EUROPEAN COOPERATION
IN THE FIELD OF SCIENTIFIC
AND TECHNICAL RESEARCH

EURO-COST

COST 273 TD (04)090
Gothenburg, Sweden
June 9-10, 2004

SOURCE: Dip. Elettronica e Informazione
Politecnico di Milano
Italy

A subspace method for soft estimation of block-fading channels in turbo equalization

B. Moschini, M. Nicoli, U. Spagnolini
Dip. Elettronica e Informazione
Politecnico di Milano
P.zza L. da Vinci, 32
I-20133 Milano
Italy
Phone: +39 02 2399 3429
Fax: +39 02 2399 3413
Email: {nicoli, spagnoli}@elet.polimi.it

A subspace method for soft estimation of block-fading channels in turbo equalization

B. Moschini, M. Nicoli, and U. Spagnolini
Dip. di Elettronica e Informazione, Politecnico di Milano, Italy
e-mail: {nicoli, spagnoli}@elet.polimi.it

Abstract— **The performance of iterative (turbo) equalization strongly relies on the quality of the channel state information (CSI). The use of soft information for channel estimation has been largely investigated in the literature to improve the reliability of the conventional training-based estimate. In this paper we propose a new soft method for the estimation of block-fading channels based on a multi-block (MB) processing. As turbo equalization usually requires a joint processing of several data-blocks, the idea is to take advantage of this latency to improve the performance of channel estimation. The MB estimator exploits the invariance of the subspace spanned by the channel vector over a frame of blocks and it estimates the corresponding basis from different blocks. In this paper the MB method is extended to incorporate the soft information provided by the turbo equalizer. The mean square error for the soft estimate is evaluated analytically and validated by simulations. The comparison with the conventional block-by-block estimate shows the benefits of the proposed approach on the turbo equalization convergence.**

I. INTRODUCTION

Iterative (turbo) equalization is a powerful technique that can be adopted at the receiver when data, protected by an error correction code, is transmitted over a frequency selective channel causing inter-symbol (ISI) and/or co-channel interference (e.g., MAI in CDMA systems). In the turbo approach the equalization and decoding tasks are performed iteratively on the same block of received signals so as to refine the data estimate and achieve near-optimal performance with reasonable complexity [1]. The turbo structure consists of two soft-input/soft-output (SISO) detectors, one for each task, exchanging soft information on the reliability of the estimated data. For both the equalizer and the decoder stage the optimal SISO detector is the maximum a-posteriori (MAP) symbol-by-symbol estimator [2]. Yet, a better complexity/performance trade-off is obtained by replacing the MAP equalizer with a linear minimum mean square error (MMSE) filter [3] [4].

The performance and the convergence behavior of turbo-equalization depends on the reliability of the channel state information (CSI). In block transmission systems the CSI is usually obtained on a block-by-block basis from the training symbols included in each block. However, the performance of this conventional estimate is limited by the number of training symbols in the block, thus the CSI provided to the turbo equalizer may be inaccurate and it may result in considerable performance loss. In this paper a new soft iterative technique is proposed to improve the CSI reliability. The method is developed for wireless communication systems under the assumption of block-fading channel (where the fading is constant within each data block, but it varies from block to block due to the terminal mobility).

Several methods have been proposed to integrate iterative channel estimation into turbo structures [5]-[8]. The basic idea is to repeat the estimation at each iteration by using the soft information fed back from the channel decoder. Soft decision feedback allows to properly weight reliable and unreliable symbols and thus to avoid the error propagation effects that usually arise in hard decision feedback. Here we derive a soft-iterative version of the multi-block (MB) estimator [9], where the initial estimate (obtained from the pilot symbols) is refined by extending the training set with soft-valued information symbols.

Since turbo processing is usually performed on a set of $L > 1$ data blocks, we propose to take advantage of this inherent latency to improve the estimate accuracy for the slowly varying channel parameters. The MB estimator is based on the assumption that the delays in the multipath channel remain constant over the considered ensemble of blocks, while the fading amplitudes vary from block to block. In order to avoid the expensive estimation of the delays, the stationarity of the delay pattern is translated into the stationarity of the channel-vector subspace (*temporal subspace*), whose basis can be estimated by averaging the signals received in the L blocks. For increasing L the estimate error for the subspace basis becomes negligible and the error on the overall channel response depends only on the fading amplitudes (that have to be calculated block by block). With respect to [9], here the soft MB estimate improves the accuracy for the amplitudes as well by exploiting the soft information provided by the turbo equalizer.

The paper is organized as follows. Sec. II presents the signal model for a single-user block-based transmission system and the structure of the turbo receiver. Soft MB estimation is in Sec. III, the analytical evaluation of the mean square error (MSE) is in Sec. IV. Sec. V shows by simulations the advantage of the proposed method and Sec. VI gives the concluding remarks.

II. SYSTEM DESCRIPTION

A. Signal model

We consider the equivalent complex baseband model for the convolutionally coded system shown in Fig. 1. A sequence of binary information symbols $d_i \in \mathcal{D} = \{+1, -1\}$ is convolutionally encoded with code rate R. The output code bits c_i are permuted by a random interleaver, $b_{i,k} = \Pi(c_i)$, with $k = 1, 2$, and mapped into quadrature phase-shift-keying (QPSK) symbols $x_{\mathrm{s},i} = (b_{i,1} + jb_{i,2})/\sqrt{2}$ of duration T_s (the analysis can

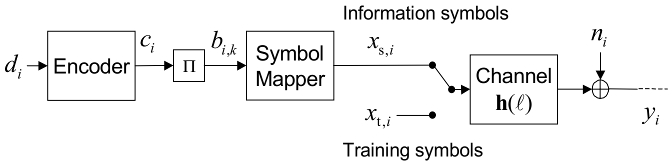

Fig. 1. Signal model for convolutionally coded system.

