

Pilot-aided ML Schemes for Joint Beamforming and Channel Estimation in WCDMA Systems

Xavier Mestre and Javier R. Fonollosa
 Department of Signal Theory and Communications
 Universitat Politècnica de Catalunya
 c/ Jordi Girona 1-3, Mòdul D-5
 Campus Nord UPC 08034 Barcelona (Spain)
 e-mail: {mestre, fono}@gps.tsc.upc.es

Abstract — This paper proposes semi-blind channel estimation and interference cancellation schemes for the reception of pilot-aided signals in WCDMA systems. It is shown that the performance of classical training-based schemes is severely degraded due to the code-multiplexing (as opposed to time-multiplexing) of traffic and pilot signals. The time-dispersivity of the channel destroys the orthogonality between the traffic signal and the pilot, and consequently the traffic signal appears as a new interfering signal for the channel estimation/beamforming algorithm. In order to avoid this effect, we propose techniques that exploit both the presence of the training sequence and the structure of the traffic signal in a semi-blind fashion.

I. INTRODUCTION

The introduction of training sequences or reference signals for channel estimation purposes is quite a common practice in mobile communication systems. Traditionally the pilot or training signal has been transmitted time-multiplexed with the traffic data—that is, occupying different temporal positions on the radio frame—so that channel estimation and data detection could be performed independently. Nevertheless, in most of the terrestrial mobile communication standards proposed today—such as WCDMA or CDMA2000—training sequence and traffic signal are transmitted simultaneously while code-multiplexed. Although the traffic signal and the pilot are in principle transmitted with perfectly orthogonal codes, the time-dispersivity of the mobile radio channel destroys this orthogonality at the reception stage. For this reason, the use of channel estimators or even interference suppression schemes based on the sole knowledge of the training sequence results in very poor performance at the basestation, no matter how good the signal to noise ratio might be.

In this paper we propose channel estimators and interference rejection schemes based on a semi-blind approach to the underlying identification problem. The proposed techniques make use of the knowledge of the training sequence while exploiting the inherent structure of the traffic signal. Semi-blind channel identification techniques yield moderate improvements over classical training-only estimators when traffic and training signals are time-multiplexed [1]. We will see here that these improvements turn out to be dramatic when these signals are code-multiplexed.

The rest of the paper is organized as follows. Section II describes the signal model for DS/CDMA-modulated signals

with emphasis on multi-rate schemes. In section III three different semi-blind channel estimators designed for pilot-aided WCDMA signals are proposed. These results are then used in Section IV to derive three corresponding semi-blind beamforming methods. Section V compares the performance of the proposed schemes.

II. RECEIVED SIGNAL MODEL

Each mobile station is assumed to map the underlying data sequence to Q distinct and synchronized spreading sequences with period N_c chips. Both the period of the spreading sequences and the chipping rate are assumed constant for all users and sequences. Let $s_q(m) \in \mathbb{C}$ represent the underlying complex symbol stream associated with the q -th sequence and assume that $N_s(q)$, $q = 1 \dots Q$ consecutive symbols are mapped to each interval of N_c chips. The spreading factor associated with a particular spreading sequence is expressed as $SF(q) = N_s(q)/N_c$ chips/bit.

The chip-sampled spreading sequences are denoted $c_1(n) \dots c_Q(n)$, all of them with non-zero support over $1 \leq n \leq N_c$, and $t(n)$ stands for the training sequence transmitted in parallel with the traffic channels. On the other hand, $h(l)$ $l = 1 \dots L$ represents the channel impulse response, where it is assumed that $L < N_c$. Using these definitions the chip rate-sampled received signal can be expressed as:

$$x(n) = x^k(n) + \sum_{q=1}^Q \sum_{k=-\infty}^{\infty} s_q(k) g_q \left(n - \left\lfloor \frac{k}{N_s(q)} \right\rfloor N_c \right) + n(n), \quad (1)$$

with $\lfloor \cdot \rfloor$ denoting the lower largest integer value and $x^k(n)$ the channel-filtered known component of the transmitted signal (typically a pilot or training sequence mapped to a code orthogonal to the traffic channels):

$$x^k(n) = \sum_{i=1}^L h(i) t(n - i + 1), \quad (2)$$

Finally $g_q(n)$ stands for the signature associated with the q -th code sequence, resulting from the discrete convolution between the q -th spreading sequence and the channel impulse response:

$$g_q(n) = \sum_{i=1}^L h(i) c_q(n - i + 1), \quad (3)$$

and $n(n)$ a component modelling thermal noise. Next we define a column vector $\mathbf{s}_q(m) \in \mathbb{C}^{N_s(q) \times 1}$ containing a fragment of the data sequence mapped to the q -th spreading code:

$$\mathbf{s}_q(m) = [s_q((m-1)N_s(q)+1) \dots s_q(mN_s(q))]^T. \quad (4)$$

Gathering N_c samples of the received signal $x(n)$ into a column vector $\mathbf{x}(m) = [x((m-1)N_c+1) \dots x(mN_c)]^T \in \mathbb{C}^{N_c \times 1}$,

