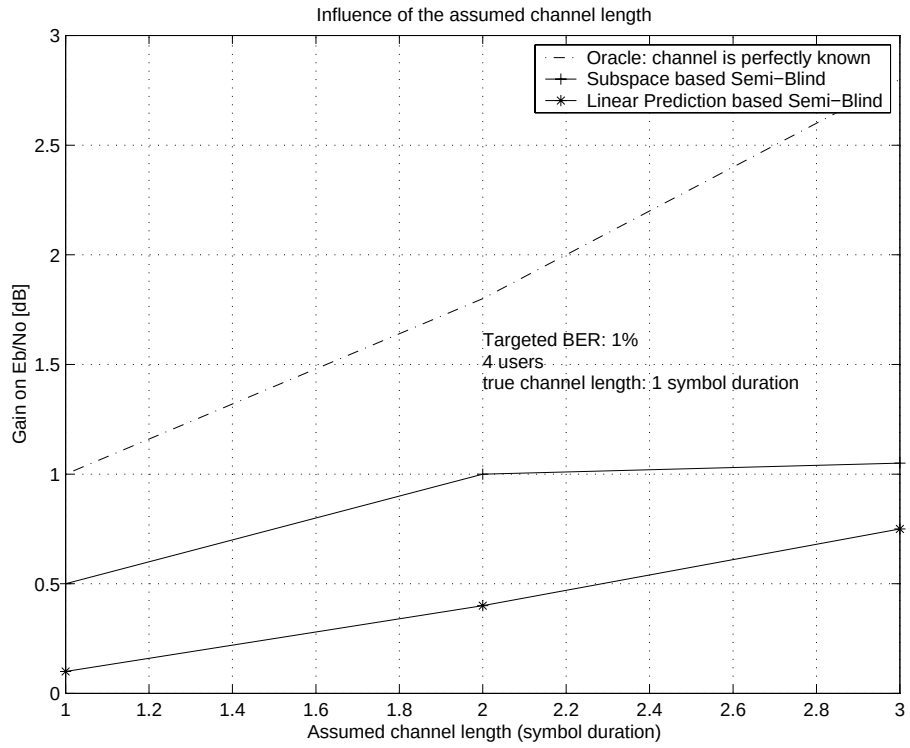


Fig. 4. Robustness of the semi-blind estimators to channel overmodeling

improvements have been stated in terms of data rate, quality of service or power consumption. For instance, semi-blind approaches allow us to achieve the same performance as the classical trained estimate but with a training sequence which is three times shorter. This corresponds to increase the data throughput by 15%.

- As for the comparison of the semi-blind schemes between themselves, it has been seen that the semi-blind subspace approach generally outperforms the weighted linear prediction one in the cases of interest. This fact was not obvious since it was shown recently [11] that the blind linear prediction approach dominates the blind subspace method in the cases where channel lengths are overestimated.

Regarding the complexity issue, it has been found that the additional computational cost due to the semi-blind estimation is reasonable [15]. The complexity depends on many factors: the

number of active users per time-slot (K), the assumed channel length (say $\hat{\ell}$ chips per user), the smoothing factor (M), the midamble length (m) and the number of regularizing constants to be tested (n_α). In the typical scenario presented in this paper ($K = 4, \hat{\ell} = 32, M = 1, m = 512, n_\alpha = 20$), the complexity of the channel estimator is increased by 5.8 with respect to the training-based channel estimator. By way of comparison, the joint minimization of the regularizing constants (with 80 test vectors) leads to a complexity increased by 25.5 w.r.t the training-based estimator.

Appendix

Comments on table III-B

- *Second-order statistics estimation:* we use the empirical observation covariance matrix defined by $\hat{\Re}_M = T^{-1} \sum_{t=0}^{T-1} \underline{Y}_M(t) \underline{Y}_M^H(t)$.
- *Noise subspace estimation:* $\hat{\boldsymbol{\pi}} = \mathbf{U}\mathbf{U}^H$ where $\mathbf{U}$ is the singular vectors matrix associated with the $N(M+1) - K(L+M+1)$ smallest singular values of $\hat{\Re}_M$.
- *Definition of* $\mathcal{D}(\boldsymbol{\pi})$:

$$\mathcal{D}(\boldsymbol{\pi}) = \begin{bmatrix} \boldsymbol{\pi}_0 & & \mathbf{0} \\ & \ddots & \\ \vdots & & \boldsymbol{\pi}_0 \\ & & \\ \boldsymbol{\pi}_M & & \vdots \\ & \ddots & \\ \mathbf{0} & & \boldsymbol{\pi}_M \end{bmatrix} \stackrel{d}{=} N(L+M+1)(M+1) \times N(L+1),$$

where $\boldsymbol{\pi} = [\boldsymbol{\pi}_0 \dots \boldsymbol{\pi}_M]$.

