Successive Interference Cancellation with SISO Decoding and EM Channel Estimation

Mari Kobayashi† and Joseph Boutros† and Giuseppe Caire⋆

April 26, 2001

† Ecole Nationale Supérieure des Télécommunications de Paris,
64 rue Barrault – 75634 Paris CEDEX, France.
Email: mari@wcs.sony.co.jp, boutros@enst.fr.
⋆ Institut Eurecom,
2229 Route des Crétes, B.P. 193
06904 Sophia-Antipolis CEDEX, France.
Email: caire@eurecom.fr

Abstract

We derive a low-complexity receiver scheme for joint multiuser decoding and parameter estimation of CDMA signals. The resulting receiver processes the users serially and iteratively, and makes use of *soft-in soft-out* single-user decoders, of *soft interference cancellation* and of *expectation-maximization* parameter estimation as the main building blocks.

Computer simulations show that the proposed receiver achieves near single-user performance at very high channel load (number of users per chip) and outperforms conventional schemes with similar complexity.

Keywords: Interference cancellation, joint data detection and parameter estimation.

1 Introduction

Among the several multiuser detection schemes proposed for CDMA [1], Serial and Parallel Interference Cancellation (SIC and PIC) are particularly attractive because they process directly the output of a bank of single-user matched filters (SUMF). The receiver front-end is identical to that of conventional detection. Therefore, these methods can be seen as an "add-on" post-processing to enhance the performance of a conventional base-station receiver when particularly high channel load is needed, and can be applied easily to either *short* and *long* spreading sequence formats [2, 3, 4].

The main performance limitation of SIC/PIC schemes are: 1) error propagation caused by feeding back erroneous symbol decisions; 2) imperfect interference cancellation due to non-ideal knowledge of channel parameters (e.g., the complex amplitudes and delays of the users' multipath channels). In this work, we propose a receiver scheme which handles successfully both problems.

SIC is both simpler and more robust than PIC with respect to error propagation, since users can be ranked according to their signal-to-interference plus noise ratio (SINR) and decoded in sequence [5, 6, 7, 8]. Hence, we focus on SIC schemes. In early works [6, 5], SIC is applied to uncoded transmission and hard decisions are used at each stage to remove the already detected users from the received signals. In order to prevent error propagation, the use of soft (or *partial*) interference cancellation and iterative SIC schemes has been proposed in different forms and by different authors (see for example [9, 10, 8]). More recently, the SIC approach has been combined with channel coding and Soft-In Soft-Out (SISO) decoding [11]. The number of works in this direction is overwhelming. Without the ambition of being exhaustive, we refer to [12, 13, 14, 15, 16, 17, 8, 18, 19, 20, 21, 22, 23] and references therein. A common feature of these algorithms is that single-user SISO decoders provide at each iteration an estimate of the *a posteriori* probabilities (APP) for the user code symbols, which are used to form a soft estimate of interference to be subtracted from the received signal. In this way, the contribution of a user is effectively subtracted from the signal only if its symbol decisions are sufficiently reliable.

A unified framework to iterative multiuser joint decoding based on factor-graphs and sum-product algorithm [24] is provided in [25]. In this framework, almost all algorithms previously proposed (notably, those of [12] and of [23]) have been re-derived in a simple direct way. Moreover, as a consequence of the sum-product approach, it is found that *extrinsic* (EXT) probabilities [26] rather than APPs should be fed back to form the soft interference estimate. As confirmed experimentally by [27], APP-based soft interference cancellation yields a biased residual interference term which tends to cancel the useful signal, and the APP-based algorithms of [12, 23] attain a worse overall spectral efficiency than their EXT-based counterparts derived and analyzed in [25].

In order to reduce parameter estimation errors, iterative SIC schemes can be naturally coupled with iterative parameter estimation in order to (hopefully) improve the estimates with the iterations, as long as the signal is "cleaned-up" from interference (see for example [28]). In [29] the trade-off between the number of users per chip (channel load) and the amount of training symbols is investigated in a general iterative joint decoder which re-estimates the channel parameters at each iteration.

We propose a low-complexity iterative soft-SIC algorithm for joint data detection and channel parameter estimation. The main building blocks of our receiver are SISO single-user decoders, soft interference cancellation stages and a channel parameter estimation updating step which is formally equivalent to one step of the Expectation-Maximization (EM) algorithm [30, 31]. The key idea to achieve polynomial complexity in the number of users is to apply EM "locally", i.e., instead of using the true a posteriori distribution of the missing data given the observation and the current parameter estimate, we use the product distribution induced by the a posteriori marginal (symbol-by-symbol) probabilities output by the SISO decoders at each receiver iteration.

We restrict our treatment to synchronous CDMA with frequency non-selective propagation channels. Users are synchronous at the chip, symbol and frame level and encoding and decoding is performed frame by frame. We assume also that the channel parameters remain constant over each frame. The reason for adopting this simple model is twofold: on one hand, this model allows the development of the algorithm in a simple and clear way, on the other hand frame-synchronous transmission with piecewise constant channel parameters is quite realistic in systems like UMTS-TDD [3], applied to indoor and picocells with slowly moving user terminals. Generalization to asynchronous transmission and continuously time-varying multipath channels is left as an interesting topic for future work.

Related work can be found, for example, in [32] (see also [31] and references therein), where EM channel estimation is applied to SIC in an uncoded system. In [33], joint parameter estimation and data detection in a multiuser multipath environment is tackled by using an alternating maximization strategy and EM is used to solve the parameter estimates updating step. In [34, 35, 36], the EM approach is applied to the joint data detection and parameter estimation in a single-user space-time coded system. In [37], the SAGE algorithm [38] is applied to joint MAP symbol-by-symbol detection and parameter estimation in an asynchronous CDMA system. The algorithms obtained in [37] have exponential complexity in the number of users as the SAGE is not applied "locally" (as opposed to what we do here). Classical references on the application of EM in communications problems are [39], where EM is applied to parameter estimation in digital receivers, and [40], where several iterative multiuser schemes for uncoded CDMA (with perfectly known parameters) are derived as applications of EM and SAGE.

The paper is organized as follows. In Section 2 the synchronous CDMA signal model is presented. In Section 3 we derive the proposed receiver structure In Section 4 we present some numerical results and in Section 5 we summarize our conclusions.

Notation conventions:

- Let $\mathbf{A}$ be a matrix, then $\mathbf{a}_n, \mathbf{a}^k$ and $a_{k,n}$ (or equivalently $[\mathbf{A}]_{k,n}$) denote the n-th column, the k-th row and the (k,n)-th element of $\mathbf{A}$.
- $\mathbf{z} \sim \mathcal{N}_{\mathbb{C}}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ indicates that the random vector $\mathbf{z}$ is complex circularly-symmetric jointly Gaussian with mean $E[\mathbf{z}] = \boldsymbol{\mu}$ and covariance $E[(\mathbf{z}-\boldsymbol{\mu})(\mathbf{z}-\boldsymbol{\mu})^H] = \boldsymbol{\Sigma}$.
- The superscript ^{H} indicates Hermitian transpose.
- $A \propto B$ indicates that A and B differ by a multiplicative term.
- $A \doteq B$ indicates that A and B differ by an additive term.
- Probability density functions (pdf) are denoted by $p(\cdot)$ and probability mass functions (pmf) are denoted by $\Pr(\cdot)$.

2 System Model

We consider the uplink of a coded direct-sequence CDMA system with synchronous transmission over frequency-non-selective channels and Nyquist chip-shaping pulses [41]. The system is frame-oriented, i.e., encoding and decoding is performed frame-by-frame and users are synchronous also at the frame level. In each frame, the complex baseband equivalent discrete-time signal originated by sampling at the chip rate the output of a chip-matched filter is given by [1]

$$\begin{cases} \mathbf{Y} = \mathbf{S}\mathbf{W}\mathbf{X} + \mathbf{N} & \text{Data transmission phase} \\ \mathbf{Y}^{(t)} = \mathbf{S}\mathbf{W}\mathbf{X}^{(t)} + \mathbf{N}^{(t)} & \text{Training phase} \end{cases} \tag{1}$$

where:

- $\mathbf{Y} \in \mathbb{C}^{L\times N}$ and $\mathbf{Y}^{(t)} \in \mathbb{C}^{L\times T}$ are the arrays of received signal samples in the data and training phases, respectively.

- $\mathbf{N} \in \mathbb{C}^{L\times N}$ and $\mathbf{N}^{(t)} \in \mathbb{C}^{L\times T}$ are the corresponding arrays of noise samples, assumed complex circularly-symmetric Gaussian i.i.d. $\sim \mathcal{N}_{\mathbb{C}}(0, N_0)$.
- $\mathbf{S} \in \mathbb{C}^{L\times K}$ contains the user spreading sequences by columns.
- $\mathbf{W} = \text{diag}(w_1, \ldots, w_K)$ contains the user complex amplitudes w_k.
- $\mathbf{X} \in \mathbb{C}^{K\times N}$ is the array of transmitted code symbols.
- $\mathbf{X}^{(t)} \in \mathbb{C}^{K\times T}$ is the array of transmitted training symbols (known at the receiver).
- N, T, L and K denote the code block length and the training sequence length (in symbols), the spreading factor (number of chips per symbol) and the number of users, respectively.

The total frame length in symbols is equal to $N + T$. Since the channel amplitudes remain constant over the whole frame and the system is synchronous, the position of training symbols in the frame is irrelevant and arbitrary. [1] With reference to the above model and to our notation conventions, $\mathbf{s}_k, \mathbf{x}^k, \mathbf{y}_n$ and $\mathbf{x}_n$ denote the k-th user spreading sequence, the k-th user code word, the received signal vector in the n-th symbol interval and the transmitted symbol vector in the n-th symbol interval, respectively. The user spreading sequences are normalized such that $|\mathbf{s}_k|^2 = 1$ for all k. Hence, the signal-to-noise ratio (SNR) of user k is given by $\text{SNR}_k = |w_k|^2/N_0$. The corresponding system block-diagram is shown in Fig. 1.

At each frame, each user encodes a sequence of information bits into a code word $\mathbf{x}^k \in \mathcal{C}_k$, where $\mathcal{C}_k$ is the code book of user k, defined over a given complex signal set (e.g., a PSK or QAM constellation). In this paper we consider non-systematic non-recursive convolutional codes with trellis termination, mapped onto BPSK, so that $x_{k,n} \in \{-1, +1\}$. Each code word is independently interleaved before transmission.