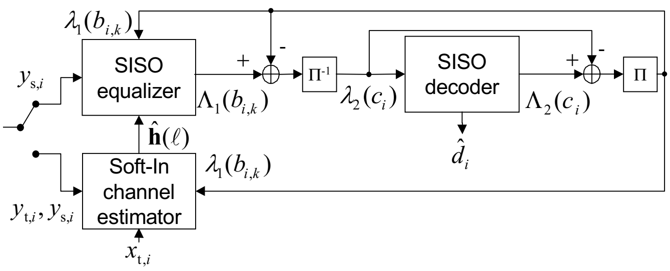

Fig. 2. Turbo receiver structure for convolutionally coded system.

be also extended to larger constellations). After mapping, the sequence $\{x_{\mathrm{s},i}\}$ is split into L blocks of N_{s} symbols each: $\mathbf{x}_{\mathrm{s}}(\ell) = [x_{\mathrm{s},0}(\ell), \ldots, x_{\mathrm{s},N_{\mathrm{s}}-1}(\ell)]$, for $\ell = 1, \ldots, L$. In order to allow channel estimation at the receiver, an uncoded training sequence $\mathbf{x}_{\mathrm{t}} = [x_{\mathrm{t},0}, \ldots, x_{\mathrm{t},N_{\mathrm{t}}-1}]$ is added as preamble within each block, yielding the overall sequence of length $N = N_{\mathrm{t}} + N_{\mathrm{s}}$: $\mathbf{x}(\ell) = [\mathbf{x}_{\mathrm{t}}, \mathbf{x}_{\mathrm{s}}(\ell)]$, with $x_i(\ell) = [\mathbf{x}(\ell)]_i$ being the ith symbol for $i = 0, \ldots, N-1$. The L blocks are then transmitted over a block-faded frequency-selective channel.

At the receiver, after matched filtering and sampling at the symbol rate, the signal measured within the ℓth block is

$$y_i(\ell) = \sum_{t=0}^{W-1} h_t(\ell) x_{i-t}(\ell) + w_i(\ell), \ \ i = 0, \ldots, N+W-2, \quad (1)$$

where the Gaussian noise $w_i(\ell)$ is white, with known variance σ_w^2: $\mathrm{E}\left[w_i^*(\ell) w_{i+k}(\ell+m)\right] = \sigma_w^2 \delta_k \delta_m$. The channel is modelled as a linear filter $\mathbf{h}(\ell) = [h_0(\ell), \ldots, h_{W-1}(\ell)]^{\mathrm{T}}$ (including transmitter/receiver filters and multipath effects) that is constant within the block interval but varying from block to block.

By gathering the samples received within the block into the vector $\mathbf{y}(\ell) = [y_0(\ell), \ldots, y_{N+W-2}(\ell)]^{\mathrm{T}}$, the discrete-time signal model (1) reduces to

$$\mathbf{y}(\ell) = \mathbf{X}(\ell) \mathbf{h}(\ell) + \mathbf{w}(\ell), \quad \ell = 1, \ldots, L. \quad (2)$$

The $(N + W - 1) \times W$ Toeplitz matrix $\mathbf{X}(\ell)$ denotes the convolution with the ℓth sequence $[\mathbf{X}(\ell)]_{m,k} = x_{m-k}(\ell)$, while the noise vector $\mathbf{w}(\ell) = [w_0(\ell), \ldots, w_{N+W-2}(\ell)]^{\mathrm{T}}$ has covariance matrix $\mathrm{E}\left[\mathbf{w}(\ell) \mathbf{w}^{\mathrm{H}}(\ell+m)\right] = \sigma_w^2 \delta_m \mathbf{I}_{N+W-1}$. The data received within the overall L blocks, $\mathbf{y} = [\mathbf{y}^{\mathrm{T}}(1) \ldots \mathbf{y}^{\mathrm{T}}(L)]^{\mathrm{T}}$, is jointly processed by the turbo receiver as described in the following section.

B. Iterative (turbo) receiver structure

The optimal receiver for the communication system depicted in Fig. 1 performs channel estimation, equalization and decoding jointly. Since the computational load required by such solution is too large for practical systems, here we perform the three tasks separately but iteratively and with exchange of reliability information. As shown in Fig. 2, the receiver structure consists of a soft-in channel estimator, a SISO equalizer and a SISO decoder separated by interleaver/de-interleaver. According to the turbo principle, channel estimation, equalization and decoding are repeated several times on the same frame of L received blocks so as to refine the estimate and ideally approach the optimal receiver's performance.

At the first step the channels $\{\mathbf{h}(\ell)\}_{\ell=1}^{L}$ are estimated from the training data contained in the L blocks. On the base of these estimates, the SISO linear equalizer [4] computes the MMSE estimate of the transmitted symbols $\hat{x}_{\mathrm{s},i}$ from the signals $\mathbf{y}$ and the a-priori log-likelihood ratio (LLR) $\lambda_1(b_{i,k}) = \log(P(b_{i,k} = +1)/P(b_{i,k} = -1))$ received as input for every code bit $b_{i,k}$. Then it outputs the extrinsic LLR $\lambda_1^{\mathrm{E}}(b_{i,k}) = \log(P(\hat{x}_{\mathrm{s},i}|b_{i,k} = +1)/P(\hat{x}_{\mathrm{s},i}|b_{i,k} = -1))$, with $\lambda_1^{\mathrm{E}}(b_{i,k}) = \Lambda_1(b_{i,k}) - \lambda_1(b_{i,k})$ and $\Lambda_1(b_{i,k})$ denoting the a-posteriori LLR. At the first step no a-priori information is available, thus $\lambda_1(b_{i,k}) = 0$.

