# Iterative Joint Detection, Decoding, and Channel Estimation for Dual Antenna Arrays in Frequency Selective Fading

**Maja Lončar*** **Ralf Müller** **Joachim Wehinger**
Telecommunications Research Center Vienna (ftw.)
Vienna, Austria
{loncar, mueller, wehinger}@ftw.at

**Tetsushi Abe**
NTT DoCoMo Inc.
Yokosuka-chi, Kanagawa, Japan
abe@mlab.yrp.nttdocomo.co.jp

## Abstract

Iterative receiver algorithms for multi-user detection and decoding are a solution for mitigating multiple-access interference (MAI) in multiple access communication systems. After several iterations they outperform non-iterative schemes, while maintaining modest complexity. In this paper, we present an iterative receiver for multiple-antenna systems in frequency selective fading. Channel knowledge is not assumed to be available at the receiver. We propose a comprehensive algorithm that performs channel estimation, multi-user detection and decoding in an iterative manner. The multipath channel is estimated using several pilot symbols and soft estimates of data symbols. The performance of the receiver is evaluated by simulations.

## INTRODUCTION

In multi-user communication systems signals are subject to the interference originating from other users' signals, i.e. MAI. As an approach to combating this effect, iterative multi-user detectors were proposed by several authors for code division multiple access (CDMA) as well as for multiple antenna systems (see [1–13] and references therein). The common realization of the detector, also adopted in this paper, is the soft interference canceller with post-filtering. Using soft symbol decisions in the cancellation process eliminates the error propagation effect that occurs in hard decision schemes. For coded transmission, interference cancellation and filtering is followed by single-user soft-in soft-out (SISO) decoding. Throughout the iterative process, soft estimates of coded symbols, obtained from the decoders' outputs in one iteration are used as feedback information for interference cancellation in the next iteration. A unified treatment of iterative joint detection and decoding for CDMA, based on the factor-graph approach, can be found in [13]. It is shown that various low-complexity interference cancellation algorithms can be obtained as approximations of the optimum high-complexity receiver. In [13] and in most of the approaches discussed in literature it is assumed that the receiver has perfect knowledge of the channel state. In real applications, however, non-perfect channel estimation has to be taken into account. Estimation of the frequency flat fading channels as a part of the iterative CDMA receiver algorithm is discussed in [8], [10] and [11]. In this paper, we consider frequency-selective fading as the more general scenario, and include the multipath channel estimator in the iterative receiver scheme. Iterative channel estimation for frequency selective multiple-input multiple-output (MIMO) channels was addressed in [12] where estimates are obtained using pilot symbols and hard decisions of "reliable" data symbols. We extend this idea and perform channel estimation using soft decisions of all the symbols in a data block. This leads to significant improvement of channel estimates, and, subsequently, of the overall receiver performance. The equivalent application of this approach to CDMA systems is provided in [14].

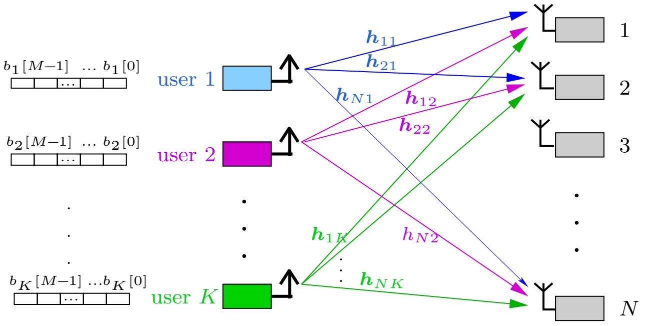

**Figure 1. MIMO communication system**

## SIGNAL MODEL

Let $K$ be the number of transmit antennas and $N$ the number of receive antennas in a MIMO system depicted in Fig. 1. It is a multiple access system, where each antenna can be viewed as a user that "sees" a single-input multiple-output (SIMO) channel. Users transmit blocks of $M$ convolutionally encoded and randomly interleaved binary symbols $b_k[m] \in \{-1, +1\}$, where $k = 1, ..., K$ is the transmit antenna (user) index and $m$ is the discrete symbol-time. The random interleavers are different for each user. The signals propagate through the frequency selective fading channel of memory length $L$. The fading processes on all receive antennas are jointly independent, which is justified by the fact that the transmitters are assumed to have different positions in a hypothetical cellular system. The channel impulse responses remain constant during transmission of one data block (block fading) and change randomly from block to block. Consecutive data blocks are separated by a guard interval, long enough to prevent inter-block interference (IBI). Each block

---

*Part of this work was supported by NTT DoCoMo, while Maja Lončar was visiting Wireless Laboratories, NTT DoCoMo, Japan.