3 Iterative joint data detection and parameter estimation

Without loss of generality, we assume that the user decoding order at each iteration is $k = 1, 2, \ldots, K$. Decoding of user k at iteration m in the soft-SIC receiver is based on the

[1] In practice, for slowly-varying frequency-selective channels it is convenient to place the training phase in the middle of each frame [3].

observed signal sequence

$$z_{k,n}^{(m)} = \underbrace{\frac{1}{\widehat{w}_k^{(m)}} \mathbf{s}_k^H \mathbf{y}_n}_{\text{SUMF output}} \underbrace{- \sum_{j=1}^{k-1} \mathbf{s}_k^H \mathbf{s}_j \frac{\widehat{w}_j^{(m)}}{\widehat{w}_k^{(m)}} \widehat{x}_{j,n}^{(m)}}_{\text{current iteration}} \underbrace{- \sum_{j=k+1}^{K} \mathbf{s}_k^H \mathbf{s}_j \frac{\widehat{w}_j^{(m)}}{\widehat{w}_k^{(m)}} \widehat{x}_{j,n}^{(m-1)}}_{\text{previous iteration}} \tag{2}$$

for $n = 1, \ldots, N$, where $\{\widehat{w}_j^{(m)} : j = 1, \ldots, K\}$ are estimates of the user amplitudes at iteration m, $\{\widehat{x}_{j,m}^{(m)} : j = 1, \ldots, k-1\}$ are estimates of the user symbols already decoded at iteration m and $\{\widehat{x}_{j,m}^{(m-1)} : j = k+1, \ldots, K\}$ are estimates of the user symbols provided by the previous iteration, since these users are not yet decoded at iteration m.

Decoding is performed by a SISO decoder, which in the case of convolutional codes can be implemented efficiently by the forward-backward BCJR algorithm [42]. Let $p(z_{k,n}^{(m)} | x_{k,n} = a)$ be the conditional pdf of $z_{k,n}$ given $x_{k,n} = a$, with $a \in \{-1, +1\}$. The SISO decoder for user k produces a marginal EXT pmf for $x_{k,n}$, given by

$$\text{EXT}_{k,n}^{(m)}(a) \propto \sum_{\mathbf{c} \in \mathcal{C}_k : c_n = a} \prod_{\ell \neq n} p(z_{k,\ell}^{(m)} | x_{k,\ell} = c_\ell) \tag{3}$$

where the normalization $\text{EXT}_{k,n}^{(m)}(+1) + \text{EXT}_{k,n}^{(m)}(-1) = 1$ is enforced. The corresponding APP is given by

$$\text{APP}_{k,n}^{(m)}(a) \propto p(z_{k,n}^{(m)} | x_{k,n} = a) \text{EXT}_{k,n}^{(m)}(a) \tag{4}$$

with again the normalization $\text{APP}_{k,n}^{(m)}(+1) + \text{APP}_{k,n}^{(m)}(-1) = 1$.

Assuming that $z_{k,n}$ is conditionally (marginally) circularly-symmetric complex Gaussian given $x_{k,n}$, the pdf $p(z_{k,n}^{(m)} | x_{k,n} = a)$ can be approximated as

$$p(z_{k,n}^{(m)} | x_{k,n} = a) \propto \exp\left(- \frac{|z_{k,n} - a|^2}{\nu_k^{(m)}} \right) \tag{5}$$

where $\nu_k^{(m)} = E[|z_{k,n}^{(m)} - x_{k,n}|^2]$ is the residual interference plus noise variance, which is independent of n under mild uniformity conditions on the user codes [25].

The SISO decoders output also APPs for the information bits, which will be used for final symbol-by-symbol decisions in the last iteration. For simplicity, we assume that the total number of iterations M is fixed for all users. In practice, M should be optimized according to the SNR and channel load K/L. Also, some dynamic stopping criterion might be used in order to minimize the number of iterations. We leave this interesting topic for future work.

Next, we address the estimation of the residual interference plus noise variance $\nu_k^{(m)}$, the estimation of the code symbols $x_{k,n}$ and the estimation of the user amplitudes w_k

used in the soft-SIC (equation (2)). We also address the initialization of the receiver with training-based parameter estimation and some methods to combine training-based and EM-based estimation. Finally, we summarize the resulting soft-SIC receiver with joint data detection and parameter estimation.

3.1 Estimation of the residual interference plus noise variance

The variance $\nu_k^{(m)}$ is unknown, and must be estimated on-line *before* each SISO decoding step. Let $\zeta_{k,n}^{(m)} = z_{k,n}^{(m)} - x_{k,n}$ denote the residual interference plus noise term in (2). A simple estimator for $\nu_k^{(m)}$ is given by [2]

$$\widehat{\nu}_k^{(m)} = \frac{1}{N}\sum_{n=1}^{N} |z_{k,n}^{(m)}|^2 - 1 \tag{6}$$

Beside its simplicity, the motivations for using (6) to estimate $\nu_k^{(m)}$ are:

1. If $\zeta_{k,n}^{(m)}$ and $x_{k,n}$ are uncorrelated, then $\widehat{\nu}_k^{(m)}$ is an unbiased estimator.

2. If $x_{k,n}$ is i.i.d., uniformly distributed on $\{-1,+1\}$ (as in our case), $\zeta_{k,m}^{(m)}$ is i.i.d. $\sim \mathcal{N}_{\mathbb{C}}(0,\nu_k^{(m)})$, and $x_{k,n}, \zeta_{k,n}^{(m)}$ are uncorrelated, then the error variance of $\widehat{\nu}_k^{(m)}$ is given by
$$E\left[\left|\nu_k^{(m)} - \widehat{\nu}_k^{(m)}\right|^2\right] = \frac{1}{N}\left(4\nu_k^{(m)} + (\nu_k^{(m)})^2\right)$$
while the error variance of the Maximum-Likelihood (ML) estimator with known $x_{k,n}$ is given by
$$E\left[\left|\nu_k^{(m)} - \frac{1}{N}\sum_{n=1}^{N}|z_{k,n}^{(m)} - x_{k,n}|^2\right|^2\right] = \frac{1}{N}(\nu_k^{(m)})^2$$
Hence, if $4\nu_k/N \ll 1$ the proposed estimator performs very close to the ML estimator for known code symbols.

3. If the complex amplitude is estimated reliably, i.e., $\widehat{w}_k^{(m)} \approx w_k$, and if $x_{k,n}$ is uncorrelated with $\widehat{x}_{j,n}$ for $j \neq k$, then $\zeta_{k,n}^{(m)}$ and $x_{k,n}$ are practically uncorrelated. Moreover, under mild conditions on the user amplitudes, for large K the residual interference term $\zeta_{k,n}^{(m)}$ is asymptotically Gaussian [43, 25]. We conclude that for large N and

[2] It is easily shown that $\widehat{\nu}_{k,n}^{(m)}$ is the Maximum-Likelihood estimator for the variance of the process $\zeta_{k,n}^{(m)}$ from the observation $z_{k,n}^{(m)} = x_{k,n} + \zeta_{k,n}^{(m)}$ when $x_{k,n}$ and $\zeta_{k,n}^{(m)}$ are white, statistically independent, and Gaussian with $x_{k,n} \sim \mathcal{N}_{\mathbb{C}}(0,1)$ and $\zeta_{k,n}^{(m)} \sim \mathcal{N}_{\mathbb{C}}(0,\nu_k^{(m)})$.

K the estimator $\widehat{\nu}_k^{(m)}$ performs very close to the ML estimator for known coded symbols.

In the actual receiver implementation, the EXT and APP pmfs (3) and (4) are calculated by using (5) where $\nu_k^{(m)}$ is replaced by its estimate $\widehat{\nu}_k^{(m)}$ given by (6).

3.2 Soft estimation of the code symbols

The (non-linear) MMSE estimate of symbol $x_{k,n}$ given the observation $\mathbf{Y}$ is given by the conditional mean [44]

$$x_{k,n}^{\rm mmse} = E[x_{k,n}|\mathbf{Y}] = +\Pr(x_{k,n} = +1|\mathbf{Y}) - \Pr(x_{k,n} = -1|\mathbf{Y}) = 2\Pr(x_{k,n} = +1|\mathbf{Y}) - 1 \quad (7)$$

where $\Pr(x_{k,n} = a|\mathbf{Y})$ is the a posteriori pmf of symbol $x_{k,n}$ given the observation $\mathbf{Y}$. We are tempted to replace $\Pr(x_{k,n} = a|\mathbf{Y})$ by $\text{APP}_{k,n}^{(m)}(a)$ given by the SISO output at iteration m and let $\widehat{x}_{k,n}^{(m)} = 2\text{APP}_{k,n}^{(m)}(+1) - 1$, and claim that this choice minimizes the residual interference variance and it is therefore optimal. Unfortunately, this reasoning is incorrect. An intuitive way of seeing this is by contradiction: if the true a posteriori pmfs $\Pr(x_{k,n} = a|\mathbf{Y})$ were available at some iteration, then *optimal* symbol-by-symbol MAP decisions could be made and there would be no need for further interference cancellation. Moreover, the *exact* calculation of a posteriori probabilities $\Pr(x_{k,n} = a|\mathbf{Y})$ is in general an NP-complete problem [1]. Therefore, if after a finite number of iterations m an iterative algorithm (with polynomial complexity in K) obtains exact values for $\Pr(x_{k,n} = a|\mathbf{Y})$ the NP-completeness would be violated. Hence, we conclude that $\text{APP}_{k,n}^{(m)}(a) \not\equiv \Pr(x_{k,n} = a|\mathbf{Y})$, for any *finite* number of iterations m.

Interestingly, the above "non-linear MMSE argument" has been used in several papers (e.g., [22, 23, 8, 18]), sometimes with claim of optimality. On the contrary, by using a rigorous derivation based on factor-graphs and on the application of the sum-product algorithm, it can be shown that [25]:

1. Even for perfectly known amplitudes and SISO input variances (i.e., $\widehat{w}_k^{(m)} = w_k$ and $\widehat{\nu}_k^{(m)} = \nu_k^{(m)}$), the residual interference term $\zeta_{k,n}^{(m)} = z_{k,n}^{(m)} - x_{k,n}$ in (2) when using $\widehat{x}_{k,n}^{(m)} = 2\text{APP}_{k,n}^{(m)}(+1) - 1$ is conditionally biased and the bias tends to cancel the useful signal, i.e.,
$$E[\zeta_{k,n}^{(m)}|x_{k,n} = a] = -\mu_{k,n}^{(m)} a$$
where $\mu_{k,n}^{(m)}$ is a non-negative quantity that may depend on k, n and on the iteration index m.

2. By using EXT-based instead of APP-based symbol estimates, i.e., by using $\widehat{x}_{k,n}^{(m)} = 2\mathrm{EXT}_{k,n}^{(m)}(+1) - 1$, the resulting residual interference term is conditionally unbiased, i.e., $E[\zeta_{k,n}^{(m)}|x_{k,n}] = 0$, and the overall soft-SIC algorithm attains better performance than its APP-based version. Remarkably, this effect is not visible for small channel load but, as K/L increases, the difference between APP-based and EXT-based soft-SIC schemes is more and more evident [27].

In passing, we notice also that a biased residual interference implies that $x_{k,n}$ and $\zeta_{k,n}^{(m)}$ are correlated (even for perfect amplitude estimation). Hence, the variance estimator (6) is asymptotically optimal for large N, K only when the symbol soft estimates are obtained from EXT pmfs.