- *Pseudo-inverse definition* (#): the matrix $\hat{\Re}_{M-1} - \sigma^2\mathbf{I}$ is in general numerically invertible whereas $\Re_{M-1} - \sigma^2\mathbf{I}$ is rank deficient for large enough M. Therefore, $(\hat{\Re}_{M-1} - \sigma^2\mathbf{I})^{\#}$ is obtained

by truncating the eigenvalue decomposition of $\hat{\Re}_{M-1} - \sigma^2\mathbf{I}$ to its significant eigenvalues. Denote $(\hat{\lambda}_k)_{k=1,NM}$ the eigenvalues of $\hat{\Re}_{M-1} - \sigma^2\mathbf{I}$ arranged in decreasing order. The truncation index k_0 is defined as the index such that $\frac{\hat{\lambda}_{k_0}}{\hat{\lambda}_{k_0+1}}$ is maximum (see [19]).

- *Estimation of* $\mathbf{H}(0)$: this matrix is simply a $N \times K$ square root of $\hat{\mathbf{D}}$. Of course, it is estimated up to a constant unitary matrix. This missing factor is identified by exploiting the specific algebraic structure of $\mathbf{H}(z)$, which is given by the CDMA codes.
- *Definition of* $\mathcal{S}(\mathbf{A})$: the Sylvester matrix associated with the prediction filter $\mathbf{A}(z)$ is given by

$$\mathcal{S}(\mathbf{A}) = \begin{pmatrix} \mathbf{A}(0) & \mathbf{0} & \dots & \mathbf{0} \\ \vdots & & \ddots & \mathbf{0} \\ \vdots & & & \mathbf{A}(0) \\ \mathbf{A}(M) & & & \vdots \\ \mathbf{0} & \ddots & & \vdots \\ \vdots & & \ddots & \vdots \\ \mathbf{0} & \dots & \mathbf{0} & \mathbf{A}(M) \end{pmatrix}. \tag{24}$$

A sketchy proof of Theorem 1 (semi-blind subspace estimation)

We want to show how the general expression of the estimation error covariance matrix (20) can be reduced to (21). Assume that the matrix $\mathbf{R}_{SS}^{(\infty)}$ is block diagonal. Then $\mathbf{M}(\underline{\alpha})$ is also block diagonal since $\mathbf{\Delta}(\underline{\alpha})$ and $\mathbf{Q}(\underline{\alpha})$ are block diagonal (by definition). As $\mathbf{W} = \mathbf{I}$, we just need to evaluate the matrix $\mathbf{\Sigma} = \mathrm{Cov}(\delta\mathbf{\Delta}\ \underline{g})$ and show that it is also block diagonal. To that end, define by $\forall (k,l) \in [1,K]^2, \mathbf{\Sigma}(k,l)$ the K^2 blocks of the matrix $\mathbf{\Sigma}$. It can be shown [16] that

$$\mathbf{\Sigma} = \sigma^2\mathbf{\Sigma}^{(1)} + \sigma^4\mathbf{\Sigma}^{(2)} \tag{25}$$

where for each $(k,l) \in [1,K]^2$,

- $\mathbf{\Sigma}^{(1)}(k,l) = \delta_{k,l}\mathbf{R}_{\boldsymbol{\pi}}$

- $\mathbf{\Sigma}^{(2)}(k,l) = \int_{-\pi}^{\pi}[\mathbf{J}_k^T\mathbf{T}(e^{i\omega})\mathbf{T}^H(e^{i\omega})\mathbf{J}_l]^* \otimes \boldsymbol{\pi}(e^{i\omega})\boldsymbol{\pi}(e^{i\omega})d\omega$