The information extracted by the equalizer is then reversed interleaved and fed to the channel decoder as a-priori LLR, $\lambda_2(c_i) = \Pi^{-1}(\lambda_1^{\mathrm{E}}(b_{i,k}))$. Next, using this information and the knowledge of the code constraints, the SISO decoder [2] computes the a-posteriori LLR $\Lambda_2(c_i)$ for each code bit and it delivers a new extrinsic information $\lambda_2^{\mathrm{E}}(c_i) = \Lambda_2(c_i) - \lambda_2(c_i)$. At the last iteration, the a-posteriori LLR for the information bit d_i are computed as well to provide the final estimate $\hat{d}_i$.

After interleaving, the refined soft information $\lambda_1(b_{i,k}) = \Pi(\lambda_2^{\mathrm{E}}(c_i))$ generated by the decoder is fed back to both the channel estimator and the equalizer to be used as a-priori information for further iterations. In the section below we focus on the channel estimator (for the equalization task refer to [4]) and we show how the proposed MB method can exploit this soft information to improve the estimate accuracy.

III. Channel estimation

A. Subspace channel model

Within the L blocks the channel is modelled as the superposition of d paths having constant delays $\boldsymbol{\tau} = [\tau_1, ..., \tau_d]$ and block-faded amplitudes $\boldsymbol{\alpha}(\ell) = [\alpha_1(\ell), ..., \alpha_d(\ell)]^{\mathrm{T}}$:

$$\mathbf{h}(\ell) = \mathbf{G}(\boldsymbol{\tau}) \boldsymbol{\alpha}(\ell), \quad \ell = 1, .., L. \quad (3)$$

The $W \times d$ constant matrix $\mathbf{G}(\boldsymbol{\tau}) = [\mathbf{g}(\tau_1), ..., \mathbf{g}(\tau_d)]$ collects the d delayed pulse waveforms (convolution between the transmitter and receiver filters) sampled at the symbol rate, while the $d \times 1$ amplitudes $\boldsymbol{\alpha}(\ell)$ vary from block to block according to the WSSUS assumption [10], with correlation $\mathrm{E}[\boldsymbol{\alpha}(\ell+m) \boldsymbol{\alpha}(\ell)^H] = \mathrm{diag}\{A_1, ..., A_d\} \delta(m)$.

Starting from (3), the channel vector can be re-parametrized according to the subspace model proposed in [9] as

$$\mathbf{h}(\ell) = \mathbf{U} \mathbf{b}(\ell) \quad (4)$$

where $\mathbf{U}$ is a $W \times r$ orthonormal basis for the column space $\mathcal{R}(\mathbf{G}(\boldsymbol{\tau}))$ of the matrix $\mathbf{G}(\boldsymbol{\tau})$ (*temporal subspace*), while $\mathbf{b}(\ell)$ is a $r \times 1$ vector that varies from block to block (*block-dependent*). The order of the temporal diversity $r = \mathrm{rank}(\mathbf{G}(\boldsymbol{\tau}))$ is a measure of the number of resolvable delays given the bandwidth of the transmitted signal [9]. Herein it is assumed to be known (or estimated elsewhere).

In the following $\mathbf{h}(\ell)$ is estimated from the L blocks $\{\mathbf{y}(\ell)\}_{\ell=1}^{L}$ by imposing the subspace model (4). This approach allows to exploit the invariance of the delays without estimating them explicitly.

B. Training-based channel estimation

The signals received within the ℓth block can be split into the first N_t training samples $\mathbf{y}_\mathrm{t}(\ell) = [y_0(\ell), \ldots, y_{N_\mathrm{t}-1}(\ell)]^\mathrm{T}$ to be used for channel estimation and the remaining $N_\mathrm{s} + W - 1$ information-bearing samples $\mathbf{y}_\mathrm{s}(\ell) = [y_{N_\mathrm{t}}(\ell), \ldots, y_{N+W-2}(\ell)]^\mathrm{T}$:

$$\mathbf{y}(\ell) = \begin{bmatrix} \mathbf{y}_\mathrm{t}(\ell) \\ \mathbf{y}_\mathrm{s}(\ell) \end{bmatrix} = \underbrace{\begin{bmatrix} \mathbf{X}_\mathrm{t} \\ \mathbf{X}_\mathrm{s}(\ell) \end{bmatrix}}_{\mathbf{X}(\ell)} \mathbf{h}(\ell) + \underbrace{\begin{bmatrix} \mathbf{w}_\mathrm{t}(\ell) \\ \mathbf{w}_\mathrm{s}(\ell) \end{bmatrix}}_{\mathbf{w}(\ell)} \quad (5)$$

(for the sake of simplicity, the edge effects between training and information sub-blocks are discarded).