Driven by the results of [25] and by the above considerations, we shall use the following soft symbol estimates

$$\widehat{x}_{k,n}^{(m)} = 2\mathrm{EXT}_{k,n}^{(m)}(+1) - 1 \tag{8}$$

which can be regarded as a "local" MMSE estimate of $x_{k,n}$ assuming that the a posteriori pmf of $x_{k,n}$ is $\mathrm{EXT}_{k,n}^{(m)}(a)$ (even if it is not true!). [3]

3.3 Estimation of the user complex amplitudes

Let $\mathbf{w} = (w_1, \ldots, w_K)^T$ denote the vector of complex amplitudes to be estimated. The ML estimate of $\mathbf{w}$ given the observation $\mathbf{Y}$ is given by

$$\mathbf{w}^{\mathrm{ML}} = \arg \max_{\mathbf{w}} \log p(\mathbf{Y}|\mathbf{w}) \tag{9}$$

where $p(\mathbf{Y}|\mathbf{w})$ is the conditional pdf of the observed signal given $\mathbf{w}$, given by

$$\begin{aligned} p(\mathbf{Y}|\mathbf{w}) &\propto \sum_{\mathbf{X}} p(\mathbf{Y}|\mathbf{X}, \mathbf{w}) \Pr(\mathbf{X}|\mathbf{w}) \\ &\propto \sum_{\mathbf{x}^1 \in \mathcal{C}_1} \cdots \sum_{\mathbf{x}^K \in \mathcal{C}_K} \exp\left(-\frac{1}{N_0} \sum_{n=1}^{N} |\mathbf{y}_n - \mathbf{S}\mathcal{X}_n \mathbf{w}|^2 \right) \end{aligned} \tag{10}$$

where we have defined the diagonal matrix $\mathcal{X}_n = \mathrm{diag}(x_{1,n}, \ldots, x_{K,n})$ and where we have used the fact that the channel input $\mathbf{X}$ is independent of the channel amplitudes, so that

[3] In [25], expression (8) is derived as a direct consequence of the application of the sum-product algorithm, without any heuristic motivation based on MMSE estimation. The fact that EXT-based algorithms perform better than APP-based algorithms just puts in evidence the power and generality of the sum-product approach to statistical inference problems on Bayesian networks (see [45] and references therein).

$\Pr(\mathbf{X}|\mathbf{w}) = \Pr(\mathbf{X}) =$ uniform on the Cartesian product of the code books $\mathcal{C}_1 \times \cdots \times \mathcal{C}_K$ and zero outside, since each user k selects its code word with uniform probability on its code book $\mathcal{C}_k$ and independently of the other users. From (10) it is clear that direct ML estimation of $\mathbf{w}$ is infeasible in any practical case, as it has complexity proportional to the total number of user code words $\prod_{k=1}^{K} |\mathcal{C}_k|$.

Now, assume that the estimate $\widehat{\mathbf{w}}^{(m)}$ and the a posteriori probability $\Pr(\mathbf{X}|\mathbf{Y}, \widehat{\mathbf{w}}^{(m)})$ are available at iteration m. Then, we can produce an updated estimate $\widehat{\mathbf{w}}^{(m+1)}$ for next iteration by following the EM approach. In the language of the EM algorithm [31], $\mathbf{Y}, \mathbf{X}$ and $\{\mathbf{Y}, \mathbf{X}\}$ play the role of *incomplete*, *missing* and *complete* data. The EM update consists of computing the expected log-likelihood function of the complete data conditionally on the incomplete data and on the current parameter estimate (E-step), and maximizing the result with respect to the parameter (M-step) [31]. In our case, the complete data log-likelihood function is given by

$$\begin{aligned} \log p(\mathbf{Y}, \mathbf{X}|\mathbf{w}) &\doteq \log p(\mathbf{Y}|\mathbf{X}, \mathbf{w}) \\ &\doteq -\frac{1}{N_0} \sum_{n=1}^{N} |\mathbf{y}_n - \mathbf{S}\mathcal{X}_n \mathbf{w}|^2 \\ &\doteq \frac{2}{N_0} \mathrm{Re}\left\{\mathbf{r}^H \mathbf{w}\right\} - \frac{1}{N_0} \mathbf{w}^H \mathbf{R} \mathbf{w} \end{aligned} \tag{11}$$

where we define the vector

$$\mathbf{r} = \sum_{n=1}^{N} \mathcal{X}_n \mathbf{S}^H \mathbf{y}_n = \sum_{n=1}^{N} \begin{bmatrix} x_{1,n} \mathbf{s}_1^H \mathbf{y}_n \\ x_{2,n} \mathbf{s}_2^H \mathbf{y}_n \\ \vdots \\ x_{K,n} \mathbf{s}_K^H \mathbf{y}_n \end{bmatrix} \tag{12}$$

and the $K \times K$ matrix

$$\mathbf{R} = \sum_{n=1}^{N} \mathcal{X}_n \mathbf{S}^H \mathbf{S} \mathcal{X}_n \tag{13}$$

with (i,j)-th element

$$[\mathbf{R}]_{i,j} = \begin{cases} N & \text{for } \; i = j \\ \mathbf{s}_i^H \mathbf{s}_j \sum_{n=1}^{N} x_{i,n} x_{j,n} & \text{for } \; i \neq j \end{cases}$$

By using (11) we obtain the E-step in the form

$$\begin{aligned} Q(\mathbf{w}, \widehat{\mathbf{w}}^{(m)}) &= E[\log p(\mathbf{Y}, \mathbf{X}|\mathbf{w}) | \mathbf{Y}, \widehat{\mathbf{w}}^{(m)}] \\ &= \sum_{\mathbf{X}} \Pr(\mathbf{X}|\mathbf{Y}, \widehat{\mathbf{w}}^{(m)}) \log p(\mathbf{Y}, \mathbf{X}|\mathbf{w}) \\ &\doteq \frac{2}{N_0} \mathrm{Re}\left\{\overline{\mathbf{r}}^H \mathbf{w}\right\} - \frac{1}{N_0} \mathbf{w}^H \overline{\mathbf{R}} \mathbf{w} \end{aligned} \tag{14}$$

where we let $\bar{\mathbf{r}} = E[\mathbf{r}|\mathbf{Y}, \widehat{\mathbf{w}}^{(m)}]$ and $\overline{\mathbf{R}} = E[\mathbf{R}|\mathbf{Y}, \widehat{\mathbf{w}}^{(m)}]$. These are given by

$$\bar{\mathbf{r}} = \sum_{n=1}^{N} \overline{\mathfrak{X}}_n \mathbf{S}^H \mathbf{y}_n \tag{15}$$

and by

$$[\overline{\mathbf{R}}]_{i,j} = \begin{cases} N & \text{for } i = j \\ \mathbf{s}_i^H \mathbf{s}_j \sum_{n=1}^{N} \overline{x_{i,n} x_{j,n}} & \text{for } i \neq j \end{cases} \tag{16}$$

where $\overline{\mathfrak{X}}_n = \text{diag}(\overline{x_{1,n}}, \ldots, \overline{x_{K,n}})$ and where $\overline{x_{k,n}}$ and $\overline{x_{k,n} x_{j,\ell}}$ denote the first and second moments of the joint a posteriori pmf $\Pr(\mathbf{X}|\mathbf{Y}, \widehat{\mathbf{w}}^{(m)})$, given by

$$\begin{aligned} \overline{x_{k,n}} &= \sum_{\mathbf{X}} x_{k,n} \Pr(\mathbf{X}|\mathbf{Y}, \widehat{\mathbf{w}}^{(m)}) \\ \overline{x_{k,n} x_{j,\ell}} &= \sum_{\mathbf{X}} x_{k,n} x_{j,\ell} \Pr(\mathbf{X}|\mathbf{Y}, \widehat{\mathbf{w}}^{(m)}) \end{aligned} \tag{17}$$

By noticing that (14) is a quadratic form in $\mathbf{w}$ and that $\overline{\mathbf{R}}$ is non-negative definite, the M-step is readily obtained as

$$\widehat{\mathbf{w}}^{(m+1)} = \arg \max_{\mathbf{w}} Q(\mathbf{w}, \widehat{\mathbf{w}}^{(m)}) = \overline{\mathbf{R}}^{-1} \bar{\mathbf{r}} \tag{18}$$

The above procedure has still complexity exponential in K, since the computation of the moments (17) is equivalent to the marginalization of the joint pmf $\Pr(\mathbf{X}|\mathbf{Y}, \widehat{\mathbf{w}}^{(m)})$, which has complexity exponential in K. Then, we shall apply the above EM step "locally", i.e., by replacing $\Pr(\mathbf{X}|\mathbf{Y}, \widehat{\mathbf{w}}^{(m)})$ by the product of the marginal APPs produced by the SISO decoders at the end of iteration m. Namely, we use the approximation

$$\Pr(\mathbf{X}|\mathbf{Y}, \widehat{\mathbf{w}}^{(m)}) \approx \prod_{k=1}^{K} \prod_{n=1}^{N} \text{APP}_{k,n}^{(m)}(x_{k,n}) \tag{19}$$

As shown in the previous section, the APPs do not coincide in general with the true marginals of the joint pmf $\Pr(\mathbf{X}|\mathbf{Y}, \widehat{\mathbf{w}}^{(m)})$. However, the utility of the *approximation* (19) is twofold: on one hand, the product pmf in the LHS is readily available from the SISO outputs. On the other hand, thanks to the product form, the exponential complexity of the moment computation is reduced to linear. In fact, the moments of the product pmf are given by

$$\begin{aligned} \widetilde{x}_{k,n} &= +\text{APP}_{k,n}^{(m)}(+1) - \text{APP}_{k,n}^{(m)}(-1) = 2\text{APP}_{k,n}^{(m)}(+1) - 1 \\ \widetilde{x_{k,n} x_{j,\ell}} &= \begin{cases} 1 & \text{for } (k,n) = (j,\ell) \\ \widetilde{x}_{k,n} \widetilde{x}_{j,\ell} & \text{otherwise} \end{cases} \end{aligned} \tag{20}$$

Finally, the proposed approximated EM updating step consists of computing (18) where $\overline{\mathbf{R}}$ and $\overline{\mathbf{r}}$ are given by (15) and by (16) when replacing the true moments (17) by their approximations (20).

The complexity of (18) is then dominated by the matrix inverse $\overline{\mathbf{R}}^{-1}$, which must be computed at each iteration. A suboptimal M-step that does not require a matrix inverse can be obtained by noticing that, under mild conditions on random interleaving and on the uniformity of user codes, the averaged symbols $\widetilde{x}_{k,n}$ are symmetrically distributed (their distribution is induced by the noise and by the random choice of the user code words over the code books). Moreover, $\widetilde{x}_{k,n}$ and $\widetilde{x}_{j,n}$ are weakly correlated for $k \neq j$. Then, $\frac{1}{N}\overline{\mathbf{R}} \approx \mathbf{I}$ for large block length N. Hence, under these conditions (18) can be approximated by

$$\widehat{\mathbf{w}}^{(m+1)} = \frac{1}{N}\overline{\mathbf{r}} \tag{21}$$

Notice that both (18) and (21) are directly computed from the SUMF outputs, since $\overline{\mathbf{r}}$ defined in (15) depends on the observed signal $\mathbf{Y}$ only through the SUMF outputs $\mathbf{s}_k^H \mathbf{y}_n$.