⁰This work was partially supported by the European Commission under IST Project METRA; the Spanish Government (CYCIT) TIC98-0412 TIC98-0703 TIC99-0849; and the Catalan Government (CIRIT) 1998SGR 00081.

we can express the global received signal in an observation window of N_c chips as:

$$\mathbf{x}(m) = \mathbf{x}^k(m) + \sum_{q=1}^Q [\mathbf{G}_q(2) \mathbf{G}_q(1)] \begin{bmatrix} \mathbf{s}_q(m) \\ \mathbf{s}_q(m+1) \end{bmatrix} + \mathbf{n}(m). \quad (5)$$

Matrices $\mathbf{G}_q(i) \in \mathbb{C}^{N_c \times N_s(q)}$ $i = 1, 2$ are the two parts of the global signature matrix:

$$\begin{bmatrix} \mathbf{G}_q(1) \\ \mathbf{G}_q(2) \end{bmatrix} = \begin{bmatrix} \mathbf{C}_q(1) \\ \mathbf{C}_q(2) \end{bmatrix} (\mathbf{I}_{N_s(q)} \otimes \mathbf{h}) \quad (6)$$

$$= \begin{bmatrix} \mathbf{g}_q^1 \mathbf{g}_q^2 \dots \mathbf{g}_q^{N_s(q)} \end{bmatrix} \in \mathbb{C}^{2N_c \times N_s(q)},$$

$$\begin{bmatrix} \mathbf{C}_q(1) \\ \mathbf{C}_q(2) \end{bmatrix} = \begin{bmatrix} \mathbf{C}_q^1 \mathbf{C}_q^2 \dots \mathbf{C}_q^{N_s(q)} \end{bmatrix} \in \mathbb{C}^{2N_c \times N_s(q)L} \quad (7)$$

$$\mathbf{g}_q^T = \begin{bmatrix} \mathbf{0}_{(r-1)SF(q) \times 1} & \dots & 0 \\ c_q((r-1)SF(q)+1) & \dots & 0 \\ \vdots & \ddots & \vdots \\ c_q(rSF(q)) & \dots & c_q((r-1)SF(q)+1) \\ \vdots & \ddots & \vdots \\ 0 & \dots & c_q(rSF(q)) \\ 0 & \dots & \mathbf{0}_{2N_c \rightarrow rSF(q)-L \times 1} \end{bmatrix} \mathbf{h},$$

$\mathbf{C}_q^T \in \mathbb{C}^{2N_c \times L}$

$$\mathbf{h} = [h(1) \dots h(L)]^T \in \mathbb{C}^{L \times 1}, \quad (8)$$

where $\mathbf{0}_{(r-1)SF(q) \times 1}$ stands for an all-zeros $(r-1)SF(q) \times 1$ column vector. The vector corresponding to the received training sequence $\mathbf{x}^k(m) = [x^k((m-1)N_c+1) \dots x^k(mN_c)]^T \in \mathbb{C}^{N_c \times 1}$ can be expressed as:

$$\mathbf{x}^k(m) = \begin{bmatrix} t((m-1)N_c+1) \dots t((m-1)N_c-L+2) \\ t((m-1)N_c+2) \dots t((m-1)N_c-L+3) \\ \vdots \\ t(mN_c) \dots t(mN_c-L+1) \end{bmatrix} \mathbf{h}.$$

$\mathbf{T}(m) \in \mathbb{C}^{N_c \times L}$

Stacking M vectors $\mathbf{x}(m)$ into a column vector $\mathbf{x} \in \mathbb{C}^{MN_c \times 1}$ we get to:

$$\mathbf{x} = [\mathbf{x}^T(1) \dots \mathbf{x}^T(M)]^T = \underbrace{\mathbf{T}\mathbf{h}}_{\mathbf{x}^k \in \mathbb{C}^{MN_c \times 1}} + \sum_{q=1}^Q \mathbf{G}_q \mathbf{s}_q + \mathbf{n}, \quad (10)$$

$$\mathbf{G}_q = \begin{bmatrix} \mathbf{G}_q(2) & \mathbf{G}_q(1) & \dots & \mathbf{0}_{N_c \times N_s(q)} \\ \mathbf{0}_{N_c \times N_s(q)} & \mathbf{G}_q(2) & \dots & \mathbf{0}_{N_c \times N_s(q)} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0}_{N_c \times N_s(q)} & \dots & \mathbf{G}_q(2) & \mathbf{G}_q(1) \end{bmatrix}_{M \times (M+1) \text{ blocks}}$$

$$= \begin{bmatrix} \mathbf{C}_q(2) & \mathbf{C}_q(1) & \dots & \mathbf{0}_{N_c \times N_s(q)} \\ \mathbf{0}_{N_c \times N_s(q)} & \mathbf{C}_q(2) & \dots & \mathbf{0}_{N_c \times N_s(q)} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0}_{N_c \times N_s(q)} & \dots & \mathbf{C}_q(2) & \mathbf{C}_q(1) \end{bmatrix} (\mathbf{I}_{(M+1)N_s(q)} \otimes \mathbf{h})$$