The initial estimate for the channel vectors $\{\mathbf{h}(\ell)\}_{\ell=1}^{L}$ is obtained from the training signals $\{\mathbf{y}_\mathrm{t}(\ell)\}_{\ell=1}^{L}$ by exploiting the knowledge of the pilot symbols $\mathbf{X}_\mathrm{t}$. According to the parameterization (4), the maximum likelihood estimation of $\mathbf{h}(\ell)$ yields the multi-block (MB) solution [9]:

$$\hat{\mathbf{h}}_\mathrm{MBt}(\ell) = \mathbf{R}_\mathrm{t}^{-1/2}\hat{\mathbf{P}}\mathbf{R}_\mathrm{t}^{1/2}\hat{\mathbf{h}}_\mathrm{SBt}(\ell), \quad (6)$$

where $\mathbf{R}_\mathrm{t} = \mathbf{X}_\mathrm{t}^\mathrm{H}\mathbf{X}_\mathrm{t}$ is the correlation matrix of the training sequence, $\hat{\mathbf{P}}$ is the estimate of the projector onto the temporal subspace $\mathcal{R}(\mathbf{R}_\mathrm{t}^{1/2}\mathbf{U})$ and $\hat{\mathbf{h}}_\mathrm{SB}(\ell)$ denotes the single-block (SB) unconstrained maximum likelihood estimate:

$$\hat{\mathbf{h}}_\mathrm{SBt}(\ell) = \mathbf{R}_\mathrm{t}^{-1}\mathbf{X}_\mathrm{t}^\mathrm{H}(\ell)\,\mathbf{y}_\mathrm{t}(\ell). \quad (7)$$

The projector $\hat{\mathbf{P}}$ in (6) is obtained from the r leading eigenvectors of the correlation matrix

$$\mathbf{R}_\mathrm{MBt} = \frac{1}{L}\mathbf{R}_\mathrm{t}^{1/2}\left(\sum_{\ell=1}^{L}\hat{\mathbf{h}}_\mathrm{SBt}(\ell)\,\hat{\mathbf{h}}_\mathrm{SBt}^\mathrm{H}(\ell)\right)\mathbf{R}_\mathrm{t}^{\mathrm{H}/2} \quad (8)$$

evaluated from the ensemble of SB estimates $\{\hat{\mathbf{h}}_\mathrm{SBt}(\ell)\}_{\ell=1}^{L}$.

C. Soft channel estimation

After the first equalization and decoding of the L blocks, the LLRs $\lambda_1(b_{i,k})$ about the code bits $b_{i,k}$ are available and they can be exploited to refine the initial channel estimate (6). Namely, the a-priori LLRs are used to compute the mean value $\bar{x}_i(\ell) = \mathrm{E}[x_i(\ell)]$ and the variance $\sigma_i^2(\ell) = \mathrm{Var}[x_i(\ell)] = 1 - |\bar{x}_i(\ell)|^2$ for every code symbol $x_i(\ell)$, for $i = N_\mathrm{t}, \ldots, N_\mathrm{t} + N_\mathrm{s} - 1$. Notice that for QPSK modulation the mean value of $x_{\mathrm{s},i}$ can be easily obtained as [4] $\bar{x}_{\mathrm{s},i} = (\tanh(\frac{\lambda_1(b_{i,1})}{2}) + j\tanh(\frac{\lambda_1(b_{i,2})}{2}))/\sqrt{2}$. A new estimate for the channel vector is then performed by exploiting both the N_t training symbols and the N_s soft-valued symbols as described herein.

Within each block a new SB estimate is first derived as in [8]

$$\hat{\mathbf{h}}_\mathrm{SB}(\ell) = \left(\bar{\mathbf{X}}^\mathrm{H}(\ell)\,\bar{\mathbf{X}}(\ell)\right)^{-1}\bar{\mathbf{X}}^\mathrm{H}(\ell)\,\mathbf{y}(\ell) \quad (9)$$

on the base of the soft values $\bar{\mathbf{X}}(\ell) = \mathrm{E}[\mathbf{X}(\ell)] = [\mathbf{X}_\mathrm{t}^\mathrm{T}, \bar{\mathbf{X}}_\mathrm{s}^\mathrm{T}(\ell)]^\mathrm{T}$, with $\bar{\mathbf{X}}_\mathrm{s}(\ell) = \mathrm{E}[\mathbf{X}_\mathrm{s}(\ell)]$ denoting the mean values of the information-bearing symbols within the ℓth block. Similarly to (6), the soft MB estimate is then obtained as

$$\hat{\mathbf{h}}_\mathrm{MB}(\ell) = \mathbf{R}_{\bar{x}}^{-1/2}\hat{\mathbf{P}}\mathbf{R}_{\bar{x}}^{1/2}\hat{\mathbf{h}}_\mathrm{SB}(\ell). \quad (10)$$

The correlation matrix is now defined as $\mathbf{R}_{\bar{x}} = \bar{\mathbf{X}}^\mathrm{H}(\ell)\,\bar{\mathbf{X}}(\ell) = \mathbf{R}_\mathrm{t} + \mathbf{R}_\mathrm{s}$, with $\mathbf{R}_\mathrm{s} = \bar{\mathbf{X}}_\mathrm{s}^\mathrm{H}(\ell)\,\bar{\mathbf{X}}_\mathrm{s}(\ell)$, while the projector $\hat{\mathbf{P}}$ onto the temporal subspace is evaluated from the r leading eigenvectors of the matrix:

$$\mathbf{R}_\mathrm{MB} = \frac{1}{L}\mathbf{R}_{\bar{x}}^{1/2}\left(\sum_{\ell=1}^{L}\hat{\mathbf{h}}_\mathrm{SB}(\ell)\,\hat{\mathbf{h}}_\mathrm{SB}^\mathrm{H}(\ell)\right)\mathbf{R}_{\bar{x}}^{\mathrm{H}/2}. \quad (11)$$