3.4 Initialization and combining with the training phase

The overall iterative soft-SIC algorithm needs a sufficiently reliable initial estimate $\widehat{\mathbf{w}}^{(0)}$ of the complex user amplitudes. Otherwise, for completely unknown $\mathbf{w}$, the SISO decoders at the first iteration yield APPs very close to 1/2, i.e., $\widetilde{x}_{k,n} \approx 0$ for all k and n. This yields $\overline{\mathbf{r}} \approx 0$ and $\overline{\mathbf{R}} = N\mathbf{I}$, which in turns yields $\widehat{\mathbf{w}}^{(1)} \approx \mathbf{0}$, so that the receiver never "bootstraps" and remains stuck at the "zero" fixed point.

For the sake of initialization, a joint ML estimate of the complex amplitudes is obtained from the training phase. This is readily given by [44]

$$\widehat{\mathbf{w}}^{(t)} = \left(\mathbf{R}^{(t)}\right)^{-1} \mathbf{r}^{(t)} \tag{22}$$

where $\mathbf{r}^{(t)}$ and $\mathbf{R}^{(t)}$ are given by (12) and by (13), respectively, when replacing N by T and the code symbols $x_{k,n}$ by the known training symbols $x_{k,n}^{(t)}$. If the training sequences are mutually orthogonal, i.e., such that $(\mathbf{X}^{(t)})^H \mathbf{X}^{(t)} = T\mathbf{I}$, we obtain $\mathbf{R}^{(t)} = T\mathbf{I}$ and no matrix inverse is needed in (22). It can be shown that this choice also minimizes the estimation error variance [46]. Then, if a set of mutually orthogonal training sequences exists, this choice should be preferred. [4]

The receiver is initialized by letting $\widehat{\mathbf{w}}^{(0)} = \widehat{\mathbf{w}}^{(t)}$. Then, at iterations $m = 1, 2, \ldots$, the receiver exploits the updated estimate $\widehat{\mathbf{w}}^{(m)}$ provided by the EM step (18) by combining it

[4] The existence of such set of training sequences depends on the training symbol alphabet and on the training length T, which must be $\geq K$. See [46] and references therein for more details.

in some way with the training-based estimate. [5] We investigate the following two methods for combining the training phase with the EM update.

The mixing method. For $m = 1, 2, \ldots$, the "local" EM estimation described above is applied to the incomplete data $\{\mathbf{Y}, \mathbf{Y}^{(t)}\}$ with missing data $\mathbf{X}$, by treating $\mathbf{X}^{(t)}$ as known parameters. The same result is obtained by including the known training symbols in the missing data and by defining their marginal pmfs as $\text{APP}_{k,n}^{(t)}(a) = 1$ if $a = x_{k,n}^{(t)}$ and 0 if $a \neq x_{k,n}^{(t)}$, so that for training symbols we have $\widetilde{x}_{k,n} = x_{k,n}^{(t)}$ [29]. After straightforward algebra, completely analogous to the derivation of the previous section and not reported here for the sake of space limitation, we obtain the *mixing* estimator as

$$\widehat{\widehat{\mathbf{w}}}_{\text{mix}}^{(m)} = \left[\ \overline{\mathbf{R}} + \mathbf{R}^{(t)} \ \right]^{-1} \left(\overline{\mathbf{r}} + \mathbf{r}^{(t)} \right) \tag{23}$$

The combining method. Assume for simplicity that the training sequences are mutually orthogonal. Then, $\widehat{\mathbf{w}}^{(t)} = \mathbf{w} + \boldsymbol{\eta}^{(t)}$ with $\boldsymbol{\eta}^{(t)} \sim \mathcal{N}_{\mathbb{C}}(\mathbf{0}, \frac{N_0}{T}\mathbf{I})$. In particular, the training-based estimator $\widehat{\mathbf{w}}^{(t)}$ is unbiased.

Now, from (15) and (1) we obtain

$$\overline{\mathbf{r}} = \sum_{n=1}^{N} \widetilde{\mathcal{X}}_n \mathbf{S}^H (\mathbf{S}\mathcal{X}_n \mathbf{w} + \mathbf{n}_n) = \mathbf{R}'\mathbf{w} + \boldsymbol{\eta}$$

where $\mathbf{R}' = \sum_{n=1}^{N} \widetilde{\mathcal{X}}_n \mathbf{S}^H \mathbf{S} \mathcal{X}_n$ and $\boldsymbol{\eta} \sim \mathcal{N}_{\mathbb{C}}(\mathbf{0}, N_0 \mathbf{R}'')$ with $\mathbf{R}'' = \sum_{n=1}^{N} \widetilde{\mathcal{X}}_n \mathbf{S}\mathbf{S}^H \widetilde{\mathcal{X}}_n$. By using this into (18) we have

$$\widehat{\mathbf{w}}^{(m)} = \overline{\mathbf{R}}^{-1} \mathbf{R}'\mathbf{w} + \mathbf{R}^{-1}\boldsymbol{\eta} \tag{24}$$

Since $\overline{\mathbf{R}} \neq \mathbf{R}'$ unless the code symbols are perfectly known, the result of EM is biased. For the sake of simplicity, we assume that N is sufficiently large so that the following approximations hold

$$\begin{aligned} \overline{\mathbf{R}} &\approx N\mathbf{I} \\ \mathbf{R}' &\approx \text{diag}\left(\sum_{n=1}^{N} x_{1,n}\widetilde{x}_{1,n}, \ldots, \sum_{n=1}^{N} x_{K,n}\widetilde{x}_{K,n} \right) \\ \mathbf{R}'' &\approx \text{diag}\left(\sum_{n=1}^{N} |\widetilde{x}_{1,n}|^2, \ldots, \sum_{n=1}^{N} |\widetilde{x}_{K,n}|^2 \right) \end{aligned} \tag{25}$$

[5] The efficient use of the available training symbols in addition to some blind parameter estimation technique is a problem common to many *semi-blind* schemes (see [47] and references therein).

(this follows by the fact that, under mild conditions, the out-of-diagonal terms are normalized empirical correlations between uncorrelated zero-mean sequences, which vanish for large N). By using (25) in (24) we obtain the biased EM estimate of user k amplitude as

$$\widehat{w}_k^{(m)} = \alpha_k w_k + \eta'_k \tag{26}$$

where

$$\alpha_k = \frac{1}{N}\sum_{n=1}^{N} x_{k,n}\widetilde{x}_{k,n}$$

and where $\eta'_k \sim \mathcal{N}_{\mathbb{C}}(0, \frac{N_0}{N}\beta_k^2)$ is the k-th component of $\overline{\mathbf{R}}^{-1}\boldsymbol{\eta}$, with

$$\beta_k^2 = \frac{1}{N}\sum_{n=1}^{N} |\widetilde{x}_{k,n}|^2$$

Now, our goal is to obtain a combined estimator in the form

$$\widehat{\widehat{w}}_{k,\text{comb}}^{(m)} = a_k \widehat{w}_k^{(m)} + b_k \widehat{w}_k^{(t)} \tag{27}$$

where the coefficients a_k and b_k are chosen in order to minimize the error variance subject to the unbiased constraint, i.e., they are the solution of

$$\begin{cases} \text{minimize} & E[|a_k\eta'_k + b_k\eta_k^{(t)}|^2] \\ \text{subject to} & a_k\alpha_k + b_k = 1 \end{cases}$$

Since η'_k and $\eta_k^{(t)}$ are mutually independent (they depend on the mutually independent noise samples $\mathbf{N}$ and $\mathbf{N}^{(t)}$ in the data and training phases), we obtain easily the solution of the above problem as

$$\begin{aligned} a_k &= \frac{\alpha_k}{\alpha_k^2 + \frac{T}{N}\beta_k^2} \\ b_k &= \frac{\frac{T}{N}\beta_k^2}{\alpha_k^2 + \frac{T}{N}\beta_k^2} \end{aligned} \tag{28}$$

One last problem is represented by the fact that α_k depends on the unknown code symbols $x_{k,n}$. Then, an estimate of α_k can be obtained as follows. We notice that

$$x_{k,n}\widetilde{x}_{k,n} = \begin{cases} |\widetilde{x}_{k,n}| & \text{for } \operatorname{sign}(\widetilde{x}_{k,n}) = x_{k,n} \\ -|\widetilde{x}_{k,n}| & \text{for } \operatorname{sign}(\widetilde{x}_{k,n}) \neq x_{k,n} \end{cases}$$

Since $\operatorname{sign}(\widetilde{x}_{k,n})$ is the maximum a posteriori symbol-by-symbol decision on the code symbol $x_{k,n}$ based on the a posteriori pmf $\text{APP}_{k,n}^{(m)}(a)$ output by the SISO decoder at iteration

m, for large N the following approximation holds

$$\alpha_k \approx (1 - 2\epsilon_k) \frac{1}{N} \sum_{n=1}^{N} |\widetilde{x}_{k,n}| \tag{29}$$

where ϵ_k is the symbol error probability (on the coded symbols, not on the information bits!) at the output of the SISO decoder for user k at iteration m. If the residual interference plus noise process $\zeta_{k,n}^{(m)}$ is Gaussian with variance $\nu_k^{(m)}$, the error probability ϵ_k is a known function of $\nu_k^{(m)}$, determined by the user code $\mathcal{C}_k$. This can be pre-computed and stored in a look-up table, and an estimate $\widehat{\epsilon}_k$ of ϵ_k can be easily obtained from the estimate $\widehat{\nu}_k^{(m)}$ given by (6). Finally, α_k can be approximated by replacing ϵ_k by $\widehat{\epsilon}_k$ in (29).

Remark. We provide a qualitative and intuitive discussion on the behavior of the mixing and combining methods.