with

- $\boldsymbol{\pi}(z) = \sum_{i=0}^{M}\boldsymbol{\pi}_i z^{-i}$
- $\mathbf{R}_{\boldsymbol{\pi}} = \int_{-\pi}^{\pi}\left[\underline{D}_{M+L}(e^{i\omega})\underline{D}_{M+L}^H(e^{i\omega})\right]^* \otimes \boldsymbol{\pi}(e^{i\omega})\boldsymbol{\pi}^H(e^{i\omega})\, d\omega$
- $\underline{D}_P(e^{i\omega}) \triangleq (1\ e^{-i\omega}\dots e^{-iP\omega})^T$
- $\mathbf{J}_k = (\mathbf{I}_{M+L} \otimes \underline{u}_k)$
- $\underline{u}_k = (0\dots0\ 1\ 0\dots0)^T \stackrel{d}{=} K\times 1$, the one is on the position k^{th}
- $\mathbf{T}(e^{i\omega}) = \sum_{p=0}^{P-1}\mathbf{T}(p)e^{-ip\omega}$
- $\Im_{M-1}^{\#}(\mathbf{H}) \triangleq [\mathbf{T}(0)\ \dots\ \mathbf{T}(M-1)]$.

From this, it is easily seen that $\mathbf{\Sigma} = \mathbf{I}_K \otimes \mathbf{R}_{\boldsymbol{\pi}} + \mathbf{O}(\sigma^4)$, which proves that $\mathbf{\Sigma}$ is block diagonal up to the σ^4-terms.

A sketchy proof of Theorem 2 (semi-blind linear prediction)

We want to show how the general expression of the estimation error covariance matrix (20) can be reduced to (23). As mentioned in the proof of Theorem 1, the only problem is to derive the expression of $\mathbf{\Sigma} = \mathrm{Cov}(\delta\mathbf{\Delta}\ \underline{g})$. The goal is to show that $\mathbf{W\Sigma\Sigma}$ is block diagonal up to the σ^4-terms. Let partition each block $\mathbf{\Sigma}(k,l)$ of $\mathbf{\Sigma}$ as:

$$\mathbf{\Sigma}(k,l) = \left[\begin{array}{cc} \mathbf{\Sigma}_{11}(k,l) & \mathbf{\Sigma}_{12}(k,l) \\ \mathbf{\Sigma}_{21}(k,l) & \mathbf{\Sigma}_{22}(k,l) \end{array}\right] \stackrel{d}{=} N(M+L+1)\times N(M+L+1)$$

where the blocks $\mathbf{\Sigma}_{11}(k,l), \mathbf{\Sigma}_{12}(k,l)$ and $\mathbf{\Sigma}_{22}(k,l)$ are $N\times N, N\times N(M+L)$ and $N(M+L)\times N(M+L)$ respectively. Using straightforward but tedious perturbation theory calculations it is possible to show that [16]

$$\mathbf{\Sigma} = \mathbf{\Sigma}^{(0)} + \sigma^2\mathbf{\Sigma}^{(1)} + \sigma^4\mathbf{\Sigma}^{(2)} \tag{26}$$

where the matrices $\mathbf{\Sigma}^{(0)}$, $\mathbf{\Sigma}^{(1)}$ and $\mathbf{\Sigma}^{(2)}$ are defined as follows:

$\forall (k,l) \in \{1,\ldots,K\}$,

$$\mathbf{\Sigma}(k,l) \triangleq \mathbf{\Sigma}^{(0)}(k,l) + \sigma^2 \mathbf{\Sigma}^{(1)}(k,l) + \sigma^4 \mathbf{\Sigma}^{(2)}(k,l),$$

and the matrices $\mathbf{\Sigma}^{(0)}(k,l)$ and $\mathbf{\Sigma}^{(1)}(k,l)$ are given by