We assume that the data symbols are uncorrelated and N_s is large enough so that $\mathbf{X}_\mathrm{s}^\mathrm{H}(\ell)\,\mathbf{X}_\mathrm{s}(\ell) \approx N_\mathrm{s}\mathbf{I}_W$. The correlation matrix $\mathbf{R}_\mathrm{s}$ is therefore approximated as $\mathbf{R}_\mathrm{s} \approx \tilde{N}_\mathrm{s}\mathbf{I}_W$, with

$$\tilde{N}_\mathrm{s} = N_\mathrm{s}(1 - \bar{\sigma}^2) \quad (12)$$

$$\bar{\sigma}^2 = \frac{1}{LN_\mathrm{s}}\sum_{i,\ell}\sigma_i^2(\ell) = 1 - \frac{1}{LN_\mathrm{s}}\sum_{i,\ell}|\bar{x}_i(\ell)|^2 \quad (13)$$

and $\mathbf{R}_{\bar{x}} = \mathbf{R}_\mathrm{t} + \tilde{N}_\mathrm{s}\mathbf{I}_W$ is considered independent of the block.

Notice that $\tilde{N}_\mathrm{s}$ represents the *effective* number of known data symbols that can be used in each block for channel estimation. If the estimated symbols are unreliable (i.e., at the first iterations of the turbo processing for moderate SNR), it is $\tilde{N}_\mathrm{s} = 0$, $\bar{\mathbf{X}}_\mathrm{s}(\ell) = 0$, and the soft MB estimate (10) equals the MB training based estimate (6). On the other hand, for perfect a-priori information (i.e., after a large enough number of iterations, if the turbo approach converges) it is $\tilde{N}_\mathrm{s} = N_\mathrm{s}$, $\bar{\mathbf{X}}_\mathrm{s}(\ell) = \mathbf{X}_\mathrm{s}(\ell)$ and therefore the soft estimate equals the training-based estimate that would be obtained from a virtual training sequence of $N_\mathrm{t} + N_\mathrm{s}$ symbols.

IV. Performance analysis

In the following the MSE for the soft iterative channel estimate is derived analytically by considering the errors $\Delta x_i(\ell) = x_i(\ell) - \bar{x}_i(\ell)$ as uncorrelated, with zero mean and variance $\bar{\sigma}^2$ equal to (13), also uncorrelated from the noise samples $w_i(\ell)$. We further assume that $L \to \infty$, which implies a perfect knowledge of the temporal subspace.

The estimate error $\Delta\mathbf{h}(\ell) = \hat{\mathbf{h}}(\ell) - \mathbf{h}(\ell)$ for the SB ($\Delta\mathbf{h}_\mathrm{SB}(\ell)$) and MB ($\Delta\mathbf{h}_\mathrm{MB}(\ell)$) methods can be written as

$$\Delta\mathbf{h}_\mathrm{SB}(\ell) = \mathbf{R}_{\bar{x}}^{-1}\bar{\mathbf{X}}^\mathrm{H}(\ell)\left[\Delta\mathbf{X}\mathbf{h}(\ell) + \mathbf{w}(\ell)\right], \quad (14)$$

$$\Delta\mathbf{h}_\mathrm{MB}(\ell) = \mathbf{R}_{\bar{x}}^{-1/2}\mathbf{P}\mathbf{R}_{\bar{x}}^{1/2}\Delta\mathbf{h}_\mathrm{SB}(\ell), \quad (15)$$

where $\mathbf{P}$ is the true projector onto $\mathcal{R}(\mathbf{R}_{\bar{x}}^{1/2}\mathbf{U})$. Notice that the first term in (14) is due to the symbol estimate error

TABLE I
MSE OF SOFT ITERATIVE SB AND MB ESTIMATES.

Est.	Corr. seq.	Uncorr. seq.
$0 \leq \bar{\sigma}^2 \leq 1$		
SB	$\sigma_w^2\,\mathrm{tr}(\mathrm{R}_{\bar{x}}^{-1}) + \sigma_u^2 \tilde{N}_\mathrm{s}\,\mathrm{tr}(\mathrm{R}_{\bar{x}}^{-2})$	$\frac{W}{\tilde{N}}(\sigma_w^2 + \sigma_u^2 \frac{\tilde{N}_\mathrm{s}}{\tilde{N}})$
MB	$\sigma_w^2\,\mathrm{tr}(\Psi) + \sigma_u^2 \tilde{N}_\mathrm{s}\,\mathrm{tr}(\Psi^2)$	$\frac{r}{\tilde{N}}(\sigma_w^2 + \sigma_u^2 \frac{\tilde{N}_\mathrm{s}}{\tilde{N}})$
$\bar{\sigma}^2 = 0$		
SB	$\sigma_w^2\,\mathrm{tr}(\mathrm{R}_x^{-1})$	$\sigma_w^2 \frac{W}{N_\mathrm{t}+N_\mathrm{s}}$
MB	$\sigma_w^2\,\mathrm{tr}(\mathrm{R}_x^{-1/2}\mathrm{P}\mathrm{R}_x^{-\mathrm{H}/2})$	$\sigma_w^2 \frac{r}{N_\mathrm{t}+N_\mathrm{s}}$
$\bar{\sigma}^2 = 1$		
SB	$\sigma_w^2\,\mathrm{tr}(\mathrm{R}_\mathrm{t}^{-1})$	$\sigma_w^2 \frac{W}{N_\mathrm{t}}$
MB	$\sigma_w^2\,\mathrm{tr}(\mathrm{R}_\mathrm{t}^{-1/2}\mathrm{P}\mathrm{R}_\mathrm{t}^{-\mathrm{H}/2})$	$\sigma_w^2 \frac{r}{N_\mathrm{t}}$