The mixing method suffers from bias in the case of large K/L and $T/N \ll 1$ (which is clearly the most interesting case, as it is usually desirable to maximize the channel load and minimize the length of the training phase). In fact, suppose that at iteration $m = 0$ the signal at the input of each SISO decoder is "very noisy", since the interference has not been removed yet and K/L is large. Then, the averaged symbols $\widetilde{x}_{k,n}$ output by the SISO decoders are all close to zero. Assuming orthogonal training sequences (the best case), the mixing method yields $\overline{\mathbf{R}} + \mathbf{R}^{(t)} \approx (N+T)\mathbf{I}$, $\overline{\mathbf{r}} \approx \mathbf{0}$ and $\mathbf{r}^{(t)} = T\mathbf{w} + \text{noise}$. The resulting estimator is

$$\widehat{\widehat{\mathbf{w}}}_{\text{mix}} \approx \frac{T}{N+T}\mathbf{w} + \text{noise}$$

which is clearly biased. In particular, if $T/N \ll 1$, the bias might prevent the whole receiver to bootstrap. [6]

On the contrary, the combining method (assuming α_k known) provides an unbiased estimate at each iteration. At the first iterations, when $\epsilon_k \approx 1/2$, then $a_k \approx 0$, $b_k \approx 1$ and $\widehat{\widehat{\mathbf{w}}}_{\text{comb}}^{(m)} \approx \widehat{\mathbf{w}}^{(t)}$, i.e., only the result of training-based estimation is used. As the soft-SIC cleans-up the signal from interference and ϵ_k becomes small (converging to the single-user performance), then $|\widetilde{x}_{k,n}| \approx 1$, $\alpha_k \approx \beta_k^2 \approx 1$ and $a_k \approx \frac{N}{N+T}$, $b_k \approx \frac{T}{N+T}$. These limiting values are precisely the maximal-ratio combining coefficient [41] for estimating $\mathbf{w}$ from the

[6] Interestingly, in [29] training symbols are used in an iterative joint decoder and channel estimation scheme according to the mixing method. The analysis in [29] is uniquely based on propagating the variances of residual interference and of channel estimation errors from one iteration to the next, and does not take into account the bias. Unfortunately, the interference cancellation algorithm of [29] is based on APPs and hence it is plagued by biased residual interference [25], and the mixing method yields biased channel estimates (as outlined here). Therefore, the results of [29] are questionable.

unbiased noisy observations $\mathbf{w} + \boldsymbol{\eta}^{(t)}$ and $\mathbf{w} + \boldsymbol{\eta}$, with $\boldsymbol{\eta}^{(t)}$ and $\boldsymbol{\eta}$ independent, Gaussian, with covariances $\frac{N_0}{T}\mathbf{I}$ and $\frac{N_0}{N}\mathbf{I}$, respectively. Comparisons between the mixing and the combining methods are provided in Section 4.

3.5 Algorithm summary

Fig. 2 shows the block diagram of the proposed receiver. The users are ranked in decreasing order of their estimated signal-to-interference ratio, given by

$$\frac{|\widehat{w}_k^{(t)}|^2}{\sum_{j \neq k} |\mathbf{s}_k^H \mathbf{s}_j|^2 |\widehat{w}_j^{(t)}|^2}$$

Without loss of generality, we assume that the decoding order is $k = 1, 2, \ldots, K$. The algorithm is initialized by letting $\widehat{\mathbf{w}}^{(0)} = \widehat{\mathbf{w}}^{(t)}$, $\widehat{x}_{k,n}^{(-1)} = 0$ for all k and n and $m = 0$. Then we have:

- User loop: For $k = 1, \ldots, K$, do
- Symbol loop: For $n = 1, \ldots, N$, do
- Compute the soft-SIC signal samples according to (2) and the estimated residual interference plus noise variance $\widehat{\nu}_k^{(m)}$ according to (6).
- Compute the k-th SISO decoder EXT and APP outputs and compute the soft interference estimate $\widehat{x}_{k,n}^{(m)}$ according to (8) and the average symbols $\widetilde{x}_{k,n}$ according to (20).
- End symbol loop.
- End user loop.
- Parameter estimation update: If the mixing method is used, compute the updated amplitude estimate according to (23). If the combining method is used, compute the EM amplitude estimate according to (18) and the updated estimate according to (27).
- If $m = M$, make symbol-by-symbol decisions on the information bits APP outputs of the SISO decoders, otherwise let $m := m + 1$ and go back to the user loop.

4 Results

In order to demonstrate the performance of the proposed soft-SIC receiver, we considered the following simulation setting, loosely inspired by the UMTS-TDD system [3]:

- Spreading factor $L = 16$, QPSK chips with "short" random spreading sequences. A new set of K sequences is generated randomly and independently with i.i.d. elements at each frame. Obviously, the BER is averaged over several frames so that the effect of the random sequences is smoothed.

- The user code is the same for all users. For the sake of simplicity, we chose the 4-state rate-1/2 convolutional code (CC) with generators $(5, 7)_8$ (octal notation [41]).

- Code block length $N = 2000$ coded symbols, corresponding to 1000 information bits per frame.

- $K = 32$ and 40 users, corresponding to channel loads of 2.0 and 2.5 users per chip, respectively.

- Training sequence lengths $T = 4$ and 32 symbols.

- Users have the same received power. The channel complex amplitudes are given by $w_k = \sqrt{RE_b}e^{j\phi_k}$ where R is the user coding rate ($R = 1/2$ in our case), E_b is the energy per information bit and ϕ_k is a uniformly distributed random variable over $[-\pi, \pi]$, independently generated for each user.

- We considered a fixed maximum number of SIC iterations $M = 10$, in all cases.

In these examples we considered only the equal-rate equal-power users for the sake of space limitation and since this is a worst-case for iterative soft-SIC decoders (see the discussion in [14, 15]). In [25], by using the technique of *density evolution*, which is now a standard tool for the analysis of iterative "message passing" algorithms (see [45] and references therein), it is shown that the soft interference cancellation algorithm considered here at target BER $= 10^{-5}$, with perfect channel parameter knowledge, CC $(5, 7)_8$ user codes and equal power users attains channel load of 3 users/chip. The required E_b/N_0 is 6 dB. Fig. 3 shows the BER curves for $K = 40$ users and perfect channel knowledge (all BER curves show the worst user performance, which in the equal power case is usually, but not necessarily, obtained by the user decoded first). The load in this case is $40/16 = 2.5$, below the limit of 3 predicted by the analysis of [25]. For $E_b/N_0 \geq 5$ dB and 10 iterations the single-user BER performance is achieved for all users. Obviously, for smaller K the convergence to the single-user BER occurs with less iterations and at lower E_b/N_0 threshold.

Fig. 4, 5 and 6 show the BER of the system with $K = 32$ users and $T = 32$ training symbols per frame, with training estimation only, and EM+training estimation with mixing and combining methods, respectively. Training-only estimation prevents the receiver to achieve the single-user BER, since interference cannot be canceled completely because of the estimation errors which do not vanish with iterations. The combining method shows faster convergence than the mixing method. This confirms the qualitative bias analysis made in the remark of Section 3.3. However, for such "light" load [7] the difference between the two methods is not very significant.

Fig. 7, 8 and 9 show the BER of the system with $K = 32$ users and $T = 4$ training symbols per frame, with training-estimation only, and EM+training estimation with mixing and combining methods, respectively. With only 4 training symbols, the degradation of system with training-only estimation is very evident (notice that for $T = 4$ and $K = 32$ it is obviously not possible to make the training sequences mutually orthogonal, and this contributes to poor channel estimation). Also, the better convergence properties of the combining method versus the mixing method are more evident: the combining method attains the single-user BER at $E_b/N_0 = 4$ dB, while the mixing method attains it at $E_b/N_0 = 6$ dB.

In order to put in evidence that the bias in the mixing method might prevent the receiver to converge to the single-user BER while the combining method still works, we consider the case $K = 40$ and $T = 32$ (again, orthogonal training sequences are not possible here). Fig. 10 and 11 show the BER of this system. The mixing method does not converge for the range of E_b/N_0 considered in our simulations, since the estimated amplitudes after the first iteration are biased by a factor $\approx 32/(2032) = 0.0157$, which prevents cancellation, and the received does not bootstrap. On the contrary, the combining method is still able to converge for $E_b/N_0 \geq 7$ dB. By comparing Fig. 3 with Fig. 11 we can quantify the degradation due to unknown channel amplitudes: with $M = 10$ iterations this is about 1.6 dB at BER $= 10^{-4}$, 0.8 dB at BER $= 10^{-5}$ and 0.0 dB at BER $\leq 4 \cdot 10^{-7}$, since in this BER range both systems achieve the single-user performance.

[7] It is worthwhile to point out here that $K = 32$ users with spreading factor $L = 16$ is a load already far beyond any conventional practical CDMA system [2, 3]. We call this load "light" since it is far from the threshold load predicted by the analysis of [25].

5 Conclusions

We proposed a low-complexity iterative soft-SIC algorithm for joint data detection and channel parameter estimation, based on SISO single-user decoders and soft interference cancellation. The channel parameters estimates are updated along with the receiver iterations. The updating operation has the form of a likelihood function expectation of followed by maximization, i.e., it is formally equivalent to the basic EM step.

Even though similar algorithms can be found (with minor variations) in several other works (see the discussion in Section 1), here we investigated in the details several new important aspects, namely: a simple and efficient way to estimate the residual interference plus noise variance at the SISO inputs; the issue of soft interference estimation based on EXT pmfs versus the conventional approach of using APPs; the correct formulation of EM estimation with channel coding, and the key approximation to bring complexity from exponential down to polynomial in the number of users; the use of training-based estimation together with EM updating. In particular, we provided a new method for combining the unbiased channel estimates provided by ML training-based estimation with the biased estimates provided by EM. The new method (referred to as "combining") provides much better convergence of the overall receiver than the more conventional method consisting of treating training symbols and unknown code symbols together (referred to as "mixing").

The full investigation of the optimal trade-off between training symbols fraction T/N and channel load K/L is out of the scope of this paper. However, from the simulation results shown here we can get some conclusions on the overall benefit of the proposed approach. With our receiver, we can fit $\alpha = 40/16 = 2.5$ users/chip with coding rate $R = 1/2$ bit/symbol at BER $= 10^{-5}$, with actual channel estimation ($T = 32$ training symbols out of $N = 2000$ coded symbol per frame) and non-recursive 4-state convolutional codes. The required E_b/N_0 is about 6.7 dB, i.e., user SNR ≈ 3.7 dB, with 10 iterations (10 SISO decoding per user per frame). In UMTS [3, 2], conventional SUMF receivers are envisaged, but very complex and powerful user channel codes are considered (either turbo-codes or 256-state convolutional codes). Consider for example a conventional system with turbo-codes of rate $R = 1/2$, optimized interleavers of size 1024 [26] (corresponding to coded block length $N = 2048$, similar to our case), recursive systematic 4-state CCs with generators $(1, 5/7)_8$ and 8 full iterations, corresponding to 16 SISO decoding per user per frame. [8]

[8] We allow more complexity in SISO decoding for the conventional system (16 SISO decoding steps instead of 10) since our system requires also interference cancellation, which involves some additional complexity

In the conventional system we assume perfect channel estimation, since channel estimation is much less critical than in the soft-SIC system. The turbo-code achieves BER $= 10^{-5}$ at SINR $= -1$ dB. The SINR at the output of the SUMF for equal-power users and random spreading sequences, in the limit for $K, L \to \infty$ with $K/L = \alpha$ [48], is given by SINR $= \frac{\text{SNR}}{1+\alpha\text{SNR}}$. Then, the limit load of the conventional turbo-encoded system is $\alpha = \frac{1}{\text{SINR}} - \frac{1}{\text{SNR}}$. By letting SINR $= -1$ dB (as required by the target BER performance) and SNR $= 3.7$ dB (as in the soft-SIC system), we obtain $\alpha = 0.83$. Even by letting SNR $\to \infty$, the maximum possible channel load is not larger than $\alpha = 1.26$. We conclude that the proposed receiver with *actual* channel estimation is able to (at least) double the cell capacity at roughly the same complexity of the conventional turbo-encoded system.