$\mathbf{C}_q \in \mathbb{C}^{2N_c M \times (M+1)N_s(q)L}$

$$\mathbf{s}_q = [\mathbf{s}_q^T(1) \dots \mathbf{s}_q^T(M+1)]^T \in \mathbb{C}^{(M+1)N_s(q) \times 1}, \quad (11)$$

with $\mathbf{T} = [\mathbf{T}^T(1) \dots \mathbf{T}^T(M)]^T$ and $\mathbf{n} = [\mathbf{n}^T(1) \dots \mathbf{n}^T(M)]^T$. The final received signal model is therefore expressed as:

$$\mathbf{x} = \mathbf{x}^k + \mathbf{G}\mathbf{s} + \mathbf{n}, \quad (12)$$

being now:

$$\begin{aligned} \mathbf{G} &= [\mathbf{G}_1 \dots \mathbf{G}_Q] \in \mathbb{C}^{N_c M \times (M+1)N_s} \\ \mathbf{C} &= [\mathbf{C}_1 \dots \mathbf{C}_Q] \in \mathbb{C}^{N_c M \times (M+1)N_s L} \\ \mathbf{s} &= [\mathbf{s}_1^T \dots \mathbf{s}_Q^T]^T \in \mathbb{C}^{(M+1)N_s \times 1} \end{aligned} \quad (13)$$

and $N_s = \sum_{q=1}^Q N_s(q)$. It is assumed that the components of the thermal noise vector are circularly symmetric jointly Gaussian distributed with zero mean and covariance $E[\mathbf{nn}^H] = \sigma^2 \mathbf{I}_{MN_c}$, $E[\mathbf{nn}^T] = \mathbf{0}$. Whenever the unknown symbols are modelled as random variables, they are also assumed circularly symmetric with zero mean and unit covariance matrix.

III. ML CHANNEL ESTIMATION METHODS

Training-only Approach

The training only estimator disregards the presence of the traffic channels, which is equivalent to setting $\mathbf{s} = \mathbf{0}$ in (12). The ML estimator under this hypothesis can be easily found as:

$$\hat{\mathbf{h}}_{to} = (\mathbf{T}^H \mathbf{T})^{-1} \mathbf{T}^H \mathbf{x}. \quad (14)$$

The normalized estimator $\sqrt{M} \hat{\mathbf{h}}_{to}$ can be shown to be unbiased Gaussian-distributed with asymptotic covariance¹:

$$\mathbf{C}_{to}^M = M \mathbf{C} \mathbf{R} \mathbf{B}_{to}^M + M (\mathbf{T}^H \mathbf{T})^{-1} \mathbf{T}^H \mathbf{G} \mathbf{G}^H \mathbf{T} (\mathbf{T}^H \mathbf{T})^{-1}, \quad (15)$$

being:

$$\mathbf{C} \mathbf{R} \mathbf{B}_{to}^M = \sigma^2 (\mathbf{T}^H \mathbf{T})^{-1} \quad (16)$$

the Cramér-Rao bound for the training-only scenario. Although traffic and training sequences are designed to be perfectly orthogonal, the time dispersivity of the channel destroys this orthogonality rendering the training-only channel estimator inefficient even at high signal to noise ratios.

The poor performance of the training-only estimator can be overcome modelling explicitly the presence of the traffic channels in the received signal. In particular, one can model the unknown data either as unknown deterministic parameters (Conditional approach) or as random variables (Unconditional Approach). These two approaches will lead to two distinct estimators that do not perform equivalently [1].

Conditional (Deterministic) Approach

If we model the unknown data as deterministic parameters, the ML estimator for both data and channel impulse response can be obtained minimizing the following negative log-likelihood function:

$$\eta_{CML} = MN_c \log(\pi \sigma^2) - \frac{1}{\sigma^2} \|\mathbf{x} - \mathbf{x}^k - \mathbf{G}\mathbf{s}\|^2. \quad (17)$$

The unknown symbols can be estimated as:

$$\hat{\mathbf{s}}_c = [\hat{\mathbf{G}}^H \hat{\mathbf{G}}]^{-1} \hat{\mathbf{G}}^H (\mathbf{x} - \mathbf{T} \hat{\mathbf{h}}_c), \quad (18)$$

with $\hat{\mathbf{h}}_c$ the CML estimate of the channel impulse response and $\hat{\mathbf{G}}$ equal to \mathbf{G} substituting \mathbf{h} for its estimation $\hat{\mathbf{h}}_c$. Plugging (18) into (17) we get to the simplified cost function:

$$\eta'_{CML} = MN_c \log(\pi \sigma^2) + \frac{1}{\sigma^2} (\mathbf{x} - \mathbf{T} \hat{\mathbf{h}}_c)^H \mathbf{P}_{\hat{\mathbf{G}}}^\perp (\mathbf{x} - \mathbf{T} \hat{\mathbf{h}}_c) \quad (19)$$