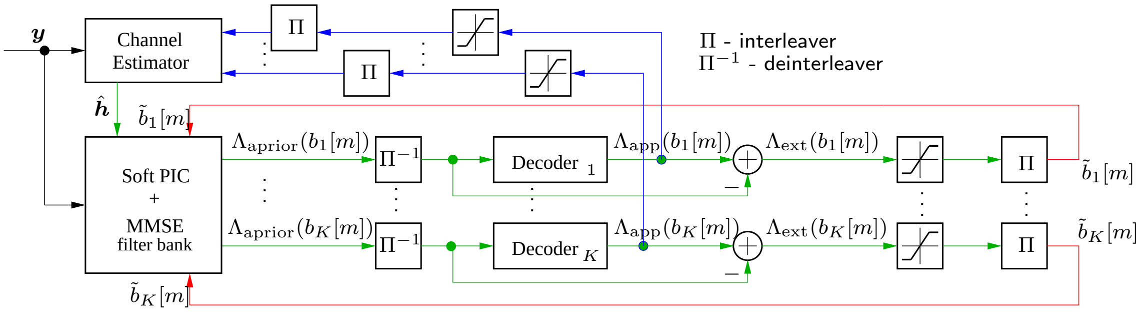

**Figure 2. Iterative receiver – block structure**

of data contains a preamble of $P$ pilot symbols known at the receiver.

The signal received at the $n$-th antenna, $n = 1, \dots, N$, in the $m$-th symbol interval is given by

$$y_n[m] = \sum_{k=1}^{K} \sum_{l=0}^{L-1} h_{nk}[l]\, b_k[m-l] + v_n[m], \qquad (1)$$

where $h_{nk}[l]$ is the $l$-th tap of the channel impulse response from the $k$-th transmit to the $n$-th receive antenna and $v_n[m]$ is spatially and temporally white Gaussian noise, with zero mean and known variance $\sigma_v^2$. Since there is no IBI, $M$-symbol long blocks transmitted from $K$ antennas will result in $M+L-1$ received symbols on the $n$-th receive antenna, defined by (1). Collecting them into a vector $\boldsymbol{y}_n = [\, y_n[0] \;\; y_n[1] \;\; \dots \;\; y_n[M+L-2]\,]^T$, we obtain:

$$\boldsymbol{y}_n = \sum_{k=1}^{K} \boldsymbol{B}_k \boldsymbol{h}_{nk} + \boldsymbol{v}_n, \qquad (2)$$

where

$$\boldsymbol{B}_k = \begin{bmatrix} b_k[0] & 0 & \cdots & 0 \\ b_k[1] & b_k[0] & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ b_k[L-1] & b_k[L-2] & \cdots & b_k[0] \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & b_k[M-1] \end{bmatrix}$$

is an $(M+L-1) \times L$ Toeplitz matrix of $k$-th user's data symbols, $\boldsymbol{h}_{nk} = [\, h_{nk}[0] \;\; h_{nk}[1] \;\; \dots \;\; h_{nk}[L-1]\,]^T$ contains the $L$ channel taps and $\boldsymbol{v}_n = [\, v_n[0] \;\; v_n[1] \;\; \dots \;\; v_n[M+L-2]\,]^T$ is the noise vector.