References

[1] S. Verdu, *Multiuser detection*, Cambridge University Press, Cambridge, UK, 1998.

[2] 3GPP, "TS 25.224 V3.1.0,"3GPP-TSG-RAN-WG1; Physical Layer Procedures (FDD)"," ETSI, December 1999.

[3] 3GPP, "TS 25.224 V3.1.0,"3GPP-TSG-RAN-WG1; Physical Layer Procedures (TDD)"," ETSI, December 1999.

[4] A. J. Viterbi, *CDMA – Principles of spread spectrum communications*, Addison-Wesley, Reading, MA, 1995.

[5] P. Patel and J. Holtzman, "Analysis of a simple successive interference cancellation scheme in a DS/CDMA system," *IEEE J. Select. Areas Commun.*, vol. 12, no. 5, pp. 796–807, June 1994.

[6] T. C. Yoon, R. Kohno, and H. Imai, "A spread-spectrum multiaccess system with cochannel interference cancellation for multipath fading channels," *IEEE J. Select. Areas Commun.*, vol. 11, no. 7, pp. 1067–1075, September 1993.

[7] M. Varanasi, "Decision feedback multiuser detection: a systematic approach," *IEEE Trans. on Inform. Theory*, vol. 45, no. 1, pp. 219–240, January 1999.

[8] A. Lampe and J. Huber, "On improved multiuser detection with soft decision interference cancellation," in *Proc. ICC 1999, Comm. Theory Mini-Conference*, Vancouver, June 1999, pp. 172–176.

[9] A. Hui and K. Ben Letaief, "Successive interference cancellation for multiuser asynchronous DS/CDMA detectors in multipath fading links," *IEEE Trans. on Commun.*, vol. 46, no. 3, pp. 384–391, March 1998.

[10] D. Divsalar, M. Simon, and D. Raphaeli, "Improved parallel interference cancellation for CDMA," *IEEE Trans. on Commun.*, vol. 46, no. 2, pp. 258–268, February 1998.

[11] S. Benedetto, D. Divsalar, G. Montorsi, and F. Pollara, "Soft-Input Soft-Output building blocks for the construction of distributed iterative decoding of code networks," *European Trans. on Commun.*, April 1998.

[12] P. Alexander, A. Grant, and M. Reed, "Iterative detection in code-division multiple-access with error control coding," *European Trans. on Telecomm.*, vol. 9, no. 5, pp. 419–425, September 1998.

[13] L. Brunel and J. Boutros, "Code division multiple access based on independent codes and turbo decoding," *Annales des Télécommunications*, vol. 54, no. 7-8, pp. 401–410, July 1999.

[14] N. Chayat and S. Shamai, "Iterative soft onion peeling for multi-acess and broadcast channels," in *Proc. PIMRC'98*, Boston, September 1998.

[15] N. Chayat and S. Shamai, "Convergence properties of iterative soft onion peeling," in *Proc. ITW 1999*, Kruger national park, South Africa, June 1999, p. 9.

[16] M. Damen, *Joint coding/decoding in a multiple access system: applications to mobile communications*, Ph.D. thesis, Ph.D Thesis, ENST Paris, 1999.

[17] N. Ibrahim, *Codage et decodage de canal pour un système de communication à accès multiple*, Ph.D. thesis, Ph.D Thesis, ENST Paris, 1999.

[18] A. Lampe, "Analytic solution to the performance of iterated soft decision interference cancellation for coded CDMA transmission over frequency selective channel," in *IEEE 6-th Int. Symp. on Spread-Spectrum Tech. and Appl., ISSSTA 2000*, NJIT, N.J., USA, September 2000.

[19] G. Woodward M. Honig and P. Alexander, "Adaptive multiuser parallel decision-feedback with iterative decoding," in *Proc. ISIT 2000*, Sorrento, Italy, June 2000, p. 335.

[20] M. Reed, C. Schlegel, P. Alexander, and J. Asenstorfer, "Iterative multiuser detection for CDMA with FEC: near single-user performance," *IEEE Trans. on Commun.*, vol. 46, no. 12, pp. 1693–1699, December 1998.

[21] C. Schlegel, "Joint detection in multiuser systems via iterative processing," in *Proc. ISIT 2000*, Sorrento, Italy, June 2000, p. 274.

[22] F. Tarkoy, "Iterative multiuser decoding for asynchronous users," in *Proc. ISIT '97*, Ulm, Germany, July 1997, p. 30.

[23] X. Wang and V. Poor, "Iterative (Turbo) soft interference cancellation and decoding for coded CDMA," *IEEE Trans. on Commun.*, vol. 47, no. 7, pp. 1047–1061, July 1999.

[24] F. Kschischang, B. Frey, and H.-A. Loeliger, "Factor graphs and the sum-product algorithm," *IEEE Trans. on Inform. Theory*, vol. 47, no. 2, pp. 498–519, February 2001.

[25] J. Boutros and G. Caire, "Iterative multiuser decoding: unified framework and asymptotic performance analysis," submitted to IEEE Trans. on Inform. Theory, also available at `www.eurecom.fr/ caire`, August 2000.

[26] C. Berrou and A. Glavieux, "Near optimum error-correcting coding and decoding: Turbo codes," *IEEE Trans. on Commun.*, vol. 44, no. 10, October 1996.

[27] S. Marinkovic, B. Vucetic, and J. Evans, "Improved iterative parallel interference cancellation," in *Intern. Symp. on Inform. Theory,* ISIT 2001, Washington DC, June 2001.

[28] H. El Gamal and E. Geraniotis, "Iterative multiuser detection for coded CDMA signals in AWGN and fading channels," *IEEE J. Select. Areas Commun.*, vol. 18, no. 1, pp. 30–41, January 2000.

[29] P. Alexander and A. Grant, "Iterative channel and information sequence estimation in CDMA," in *IEEE 6-th Int. Symp. on Spread-Spectrum Tech. and Appl., ISSSTA 2000*, NJIT, N.J., USA, September 2000.

[30] A. Dempster, N. Laird, and D. Rubin, "Maximum-likelihood from incomplete data via the EM algorithm," *J. Royal Statistics Soc., Ser. B*, vol. 39, no. 1, pp. 1–38, January 1977.

[31] T. Moon, "The expectation-maximization algorithm," *IEEE Signal Processing Magazine*, vol. 13, pp. 47–60, November 1996.

[32] M. Guernach and L. Vanderdorpe, "Performance analysis of joint EM/SAGE rstimation and multistage detection in UTRA-WCDMA uplink," in *Intern. Conf. on Commun. ICC '2000*, New Orleans, June 2000, pp. 638–640.

[33] U. Fawer and B. Aazhang, "A multiuser receiver for code division multiple access communications over multipath channels," *IEEE Trans. on Commun.*, vol. 43, no. 2/3/4, pp. 1556–1565, Feb./Mar./Apr. 1995.

[34] C. Cozzo and B. Hughes, "The Expectation-Maximization algorithm for space-time communications," in *Proc. ISIT 2000*, Sorrento, Italy, June 2000, p. 338.

[35] Y. Li, C. Georghiades, and G. Huang, "EM-based sequence estimation for space-time codes systems," in *Proc. ISIT 2000*, Sorrento, Italy, June 2000, p. 315.

[36] J. Boutros, F. Boixadera, and C. Lamy, "Bit-interleaved coded modulation for multiple-input multiple-output channels," in *IEEE 6-th Int. Symp. on Spread-Spectrum Tech. and Appl., ISSSTA 2000*, NJIT, N.J., USA, September 2000.

[37] A. Logothetis and C. Carlemalm, "SAGE algorithms for multipath detection and parameters estimation in asynchronous CDMA systems," *IEEE Trans. on Signal Processing*, vol. 48, no. 11.

[38] "Space-alternating generalized expectation-maximization algortihm," *IEEE Trans. on Sig. Proc.*, vol. 42, no. 9, pp. 2664–2677, October 1994.

[39] C. Georghiades and J. Choong Han, "Sequence estimation in the presence of random parameters via th EM algorithm," *IEEE Trans. on Commun.*, vol. 45, no. 3.

[40] L. Nelson and V. Poor, "Iterative multiuser receivers for CDMA channels: an EM-based approach," *IEEE Trans. on Commun.*, vol. 44, no. 12, pp. 1700–1710, December 1996.

[41] J. Proakis, *Digital communications, 3rd Ed.*, McGraw-Hill, New York, 1995.

[42] L. Bahl, J. Cocke, F. Jelinek, and J. Raviv, "Optimal decoding of linear codes for minimizing symbol error rate," *IEEE Trans. on Inform. Theory*, vol. 20, no. 3, pp. 284–287, March 1974.

[43] S. Verdu and S. Shamai, "Spectral efficiency of CDMA with random spreading," *IEEE Trans. on Inform. Theory*, vol. 45, no. 2, pp. 622–640, March 1999.

[44] V. Poor, *An introduction to signal detection and estimation*, Springer-Verlag, New York, 1988.

[45] "Special issue on iterative decoding," *IEEE Trans. on Inform. Theory*, vol. 47, no. 2, February 2001.

[46] G. Caire and U. Mitra, "Structured multiuser channel estimation for block-synchronous ds-cdma," *IEEE Trans. on Commun.*, vol. (to appear), 2001.

[47] E. de Carvalho and D. Slock, *"Semi-Blind Methods for FIR Multichannel Estimation", in: Signal Processing Advances in Communications, Volume 1: Trends in Channel Estimation and Equalization*, G. Giannakis and P. Stoica and Y. Hua and L. Tong, editors, Prentice Hall, 2000.

[48] D. Tse and S. Hanly, "Linear multiuser receivers: Effective interference, effective bandwidth and capacity," *IEEE Trans. on Inform. Theory*, vol. 45, no. 2, pp. 641–675, March 1999.

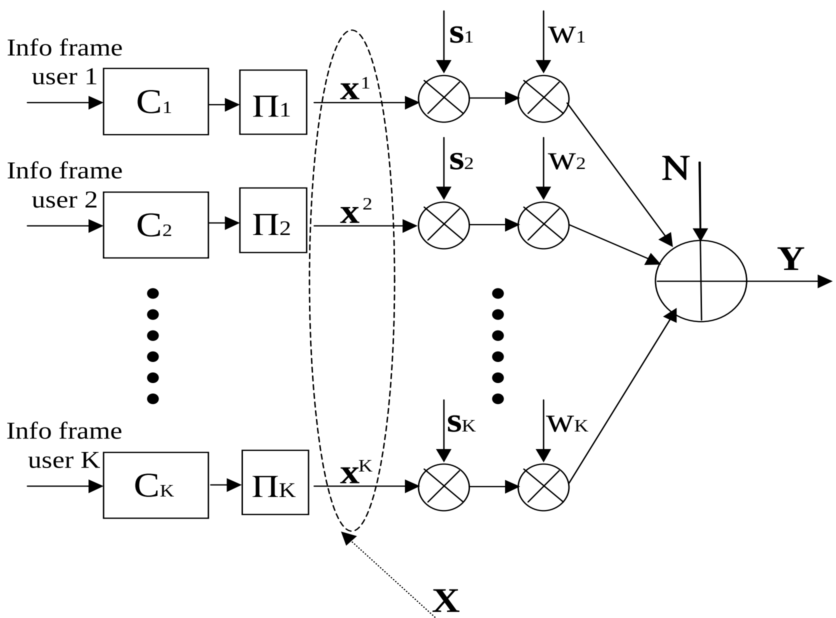

Figure 1: Coded synchronous DS/CDMA system (Π_k denotes interleaving, different for each user).