¹Note that only the circulant part of the covariance matrix is different from zero.

with $\mathbf{P}_G^\perp = \mathbf{I}_{MN_c} - \mathbf{G} [\mathbf{G}^H \mathbf{G}]^{-1} \mathbf{G}^H$ denoting the orthogonal projection matrix onto the null space of the columns of \mathbf{G} . Estimations of the channel impulse response and the noise power can be readily obtained minimizing either (17) or (19):

$$\begin{aligned} \hat{\sigma}_c^2 &= \frac{1}{MN_c} (\mathbf{x} - \mathcal{T} \hat{\mathbf{h}}_c)^H \mathbf{P}_G^\perp (\mathbf{x} - \mathcal{T} \hat{\mathbf{h}}_c) \\ \hat{\mathbf{h}}_c &= (\hat{\mathbf{R}}^H \hat{\mathbf{R}})^{-1} \hat{\mathbf{R}}^H \mathbf{x} \end{aligned} \quad (20)$$

being $\hat{\mathbf{R}} = \mathcal{T} + \mathcal{C}(\hat{\mathbf{s}}_c \otimes \mathbf{I}_L)$ a structured reconstruction of the transmitted signal convolution matrix. The normalized estimator $\sqrt{M} \hat{\mathbf{h}}_c$ can be shown to be strongly consistent and asymptotically circularly symmetric Gaussian-distributed with zero mean and covariance matrix shown in (21) on top of the next page, where $\mathcal{C}(i) = \mathcal{C}_{:, (i-1)L+1:iL} \in \mathbb{C}^{2N_c M \times L}$ and CRB_c^M is the asymptotic Deterministic Cramér-Rao Bound:

$$\sigma^2 (\text{CRB}_c^M)^{-1} = \mathcal{T}^H \mathbf{P}_G^\perp \mathcal{T} + \sum_{i=1}^{(M+1)N_s} \mathcal{C}^H(i) \mathbf{P}_G^\perp \mathcal{C}(i) \quad (22)$$

As we see from (21), the deterministic channel estimator is inefficient at finite values of the signal to noise ratio. This is due to the finite number of samples per symbol (finite spreading factors), from which a consistent estimation of the unknown symbols is not possible.

Gaussian Approach

According to the GML approach, unknown symbols are modelled as jointly Gaussian random variables with zero mean and unit variance. The likelihood function to be minimized becomes:

$$\eta_{GML} = \log \det(\pi \mathbf{C}_x) + \text{tr}(\mathbf{C}_x^{-1} \tilde{\mathbf{C}}_x), \quad (23)$$

where now \mathbf{C}_x and $\tilde{\mathbf{C}}_x$ stand for the temporal covariance matrix of the received signal and its rank-one sample estimate respectively:

$$\mathbf{C}_x = \mathbf{G} \mathbf{G}^H + \sigma^2 \mathbf{I}_{MN_c}, \quad (24)$$

$$\tilde{\mathbf{C}}_x = (\mathbf{x} - \mathcal{T} \mathbf{h})(\mathbf{x} - \mathcal{T} \mathbf{h})^H. \quad (25)$$

Note that the GML channel estimator tends to be equivalent to the CML scheme as the signal to noise ratio increases without bound, so that $\mathbf{C}_x^{-1} \rightarrow \frac{1}{\sigma^2} \mathbf{P}_G^\perp$. Taking derivatives with respect to the channel taps, the GML solution for the channel impulse response can be expressed as follows [3]:

$$\hat{\mathbf{h}}_g = [\hat{\mathbf{D}} + \mathcal{T}^H \hat{\mathbf{C}}_x^{-1} \mathcal{T}]^{-1} \mathcal{T}^H \hat{\mathbf{C}}_x^{-1} \mathbf{x}, \quad (26)$$

with:

$$\hat{\mathbf{D}} = \sum_{i=1}^{(M+1)N_s} \mathcal{C}^H(i) [\hat{\mathbf{C}}_x^{-1} - \hat{\mathbf{C}}_x^{-1} \tilde{\mathbf{C}}_x \hat{\mathbf{C}}_x^{-1}] \mathcal{C}(i), \quad (27)$$

and $\hat{\mathbf{C}}_x$ equal to \mathbf{C}_x replacing \mathbf{h} for its estimation. It can be shown that the GML channel estimator is asymptotically non-circular Gaussian-distributed and efficient. The expression of the asymptotic covariance matrix (Cramér-Rao bound) is omitted here for lack of space, but can be found in [4]. For a comparison of the theoretical performance of the semi-blind ML channel estimators, see also [2].