$\Delta\mathbf{X} = \mathbf{X}(\ell) - \bar{\mathbf{X}}(\ell) = [\mathbf{0}_{W \times N_\mathrm{t}}\ \Delta\mathbf{X}_\mathrm{s}^\mathrm{T}(\ell)]^\mathrm{T}$. We can model this term as an equivalent white noise, $\mathbf{u}(\ell) = \Delta\mathbf{X}_\mathrm{s}(\ell)\,\mathbf{h}(\ell)$, with zero mean, covariance $\Phi = \mathrm{E}[\mathbf{u}(\ell)\,\mathbf{u}^\mathrm{H}(\ell)] = \sigma_u^2 \mathbf{I}_{N_\mathrm{s}}$ and power $\sigma_u^2 = \bar{\sigma}^2 \mathrm{E}_h$, where $\mathrm{E}_h = \mathrm{E}[||\mathbf{h}(\ell)||^2]$ denotes the channel energy (this assumption is fulfilled for WSSUS channels having symbol-spaced delays and Nyquist impulse waveform). The covariance matrix $\mathrm{Cov}(\hat{\mathbf{h}}) = \mathrm{E}[\Delta\mathbf{h}(\ell)\,\Delta\mathbf{h}^\mathrm{H}(\ell)]$, averaged over fading and noise, is thus given by

$$\mathrm{Cov}(\hat{\mathbf{h}}_\mathrm{SB}) = \sigma_w^2 \mathbf{R}_{\bar{x}}^{-1} + \tilde{N}_\mathrm{s}\sigma_u^2 \mathbf{R}_{\bar{x}}^{-2}, \quad (16)$$

$$\mathrm{Cov}(\hat{\mathbf{h}}_\mathrm{MB}) = \sigma_w^2 \Psi + \tilde{N}_\mathrm{s}\sigma_u^2 \Psi^2, \quad (17)$$

where $\Psi = \mathbf{R}_{\bar{x}}^{-1/2}\mathbf{P}\mathbf{R}_{\bar{x}}^{-\mathrm{H}/2}$. The MSE can be derived from (16)-(17) as MSE=$\mathrm{tr}(\mathrm{Cov}(\hat{\mathbf{h}}))$ yielding the results in Table I.

The MSE largely simplifies for uncorrelated training sequences, i.e. for $\mathbf{R}_\mathrm{t} = N_\mathrm{t}\mathbf{I}_W$ and thus $\mathbf{R}_{\bar{x}} = \tilde{N}\mathbf{I}_W$, as shown in the third column of Table I. As expected, in this case the performance depends only on the ratio between the number of channel unknowns and the number of *effective* training symbols within each block ($\tilde{N} = N_\mathrm{t} + \tilde{N}_\mathrm{s}$). The number of unknowns is W for the SB estimator, while for the MB estimator it is reduced to the number r of block-dependent amplitudes $\mathbf{b}(\ell)$ [9], as the basis $\mathbf{U}$ (as well as the projector $\mathbf{P}$) is perfectly estimated for $L \to \infty$. Table I also shows the performance for two extreme conditions: missing prior information (i.e., at the first iteration) for $\bar{\sigma}^2 = 1$, $\tilde{N} = N_\mathrm{t}$, and $\mathbf{R}_{\bar{x}} = \mathbf{R}_\mathrm{t}$ (rows 5-7); perfect prior information (i.e., close to the convergence of the iterative approach) for $\bar{\sigma}^2 = 0$, $\tilde{N} = N_\mathrm{t} + N_\mathrm{s}$, and $\mathbf{R}_{\bar{x}} = \mathbf{R}_x = \mathbf{R}_\mathrm{t} + \mathbf{X}_\mathrm{s}^\mathrm{H}(\ell)\,\mathbf{X}_\mathrm{s}(\ell) = \mathbf{R}_\mathrm{t} + N_\mathrm{s}\mathbf{I}_W$ (rows 8-10).

V. SIMULATION RESULTS

The performance for the SB and MB estimation methods described in Sec. III are compared by simulating the following block-transmission system. A frame of 4000 information bits (including the tail bits) is coded by a 4-state convolutional code with generators $(7,5)_o$ and it is then permuted by a random interleaver. Next, the code bits are mapped into 4000 QPSK symbols and arranged into $L = 20$ blocks of $N_s = 200$ symbols each. A training sequence of $N_t = 31$ QPSK symbols is obtained as $x_{\mathrm{t},i} = j^i t_i$ from the sequence $\{t_i\}$ used in [8] and it is added in each block. To avoid border effects a cyclic prefix of $W - 1$ symbols is used yielding an overall training length of $N_t + W - 1 = 46$. The L blocks are then transmitted over a block-faded Rayleigh channel with temporal support $W = 16$, generated by 2 clusters of scatterers having 3 resolvable paths each ($r = 6$) and exponential power delay profile.