$$\begin{aligned}
\mathbf{\Sigma}_{22}^{(0)}(k,l) &= \delta_{k,l} \times (\mathbf{I}_{M+L} \otimes \mathbf{D}) \\
\mathbf{\Sigma}_{11}^{(1)}(k,l) &= \delta_{k,l} \times \left(\int_{-\pi}^{\pi} \boldsymbol{\pi}_{D^\perp} \mathbf{A}(e^{i\omega}) \mathbf{A}^H(e^{i\omega}) \boldsymbol{\pi}_{D^\perp} d\omega \right) \\
\mathbf{\Sigma}_{12}^{(1)}(k,l) = \mathbf{\Sigma}_{21}^{(1),H}(k,l) &= \delta_{k,l} \times \left(\int_{-\pi}^{\pi} e^{-i\omega} \underline{D}_{M+L-1}^T(\omega) \otimes \boldsymbol{\pi}_{D^\perp} \mathbf{A}(e^{i\omega}) \mathbf{A}^H(e^{i\omega}) d\omega \right) \\
\mathbf{\Sigma}_{22}^{(1)}(k,l) &= \int_{-\pi}^{\pi} \left[\mathbf{J}_k \mathbf{T}(e^{i\omega}) \mathbf{T}^H(e^{i\omega}) \mathbf{J}_l \right]^* \otimes \mathbf{D} d\omega \\
&\quad + \left[\underline{D}_{M+L-1}(\omega) \underline{D}_{M+L-1}^H(\omega) \right]^* \otimes \mathbf{A}(e^{i\omega}) \mathbf{A}^H(e^{i\omega}) \quad \delta_{k,l}\, d\omega.
\end{aligned}$$

The matrix $\mathbf{\Sigma}^{(2)}(k,l)$ depends on $\boldsymbol{\pi}_{D^\perp}$, $\mathbf{A}$, $\mathbf{H}^{\#}(0)$ and on the true channel itself via the matrix $\mathbf{T}$.

Apart from the many details given here, what is essentially to note is the effect of the weighting matrices $\forall k = 1, \ldots, K$, $\mathbf{W}_k(\mathbf{D}, \sigma^2) = \mathbf{I}_{M+L+1} \otimes (\boldsymbol{\pi}_{\mathbf{D}^\perp} + \sigma^2 \mathbf{D}^{\#})$ on the matrix $\mathbf{\Sigma}$. As $\boldsymbol{\pi}_{\mathbf{D}^\perp} \mathbf{D} = \mathbf{0}$, it can be checked that the constant term associated with $\mathbf{\Sigma}^{(0)}$ is cancelled. The other major effect of the proposed weighting is to cancel the first term of $\mathbf{\Sigma}_{22}^{(1)}(k,l)$, which makes this matrix both block diagonal and independent of the true channels. Therefore, up to the σ^4-terms (represented by the matrix $\mathbf{\Sigma}^{(2)}$), the matrix $\mathbf{W\Sigma W}$ is block diagonal, which concludes the proof.

References

[1] TSG RAN WG1 V 3.7.0 TS 25.221. Physical channels and mapping of transport channels onto physical channels (TDD). Technical report, 3GPP, 1999.

[2] TSG RAN WG1 V 3.1.0 TS 25.224. Physical layer procedures (TDD). Technical report, 3GPP, 1999.

[3] J. Ayadi, É. De Carvalho, and D. Slock. Blind and semi-blind maximum likelihood methods for fir multichannel identification. In *ICASSP*, Seattle, May 1998.

[4] S. Bensley and B. Aazhang. Subspace-based Channel Estimation for Code Division Multiple Access Communication Systems. *IEEE Transactions on Communications*, 44(8):1009–1020, August 1996.

[5] V. Buchoux, O. Cappé, É. Moulines, and A. Gorokhov. On the Performance of Semi-Blind Subspace-Based Channel Estimation . *IEEE Transactions on Signal Processing*, 48(6):1750–1759, June 2000.

[6] É. De Carvalho and D. Slock. Semi-blind maximum likelihood multichannel estimation with gausssian prior for the symbols using soft decision. In *VTC*, Ottawa, May 1998.

[7] H. Cirpan and M. Tsatsanis. Stochastic Maximum Likelihood Methods for Semi-Blind Channel Estimation. *IEEE Trans. on Signal Processing*, 5(1), January 1998.

[8] P. Dent, G. Bottomley, and T. Croft. Jakes Fading Model Revisited. *Electronics Letter*, 29(13):1162–1163, June 1993.

[9] A. Gorokhov and Ph. Loubaton. Semi-Blind Second Order Identification of Convolutive Channels. *Proc. ICASSP*, pages 3905–3908, 1997.

[10] A. Gorokhov and Ph. Loubaton. Blind identification of MIMO-FIR systems: A generalized linear prediction approach. *Signal Processing (Elsevier)*, (73):105–124, 1999.