IV. BEAMFORMING DESIGN

Let us now consider the design of a beamformer in order to suppress the contribution from all the undesired components of the received signal, consisting of both noise and multi-user interference. First, we generalize the received signal model in (12) to the multi-antenna case. Assuming the model in (12) for the received signal at the p -th antenna:

$$\mathbf{x}_p = \mathcal{T} \mathbf{h}_p + \mathcal{C}(\mathbf{I}_L \otimes \mathbf{s}) \mathbf{h}_p + \mathbf{n}_p, \quad (28)$$

we stack the P received signal vectors next to one another to obtain the spatial-temporal received signal matrix:

$$\mathbf{X} = [\mathbf{x}_1 \cdots \mathbf{x}_P] = \mathcal{T} \mathbf{H} + \mathcal{C}(\mathbf{I}_L \otimes \mathbf{s}) \mathbf{H} + \mathbf{N} \in \mathbb{C}^{MN_c \times P}. \quad (29)$$

Matrix $\mathbf{H} \in \mathbb{C}^{L \times P}$ contains on each of its columns the channel impulse response at the corresponding antenna, and $\mathbf{N} = [\mathbf{n}_1 \cdots \mathbf{n}_P] \in \mathbb{C}^{MN_c \times P}$ contains noise and interference from mobile stations other than the desired one.

Assume that we have an estimation of the spatial-temporal channel matrix after beamforming $\mathbf{h} = \mathbf{H} \mathbf{w}$. Our objective is to design a narrowband beamformer $\mathbf{w} = [w_1 \cdots w_P]^T \in \mathbb{C}^{P \times 1}$ that minimizes the mean squared error between its output and the training sequence once filtered by the equivalent channel:

$$\hat{\mathbf{w}} = \arg \min_{\mathbf{w}} (\mathbf{X} \mathbf{w} - \mathcal{T} \mathbf{h})^H \mathbf{M} (\mathbf{X} \mathbf{w} - \mathcal{T} \mathbf{h}). \quad (30)$$

Note that we have introduced a weighting matrix $\mathbf{M} \in \mathbb{C}^{MN_c \times MN_c}$ to allow for generalized least squares solutions. In practice, both the weighting matrix \mathbf{M} and the channel estimator \mathbf{h} will be chosen according to the statistical model for the unknown data. Moreover, the channel estimation will depend on the beamforming and consequently (30) will have to be solved forcing a constraint on the beamvector. Let us assume that the channel estimation depends linearly on the observation $\mathbf{h} = \Phi \mathbf{X} \mathbf{w}$; if we impose a constraint on the power received from the direction of arrival of the training sequence²:

$$\mathbf{w}^H \mathbf{X}^H \Phi^H \mathbf{M} \Phi \mathbf{X} \mathbf{w} = \alpha \quad (31)$$

the beamvector can be obtained as the generalized eigenvector associated with the minimum generalized eigenvalue of the following eigensystem:

$$\mathbf{X}^H \mathbf{M} \mathbf{X} \mathbf{w} = \lambda_{\min} \mathbf{X}^H \Phi^H \mathbf{M} \Phi \mathbf{X} \mathbf{w}, \quad (32)$$

where we have assumed that matrices \mathbf{M} and Φ do not depend on the weight vector \mathbf{w} . If they did, the solution to (30) would have to be found iteratively, solving (32) and using the solution to re-initialize matrices \mathbf{M} and Φ .

Training-only Solution

If we disregard the presence of unknown data and use training-only estimator of the channel impulse response, we get to the following cost function:

$$\hat{\mathbf{w}} = \arg \min_{\mathbf{w}} \mathbf{w}^H \mathbf{X}^H \mathbf{P}_T^\perp \mathbf{X} \mathbf{w},$$

$$\text{subject to : } \mathbf{w}^H \mathbf{X}^H \mathbf{P}_T^\perp \mathbf{X} \mathbf{w} = \alpha \quad (33)$$

with $\mathbf{P}_T = \mathcal{T} (\mathcal{T}^H \mathcal{T})^{-1} \mathcal{T}^H$ and $\mathbf{P}_T^\perp = \mathbf{I}_{MN_c} - \mathbf{P}_T$. Note that we have chosen the weighting matrix \mathbf{M} as:

$$\mathbf{M}_{\text{to}}^{-1} = \frac{1}{\sigma^2} E [(\mathbf{X} \mathbf{w} - \mathcal{T} \mathbf{h})(\mathbf{X} \mathbf{w} - \mathcal{T} \mathbf{h})^H] = \mathbf{I}_{MN_c}, \quad (34)$$

²Note that, due to the structure of the channel estimator, $\mathbf{X}^H \Phi^H \mathbf{M} \Phi \mathbf{X}$ can be interpreted as an approximation to the spatial covariance matrix corresponding to the user of interest.