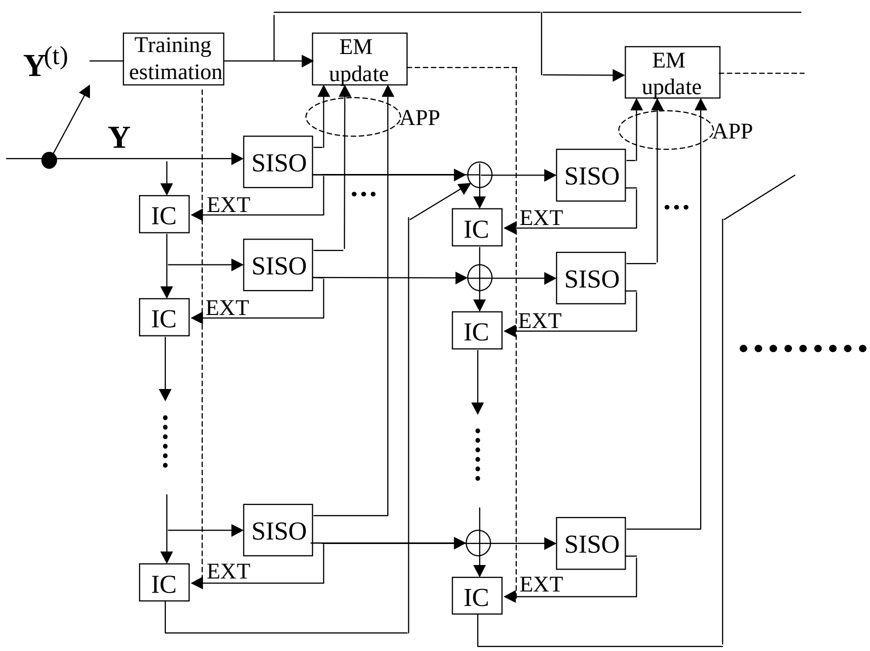

Figure 2: Block diagram of the proposed soft-SIC receiver with iterative EM channel estimation (only two iteration stages are shown for simplicity). APP and EXT denote soft code symbol estimates obtained from APP and EXT SISO outputs. The "IC" blocks denote interference cancellation and matched filtering.

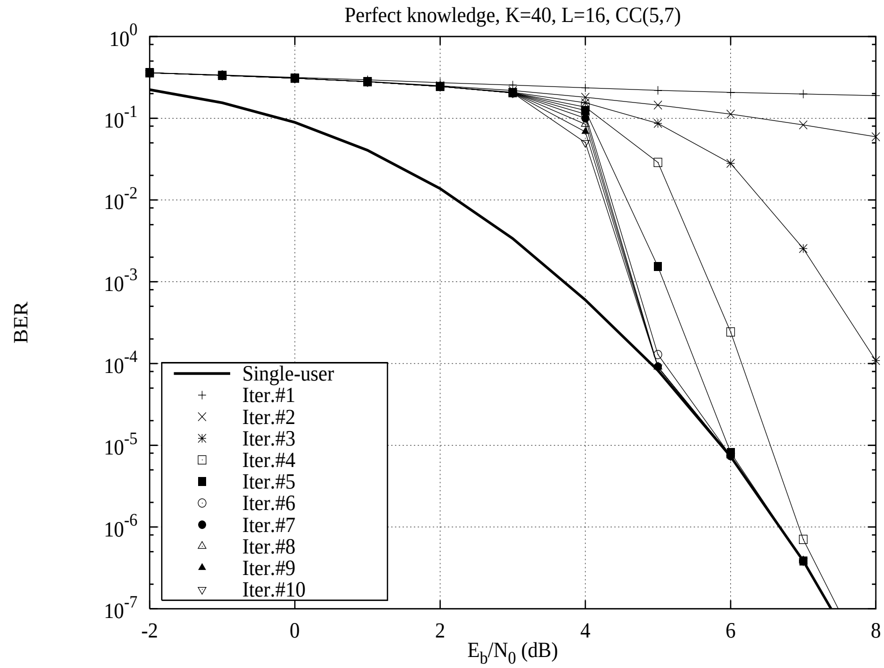

Figure 3: $K = 40, L = 16,\ N = 2000$, CC $(5,7)_8$, perfect channel knowledge.

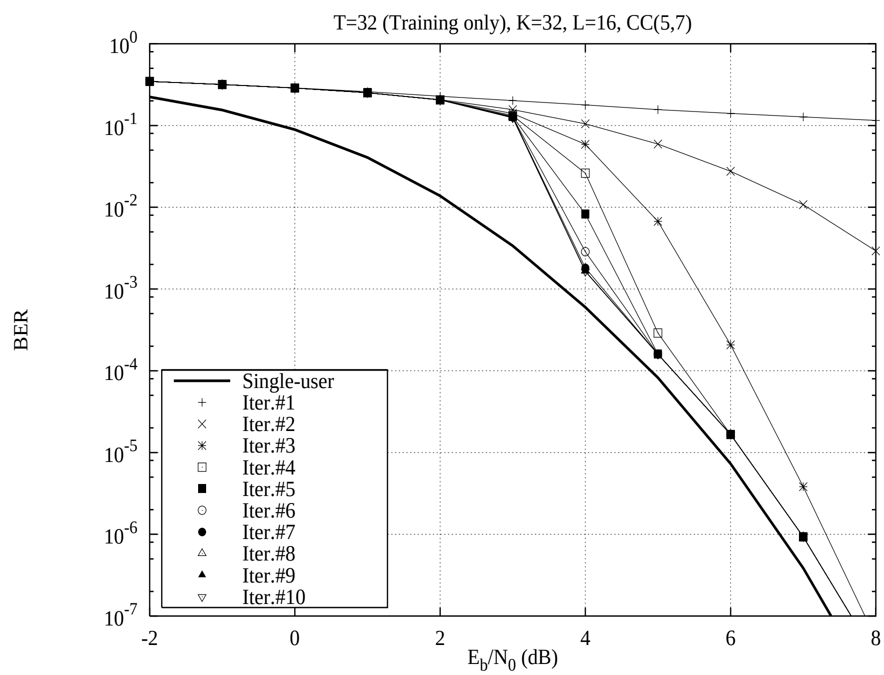

Figure 4: $K = 32, L = 16,\ N = 2000$, CC $(5,7)_8$, training-only estimation with $T = 32$.

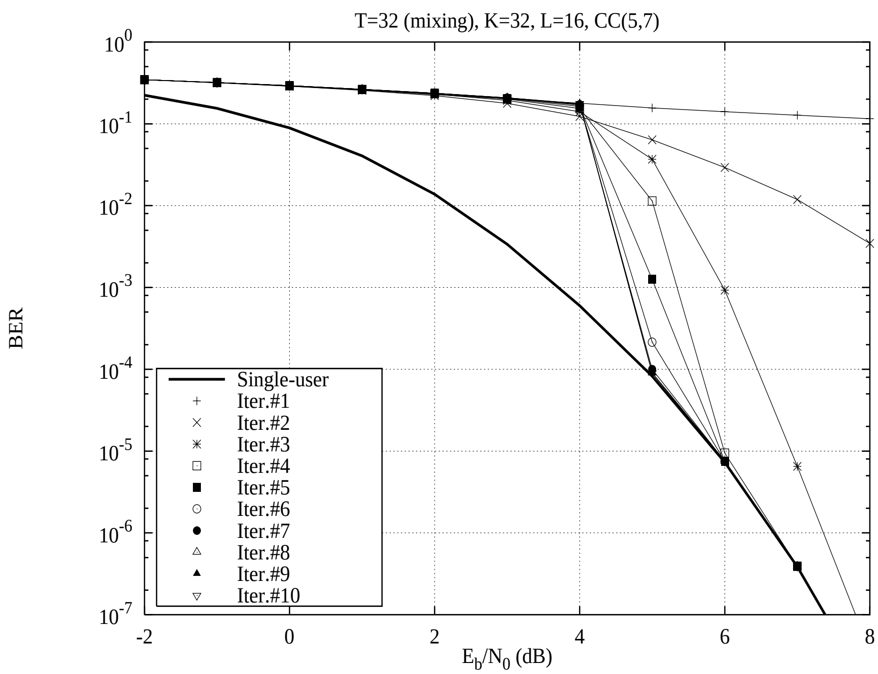

Figure 5: $K = 32, L = 16$, $N = 2000$, CC $(5, 7)_8$, EM+training estimation with $T = 32$ and the mixing method.

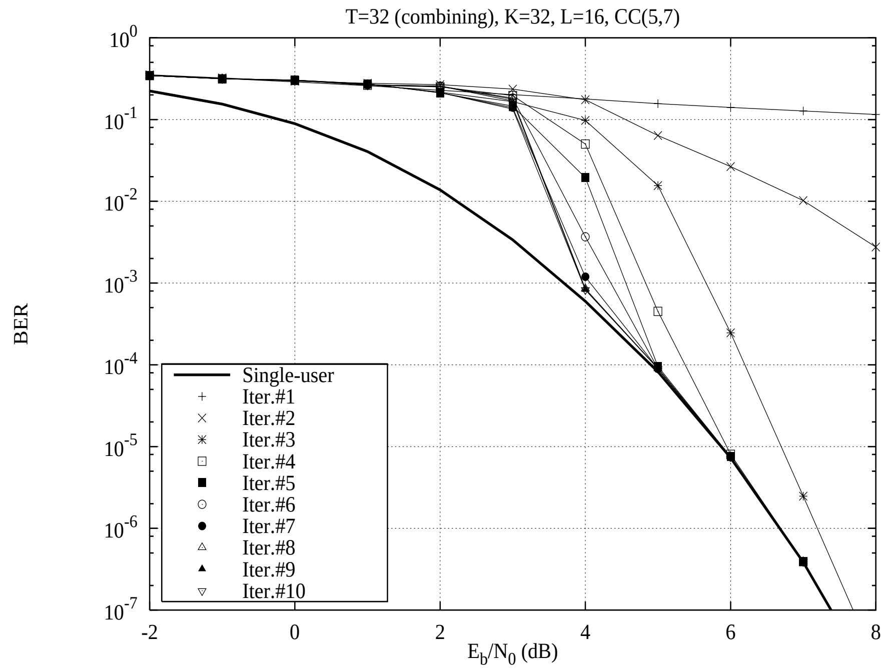

Figure 6: $K = 32, L = 16$, $N = 2000$, CC $(5,7)_8$, EM+training estimation with $T = 32$ and the combining method.