[11] S. Houcke and Ph. Loubaton. A weighted linear prediction approach for the blind cdma forward link channel estimation. In *ICASSP*, Istanbul, June 2000.

[12] European Telecommunications Standards Institute. *UTRA TDD, Transport channels and Physical channels description.* Sophia Antipolis FRANCE.

[13] A. Klein, G. Kaleh, and P. Baier. Zero Forcing and Minimum Mean-Square-Error Equalization for Multiuser Detection in Code-Division Multiple-Access Channels. *IEEE Trans. on Vehicular Technology*, 45(2), May 1996.

[14] A. Kusminsky, L. Féty, P. Forster, and S. Mayrargue. Regularized semi-blind estimation of spatio-temporal filter coefficientsfor mobile radio-communications. In *GRETSI*, pages 127–130, Grenoble, September 1997.

[15] S. Lasaulce. Semi-blind channel estimation: performance and complexity analysis. Technical report, Motorola Labs, Centre de recherche de Motorola, 91193 Gif-sur-Yvette, November 1999.

[16] S. Lasaulce. *Channel estimation and multiuser detection for TD-CDMA systems.* PhD thesis, ENST Paris, 46, rue Barrault 75013 Paris, November 2001.

[17] S. Lasaulce, K. Gosse, Ph. Loubaton, and É. Moulines. Performance of a semi-blind receiver for the uplink of the umts tdd mode. In *VTC Fall*, Amsterdam, September 1999.

[18] S. Lasaulce, P. Loubaton, É. Moulines, and S. Buljore. Training-based channel estimation and de-noising in the UMTS-TDD mode. In *VTC Fall*, Atlantic City, October 2001.

[19] A. Liavas, P. Regalia, and J-P. Delmas. Blind channel approximation: effective channel order determination. In *Thirty-second Asilomar Conference on Signals, Systems and Computers*, 1998.

[20] H. Liu and G. Xu. A subspace method for signal waveform estimation in synchronous cdma systems. *IEEE Transactions on Communications*, 44:1346–1354, October 1996.

[21] P. Loubaton and É. Moulines. On Blind Multiuser Forward Link Channel Estimation be the Subspace Method: Identifiability Results. *IEEE Transactions on Signal Processing*, 48(8):2366–2376, August 2000.

[22] J. Miguez and L. Castedo. Blind and semi-blind optimum demodulation in multiuser communications. In *ICASSP*, volume 5, pages 2917–2920, Istanbul, June 2000.

[23] É. Moulines, P. Duhamel, JF. Cardoso, and S. Mayrargue. Subspace Methods for the Blind Identification of Multichannel FIR Filters. *IEEE Transactions on Signal Processing*, 43:516–525, February 1995.

[24] R. Pacheco and D. Hatzinakos. Semi-blind suppression of MAI in multipath CDMA channels. In *Thirty-second Asilomar Conference on Signals, Systems and Computers*, volume 2, pages 1138–1142, November 1999.

[25] A. Swindelhurst and J. Gunther. Direct semi-blind symbol estimation for multipath channel. In *Thirty-second Asilomar Conference on Signals, Systems and Computers*, volume 2, pages 1124–1128, November 1998.

[26] M. Torlak and G. Xu. Blind multiuser channel estimation in asynchronous CDMA systems. *IEEE Transactions on Signal Processing*, 45(1), January 1997.

[27] M. Tsatsanis and G. Giannakis. Optimal Decorrelating Receivers for DS-CDMA Systems : A Signal Processing Framework. *IEEE Trans. on Signal Processing*, 44(12):3044–3055, December 1996.

[28] A. Van Der Veen, S. Talwar, and A. Paulraj. Joint angle and delay estimation using shift invariance properties. *IEEE Transactions on Signal Processing*, 45(1):173–190, January 1997.

[29] X. Wang and V. Poor. Blind Multiuser Detection: A Subspace Approach. *IEEE Transactions on Information Theory*, 44(2):677–690, March 1998.

[30] X. Zhuang and A. Swindlehurst. A space-time semi-blind equalizer based on constant modulus and decision-direction. In *Thirty-third Asilomar Conference on Signals, Systems and Computers*, volume 2, pages 1017–1021, 1999.