$$\mathbf{C}_c^M = M\mathbf{C}\mathbf{R}\mathbf{B}_c^M + M\mathbf{C}\mathbf{R}\mathbf{B}_c^M \left[\sum_{j=1}^{(M+1)N_s} \sum_{i=1}^{(M+1)N_s} \sigma^2 \left[\left(\mathbf{G}^H \mathbf{G} \right)^{-1} \right]_{j,i} \mathbf{C}^H(i) \mathbf{P}_G^\perp \mathcal{C}(j) \right] \mathbf{C}\mathbf{R}\mathbf{B}_c^M \quad (21)$$

where the expected value is taken according to the training-only assumption. The solution is obtained solving the eigen-system in (32):

$$\mathbf{X}^H \mathbf{X} \hat{\mathbf{w}}_{to} = \lambda_{\min}^c \mathbf{X}^H \mathbf{P}_T \mathbf{X} \hat{\mathbf{w}}_{to}, \quad (35)$$

which corresponds to the beamforming scheme proposed in [5, 6]. According to the training-only approach, the beamformer identifies the desired signal as the component of the received snapshots lying on the training sequence subspace and tries to null out any other contribution. In fact, the traffic signal corresponding to the user of interest is identified as an interference coming from the direction of arrival of the desired user. We will see in the next section that this causes a severe degradation of the beamformer performance, especially at high input signal to noise ratios.

Gaussian Solution

Let us now model the unknown symbols as circularly symmetric random variables. The weighting matrix \mathbf{M} is chosen accordingly:

$$\mathbf{M}_g^{-1} = \frac{1}{\sigma^2} E \left[(\mathbf{X}\mathbf{w} - \mathbf{T}\mathbf{h}) (\mathbf{X}\mathbf{w} - \mathbf{T}\mathbf{h})^H \right] = \frac{1}{\sigma^2} \mathbf{C}_x, \quad (36)$$

where expectation is taken with respect to both noise and unknown data. Concerning the channel impulse response, note that the estimator obtained for the Gaussian model can be expressed as:

$$\mathbf{h}_g = \mathbf{H}\mathbf{w}_g = \left[\mathbf{T}^H \mathbf{C}_x^{-1} \mathbf{T} \right]^{-1} \mathbf{T}^H \mathbf{C}_x^{-1} \mathbf{X}\mathbf{w}_g + O_p \left(M^{-1/2} \right), \quad (37)$$

with $O_p(\cdot)$ the in probability version of the corresponding deterministic notation. Disregarding the presence of the second term in (37), we can obtain an asymptotically valid solution for the beamvector under the Gaussian hypothesis:

$$\mathbf{X}^H \mathbf{C}_x^{-1} \mathbf{X} \hat{\mathbf{w}}_g = \lambda_{\min}^g \mathbf{X}^H \mathbf{C}_x^{-1} \mathbf{T} \left[\mathbf{T}^H \mathbf{C}_x^{-1} \mathbf{T} \right]^{-1} \mathbf{T}^H \mathbf{C}_x^{-1} \mathbf{X} \hat{\mathbf{w}}_g. \quad (38)$$

Because of the dependence of matrix \mathbf{C}_x on the beamforming weights, the solution to (38) will have to be obtained iteratively. We will show in the next section that this beamformer performs considerably better than the training-only one, thanks to the explicit modelling of the traffic signal in matrix \mathbf{C}_x .

Deterministic Solution

The proposed beamformer under the Deterministic approach can be obtained in the same way as the Gaussian solution. However, as shown before, the Conditional solution is no other than the limit of the Gaussian Solution as the input signal to noise ratio tends to infinity, namely $\mathbf{C}_x^{-1} \rightarrow \frac{1}{\sigma^2} \mathbf{P}_G^\perp$. Therefore:

$$\begin{aligned} \mathbf{M}_c &= \mathbf{P}_G^\perp, \\ \mathbf{h}_g &= \left[\mathbf{T}^H \mathbf{P}_G^\perp \mathbf{T} \right]^{-1} \mathbf{T}^H \mathbf{P}_G^\perp \mathbf{X}\mathbf{w}_g + O_p \left(M^{-1/2} \right), \end{aligned} \quad (39)$$

and the beamforming is readily obtained as:

$$\mathbf{X}^H \mathbf{P}_G^\perp \mathbf{X} \hat{\mathbf{w}}_c = \lambda_{\min}^c \mathbf{X}^H \mathbf{P}_G^\perp \mathbf{T} \left[\mathbf{T}^H \mathbf{P}_G^\perp \mathbf{T} \right]^{-1} \mathbf{T}^H \mathbf{P}_G^\perp \mathbf{X} \hat{\mathbf{w}}_c. \quad (40)$$

Again, the optimum value of $\hat{\mathbf{w}}_c$ will have to be obtained iteratively.

V. PERFORMANCE ASSESSMENT

In this section we will compare the performance of the channel estimation and beamforming algorithms in terms of relative mean squared error and output signal to noise ratio respectively.

Let us first concentrate on the channel estimation algorithms. Figure 1 represents the relative MSE for the three proposed methods, defined as:

$$RMSE = \frac{1}{M \|\mathbf{h}\|^2} \text{tr} \left(\mathbf{C}^M \right) \quad (41)$$

with \mathbf{C}^M the asymptotic covariances as defined in Section III. The channel was generated with a Vehicular-A delay profile as specified by ETSI with $L = 17$ chips and the codes were randomly selected from the large Kasami set of length $N_c = 256$. Only one traffic channel was used by the user of interest ($Q = 1$), and it was transmitted with the same power as the training sequence. The spreading factor was set to $SF = 8$.