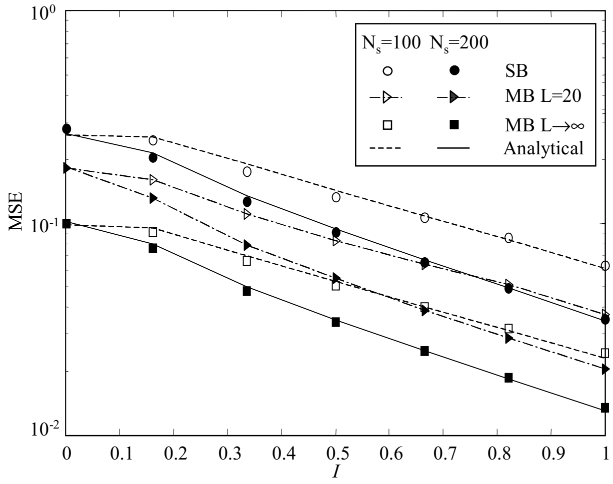

Fig. 3. MSE of the SB-MB soft estimate vs. the mutual information $\mathcal{I}$.

Figures 3-4 compare the MSE of the soft SB and MB estimates for different values of: number of blocks L used for the projector evaluation in the MB estimate; number of soft-valued symbols N_s used for channel estimation; mutual information $\mathcal{I} = \mathcal{I}(b_{i,k}, \lambda_1(b_{i,k}))$ between the transmitted code bits $b_{i,k}$ and the a-priori information $\lambda_1(b_{i,k})$ used as input [10]. Notice that the soft SB estimate here is equivalent to the method proposed in [8]. According to [11], the a-priori information $\lambda_1(b_{i,k})$ is modelled as Gaussian and E_b/N_0 (defined for the coded system) is set to 3dB. The simulated MSE values (markers) are also compared with the analytical results (continuous lines) of Table I.

Figure 3 shows the MSE for varying $\mathcal{I}$ (or, equivalently, for varying $\bar{\sigma}^2$) and for $N_\mathrm{s} = \{100{,}200\}$. It can be seen that the soft iterative channel estimate becomes more accurate for increasing $\mathcal{I}$, from $\mathcal{I} = 0$ (i.e., estimation from training symbols only, $\bar{\sigma}^2 = 1$) to $\mathcal{I} = 1$ (i.e., estimation from the overall block of N known symbols, $\bar{\sigma}^2 = 0$). The maximum gain with respect to the training based approach (MSE$_\mathrm{t}$) is reached for $\mathcal{I} = 1$ and it is MSE$_\mathrm{t}$/MSE= N/N_t as confirmed by simulations. For every value of mutual information, we observe that the soft MB estimate outperforms the soft SB estimate by a factor MSE$_\mathrm{SB}$/MSE$_\mathrm{MB}$ $\approx W/r = 4.26$dB (the training sequence is nearly uncorrelated).

Fig. 4 shows the MSE of the channel estimates vs. the number of data symbols N_s used for the soft-decision feedback. In Fig. 4-a two different values of mutual information are considered, $\mathcal{I} = \{0.5, 1\}$. For every value of N_s the performance gain between the soft and the training-based (MSE$_\mathrm{t}$) estimate is MSE$_\mathrm{t}$/MSE= $\tilde{N}/N_\mathrm{t}$. Fig. 4-b compares the performance of the SB estimate and the MB estimate with $L = \{10{,}20{,}40{,}100\}$ (here it is $\mathcal{I} = 1$). It can be seen that for increasing L the performance of the MB method converges to the analytical MSE given in Table I for $L \to \infty$.

In Fig. 5 we evaluate the BER performance of the complete turbo receiver with soft channel estimation, sliding win-

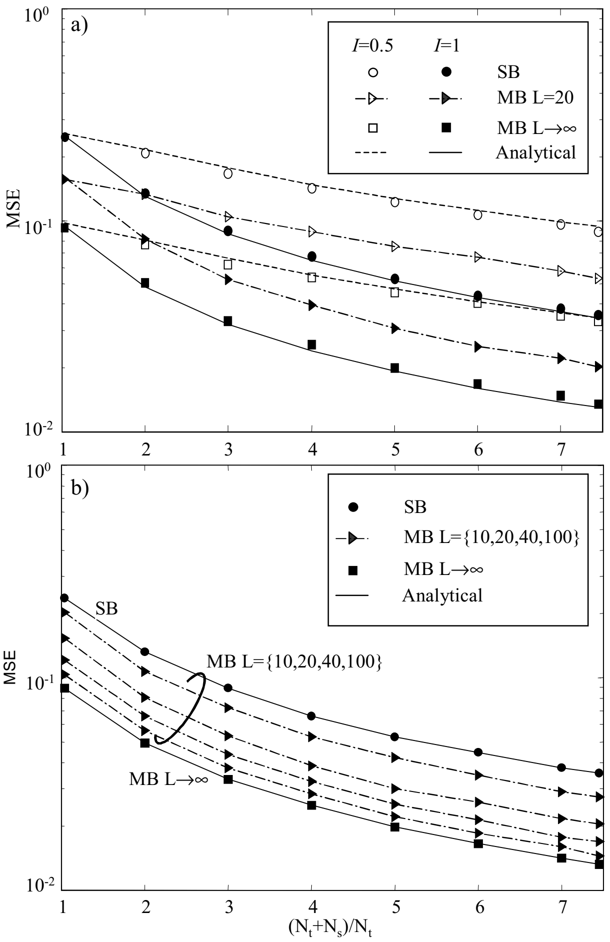

Fig. 4. MSE of the SB and MB soft iterative channel estimate for varying N_s.