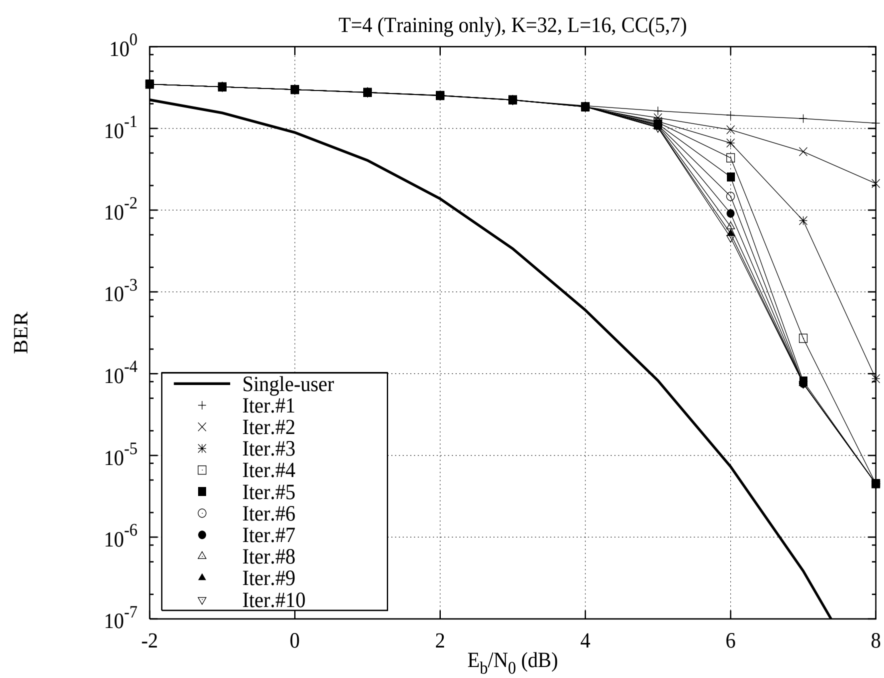

Figure 7: $K = 32, L = 16$, $N = 2000$, CC $(5,7)_8$, training-only estimation with $T = 4$.

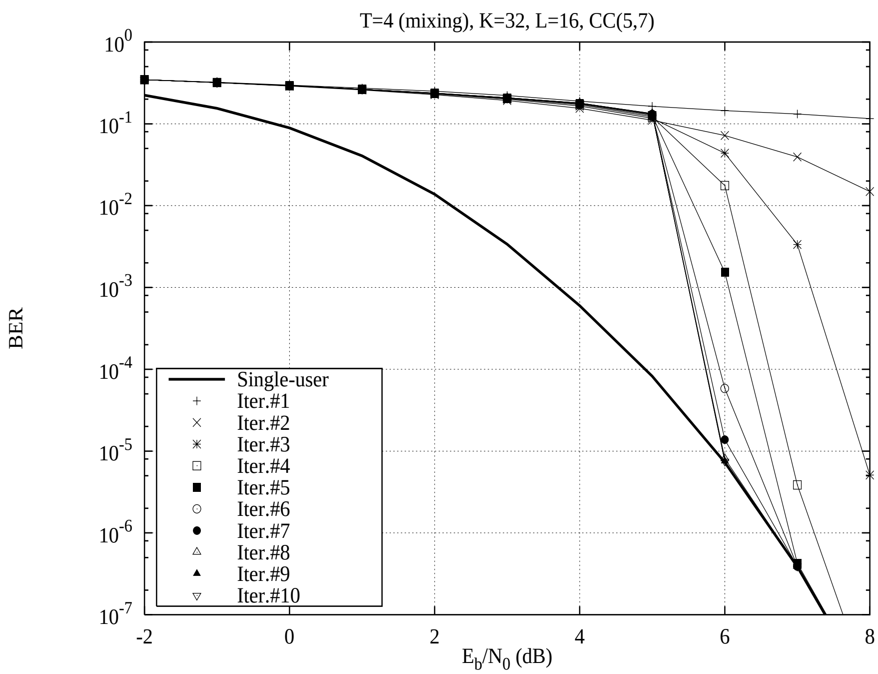

Figure 8: $K = 32, L = 16$, $N = 2000$, CC $(5, 7)_8$, EM+training estimation with $T = 4$ and the mixing method.

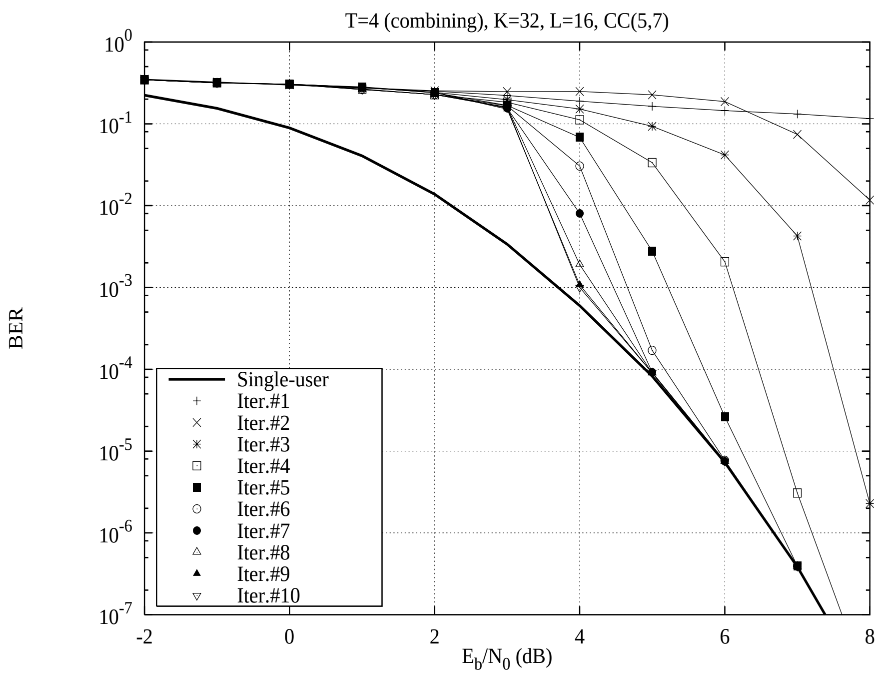

Figure 9: $K = 32, L = 16$, $N = 2000$, CC $(5,7)_8$, EM+training estimation with $T = 4$ and the combining method.

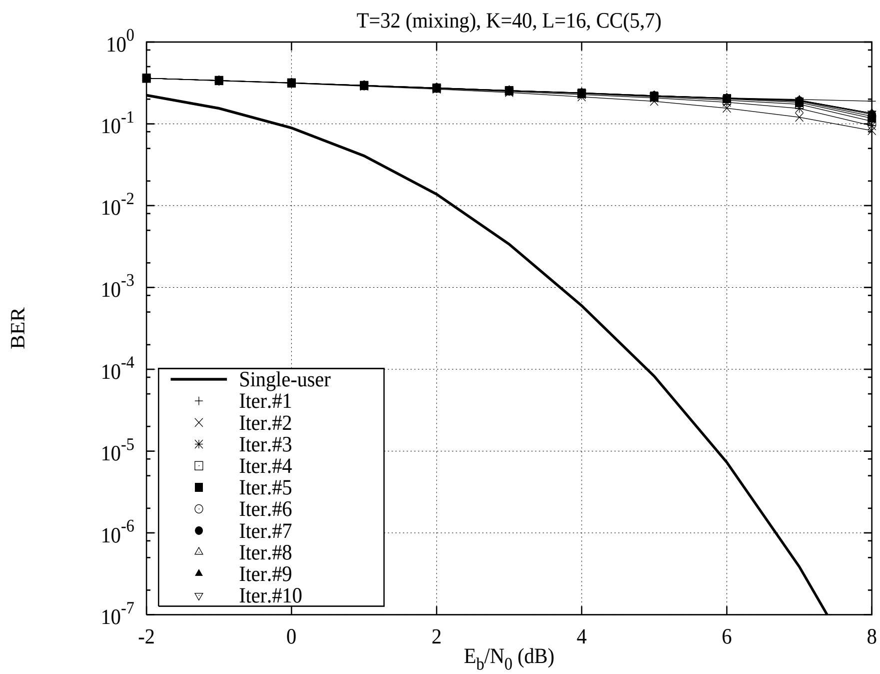

Figure 10: $K = 40, L = 16$, $N = 2000$, CC $(5, 7)_8$, EM+training estimation with $T = 32$ and the mixing method.

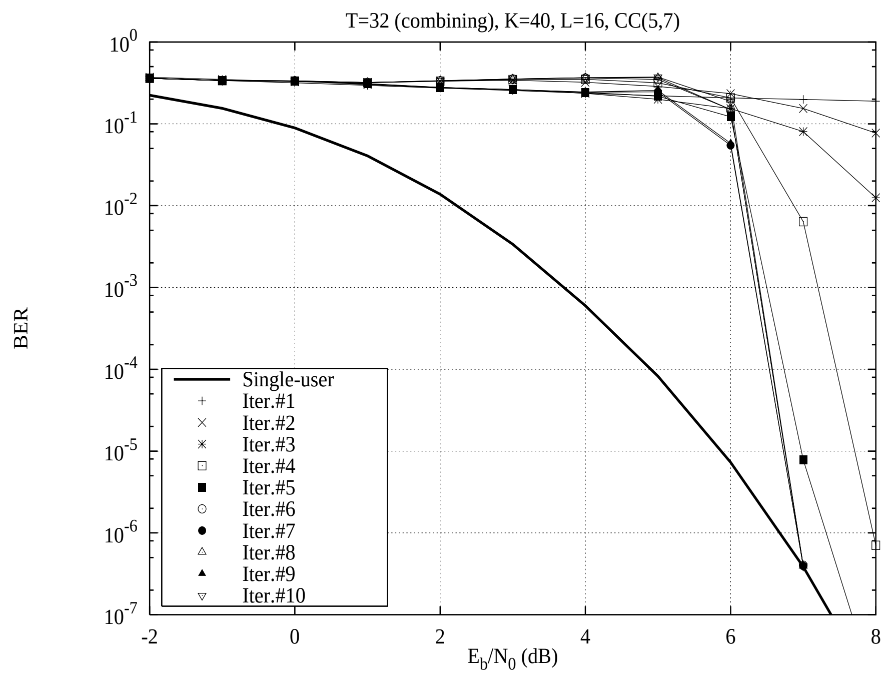

Figure 11: $K = 40, L = 16$, $N = 2000$, CC $(5,7)_8$, EM+training estimation with $T = 32$ and the combining method.

List of Figures