We see that as the input signal to noise ratio increases, the relative MSE of the training-only channel estimator tends to a constant instead of following the corresponding Cramér-Rao Bound. As shown in (15), this behavior is a consequence of the presence of the traffic signal, which renders the estimator inefficient even at high signal to noise ratios. The auto-interfering effect of the traffic channel is overcome with either of the semi-blind schemes, which in fact perform very similarly (especially at high signal to noise values).

Turning now to the performance of the three beamforming methods proposed in Section IV, Figure 2 represents the output signal to noise plus interference ratios as a function of the input signal to noise ratio for a scenario consisting of three users and an antenna array of $P = 4$ elements. Along with the values obtained with Monte-Carlo simulations (asterisks), we represent the following theoretical approximation for the output signal to noise ratio³:

$$E \left[\frac{SINR(\mathbf{w})}{SINR(\mathbf{w}_{opt})} \right] \simeq 1 - \text{tr} \left[\mathbf{T}_1^H \mathbf{K}^T \mathbf{T}_1 \mathbf{T}_2 \right] \quad (42)$$

with the signal to noise plus interference ratio defined as $SINR(\mathbf{w}) = \frac{\mathbf{w}^H \mathbf{R}_d \mathbf{w}}{\mathbf{w}^H \mathbf{R}_n \mathbf{w}}$, $(\cdot)^T$ denoting Moore-Penrose pseudo-inverse and:

$$\begin{aligned} \mathbf{T}_1 &= \mathbf{I}_Q - \frac{\mathbf{w}_{opt} \mathbf{w}_{opt}^H \mathbf{R}_d}{\mathbf{w}_{opt}^H \mathbf{R}_d \mathbf{w}_{opt}} \in \mathbb{C}^{Q \times Q} \\ \mathbf{R}_d &= E \left[\mathbf{X}^H \Phi^H \mathbf{M} \Phi \mathbf{X} \right] \in \mathbb{C}^{Q \times Q} \end{aligned}$$

³A sketch of the derivation of (42) is given in the Appendix

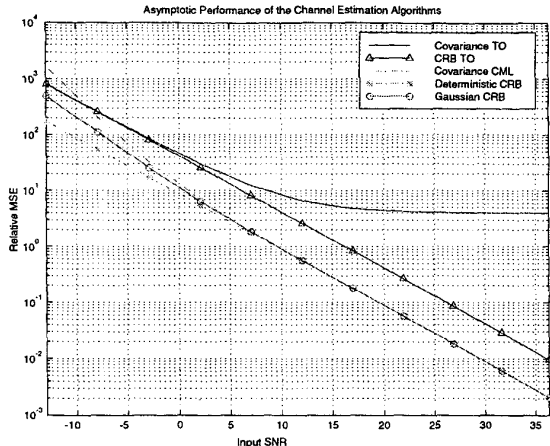


Figure 1: Performance of the channel Estimation algorithms.

$$\begin{aligned} \mathbf{K} &= [\mathbf{R}_{nd} - (1 - \lambda_{\min}^{\text{opt}})\mathbf{R}_d] \in \mathbb{C}^{Q \times Q} \\ \mathbf{R}_{nd} &= E[\mathbf{X}^H (\mathbf{M} - \Phi^H \mathbf{M} \Phi) \mathbf{X}] \in \mathbb{C}^{Q \times Q} \\ \mathbf{T}_2 &= \frac{E[(\mathbf{X} - \mathbf{T}\mathbf{H})^H \mathbf{T}_3 \mathbf{C}_x \mathbf{T}_3 (\mathbf{X} - \mathbf{T}\mathbf{H})]}{\mathbf{w}_{\text{opt}}^H \mathbf{R}_{nd} \mathbf{w}_{\text{opt}}} \in \mathbb{C}^{Q \times Q} \\ \mathbf{T}_3 &= \mathbf{M} - \lambda_{\min} \Phi^H \mathbf{M} \Phi \in \mathbb{C}^{M N_c \times M N_c}. \end{aligned} \quad (43)$$

Note that, due to the definitions of \mathbf{R}_d and \mathbf{R}_{nd} , our $SINR(\mathbf{w})$ depends on the modelling assumption for the statistics of the unknown data. In Figure 2 matrices \mathbf{R}_d and \mathbf{R}_{nd} are defined under the Gaussian assumption, but very similar results are obtained with deterministic definitions. Because of the presence of the traffic channel, the performance of the training-only beamformer levels off for increasingly high input signal to noise ratios. Once again, this behavior is avoided with the semi-blind techniques, whose performance is very close to the optimum one. Note that in Figure 2 the theoretical values for the training-only scheme are not plotted for high signal to noise ratios. This is due to the fact that (42) is only a valid approximation at $SINR(\mathbf{w})$ values close to the optimum ones. When this hypothesis is not met, the approximation may give negative values (as it is the case in Figure 2).