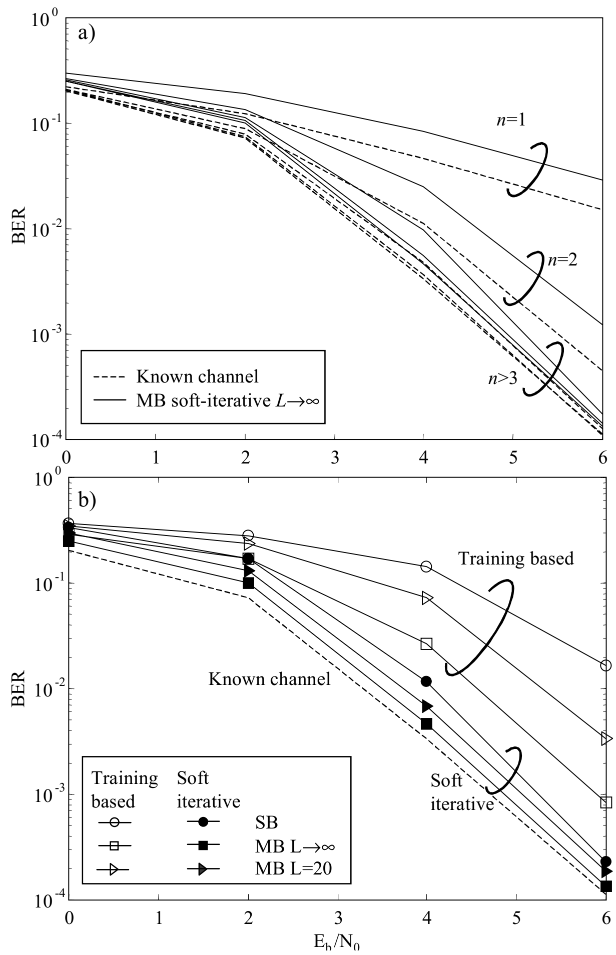

Fig. 5. BER performance for the turbo equalizer vs. E_b/N_0 for varying number n of iterations (top) and at the 5th iteration (bottom).

dow MMSE SISO equalization [4] and log-MAP SISO decoding [2]. Fig. 5-a compares the performance of the turbo equalizer with MB soft channel estimation (for known projector $\mathbf{P}$ or $L \to \infty$) with the case of known channel. It can be noticed that $n = 5$ iterations are enough for the MB iterative channel estimator to approach the performance of known channel. Fig. 5-b shows the performance of the turbo equalizer with SB and MB estimation after 5 iterations. Both training based and soft-iterative approaches are used. For the MB method the projector $\mathbf{P}$ onto the temporal subspace is estimated either from $L = 20$ and $L \to \infty$ blocks. We observe that the MB soft method outperforms both the training-based and the SB soft methods. Its performance at the 5th iteration is close to that obtained by the turbo equalizer for known channel.

VI. Concluding remarks

This paper proposes the integration of MB channel estimation for block-fading channels [9] with turbo equalization. The MB method exploits the invariance of the temporal subspace across blocks and it estimates the channel using the soft statistics fed back by the decoder. The analytical evaluation of the MSE for the channel estimate and the simulation results on the BER for the complete iterative receiver show that the proposed method improves the performance of turbo equalization.

References

[1] C. Douillard, M. Jézéquel, C. Berrou, "Iterative correction of intersymbol interference, turbo-equalization", *Eur. Trans. Telecommun.*, Vol. 6, pp. 507-511, Sept.-Oct. 1995.

[2] L. R. Bahl, J. Cocke, F. Jelinek, J. Ravin, "Optimal decoding of linear codes for minimizing symbol error rate", *IEEE Trans. Infor. Theory*, Vol. 20, pp. 284-287, Mar. 1974.

[3] X. Wang, V. Poor, "Iterative (turbo) soft interference cancellation and decoding for coded CDMA", *IEEE Trans. Commun.*, vol. 47, pp. 1046-1061, July 1999.

[4] M. Tüchler, A. C. Singer, "Minimum mean squared error equalization using *a priori* information", *IEEE Trans. Signal Processing*, Vol. 50, pp. 673-683, Mar. 2002.

[5] A. O. Berthet, B. S. Unal, R. Visoz, "Iterative decoding of convolutionally encoded signals over multipath Rayleigh fading channels", *IEEE Journal on Selected Areas in Commun.*, Vol. 19, N. 9, pp. 1729 -1743, Sept. 2001.

[6] S. Song, A. Singer, K-M. Sung, "Turbo equalization with an unknown channel", *Proc. IEEE ICASSP* 2002, Vol. 3 , pp. 2805-2808.

[7] K.-D. Kammeyer, V. Kuhn, T. Petermann, "Blind and non-blind turbo estimation for fast fading GSM channels", *IEEE Journal on Selected Areas in Commun.*, Vol. 19, pp. 1718-1728, Sept. 2001.

[8] M. Tuchler, R. Otnes, A. Schmidbauer, "Performance of soft iterative channel estimation in turbo equalization", *Proc. IEEE ICC 2002*, pp. 1858-1862.

[9] M. Nicoli, O. Simeone, U. Spagnolini, "Multi-slot estimation of frequency-selective fast-varying channels", *IEEE Trans. Commun*, Vol. 51, No. 8, pp. 1337-1347, August 2003.

[10] J. Proakis, *Digital Communications*, 3rd ed. McGraw-Hill, 1995.

[11] S. ten Brink, "Convergence behavior of iteratively decoded parallel concatenated codes", *IEEE Trans. Commun.*, Vol. 49, N. 10, pp. 1727-1737, Oct. 2001.