VI. CONCLUSIONS

This paper proposes semi-blind channel identification and beamforming techniques for the reception of pilot-aided signals in WCDMA systems. First, Maximum Likelihood Conditional and Gaussian estimation techniques are particularized to the WCDMA scenario. Then, an asymptotic approximation of these channel estimators is used to propose semi-blind narrowband beamformers which are robust to the presence of the traffic signal. The performance of the proposed algorithms has been validated theoretically and via simulations, and results show substantial improvements with respect to traditional training-only approaches. Further results on the performance of the semi-blind algorithms and a their dependence on system parameters –such as Spreading Factors, ratio of training to traffic power or signal to noise ratio– will be included in [4].

APPENDIX

In this Appendix we give a brief description of the approximations made in order to obtain expression (42). Assuming

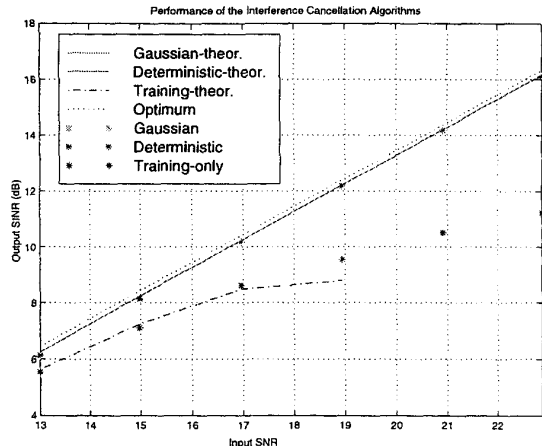


Figure 2: Performance of the Beamforming Algorithms.

that the $SINR(\mathbf{w})$ is close to its optimum value $SINR(\mathbf{w}_{\text{opt}})$ we approximate its value with the first term of its Taylor series development:

$$\frac{SINR(\mathbf{w})}{SINR(\mathbf{w}_{\text{opt}})} = 1 - \frac{\tilde{\mathbf{w}}^H \mathbf{K} \tilde{\mathbf{w}}}{\mathbf{w}_{\text{opt}}^H \mathbf{R}_{nd} \mathbf{w}_{\text{opt}}} + o_p(\|\tilde{\mathbf{w}}\|^2) \quad (44)$$

with \mathbf{K} defined in (43) and $\tilde{\mathbf{w}} = \mathbf{w} - \mathbf{w}_{\text{opt}}$. Applying classical results of eigensystem perturbation theory to (44) and disregarding the terms of higher order:

$$\frac{SINR(\mathbf{w})}{SINR(\mathbf{w}_{\text{opt}})} \simeq 1 - \text{tr} \left[\mathbf{T}_1^H \mathbf{K} \mathbf{T}_1 \frac{\mathbf{B} \mathbf{w}_{\text{opt}} \mathbf{w}_{\text{opt}}^H \mathbf{B}}{\mathbf{w}_{\text{opt}}^H \mathbf{R}_{nd} \mathbf{w}_{\text{opt}}} \right] \quad (45)$$

where $\mathbf{B} = \mathbf{X}^H (\mathbf{M} - \lambda_{\min}^{\text{opt}} \Phi^H \mathbf{M} \Phi) \mathbf{X}$. Finally applying the formula for the expectation of four Gaussian-distributed random matrices given in [7] we get to (42).

REFERENCES

- [1] E. de Carvalho and D. Slock, "Cramér-Rao bounds for semi-blind, blind a training sequence-based channel estimation," in *Proc. IEEE SP Workshop on SPAWC*, Paris (France), Apr. 1997, pp. 129–132.
- [2] E. de Carvalho and D.T.M. Slock, "Asymptotic performance of ML methods for semi-blind channel estimation," in *Proc. of the IEEE Asilomar Conf.*, 1998.
- [3] X. Mestre, M. Nájar, and J.R. Fonollosa, "Joint beamforming and channel estimation for pilot-aided WCDMA systems," in *Proc. of the IEEE International Conference on Acoustics, Speech and Signal Processing*, Istanbul (Turkey), 2000.
- [4] X. Mestre and J. Fonollosa, "ML approaches to semi-blind channel identification for WCDMA systems," To be submitted to *IEEE Transactions on Signal Processing*.
- [5] M.A. Lagunas, A. Pérez, and J. Vidal, "Joint beamforming and viterbi equalization in wireless communications," in *Proc. 31st Asilomar Conference on Signals Systems and Computers*, Pacific Grove (CA), November 1997.
- [6] F.P. Pípon, P. Chevalier, P. Villa, and J.-J. Monot, "Joint spatial and temporal equalization for channels with ISI and CCI: Theoretical and experimental results for a base station reception," in *Proc. IEEE Signal Processing Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, Paris (France), April 1997, pp. 309–312.
- [7] P.H.M. Janssen and P. Stoica, "On the expectation of the product of four matrix-valued Gaussian random variables," *IEEE Transactions on Automatic Control*, vol. 33, no. 9, pp. 867–870, Sept. 1988.