The equation (2) can be written in matrix form as

$$\boldsymbol{y}_n = \underbrace{[\boldsymbol{B}_1 \; \boldsymbol{B}_2 \dots \boldsymbol{B}_K]}_{\substack{\triangleq \\ \boldsymbol{B}}} \underbrace{\begin{bmatrix} \boldsymbol{h}_{n1} \\ \vdots \\ \boldsymbol{h}_{nK} \end{bmatrix}}_{\substack{\triangleq \\ \boldsymbol{h}_n}} + \boldsymbol{v}_n = \boldsymbol{B}\, \boldsymbol{h}_n + \boldsymbol{v}_n \,.$$

Finally, we stack these vectors for all $N$ receive antennas into a single $N(M+L-1) \times 1$ vector $\boldsymbol{y}$:

$$\boldsymbol{y} = (\boldsymbol{I}_N \otimes \boldsymbol{B})\, \boldsymbol{h} + \boldsymbol{v} \triangleq \boldsymbol{\mathcal{B}}\, \boldsymbol{h} + \boldsymbol{v}, \qquad (3)$$

where $\boldsymbol{y} = \left[\boldsymbol{y}_1^T \; \boldsymbol{y}_2^T \dots \boldsymbol{y}_N^T\right]^T$, $\boldsymbol{h} = \left[\boldsymbol{h}_1^T \; \boldsymbol{h}_2^T \dots \boldsymbol{h}_N^T\right]^T$, $\boldsymbol{v} = \left[\boldsymbol{v}_1^T \; \boldsymbol{v}_2^T \dots \boldsymbol{v}_N^T\right]^T$, $\boldsymbol{I}_N$ is the $N \times N$ identity matrix, $\otimes$ is the Kronecker product, and $\boldsymbol{\mathcal{B}}$ is an $N(M+L-1) \times NKL$ block diagonal data matrix. This signal model is suitable for channel estimation since all the $NKL$ unknown channel parameters are collected in the vector $\boldsymbol{h}$.

## ITERATIVE RECEIVER ALGORITHM

The structure of the investigated iterative receiver is shown in Fig. 2. It consists of three major blocks: channel estimator, multi-user detector and a bank of $K$ single-user SISO decoders. The detector is realized as a soft parallel interference canceller (PIC) followed by $K$ linear MMSE filters, separate for each user. Each block will be explained in the following.

### Channel Estimation

*Maximum A Posteriori (MAP) Joint Channel and Data Estimation*

Based on $N(M+L-1)$ observations (collected in the received vector $\boldsymbol{y}$), we need to estimate $KNL$ channel coefficients $h_{nk}[l]$ and $KM$ data symbols $b_k[m]$. The optimum approach is joint estimation of all the parameters. The joint estimator does not take into account the code constraints. The solution to this estimation problem is obtained by maximizing the joint posterior probability density function (pdf) of $\boldsymbol{\mathcal{B}}$ and $\boldsymbol{h}$:

$$(\boldsymbol{h},\, \boldsymbol{\mathcal{B}}) = \underset{\boldsymbol{h},\boldsymbol{\mathcal{B}}}{\arg\max} \left\{ f(\boldsymbol{h}, \boldsymbol{\mathcal{B}} \,|\, \boldsymbol{y}) \right\}.$$

Using Bayes' rule and assuming that the channel realizations are independent of the transmitted data we can write:

$$f(\boldsymbol{h}, \boldsymbol{\mathcal{B}} \,|\, \boldsymbol{y}) \propto f(\boldsymbol{y} \,|\, \boldsymbol{h}, \boldsymbol{\mathcal{B}})\, f(\boldsymbol{h})\, \Pr(\boldsymbol{\mathcal{B}}), \qquad (4)$$

where $\propto$ denotes proportionality up to a constant. The conditional distribution of the received sequence given the channel state and the data symbols is an $N(M+L-1)$–dimensional complex Gaussian distribution:

$$f(\boldsymbol{y}|\boldsymbol{\mathcal{B}},\boldsymbol{h}) = \left(\pi\sigma_v^2\right)^{-N(M+L-1)} e^{-\sigma_v^{-2}(\boldsymbol{y}-\boldsymbol{\mathcal{B}}\boldsymbol{h})^H(\boldsymbol{y}-\boldsymbol{\mathcal{B}}\boldsymbol{h})}. \quad (5)$$

Since the coded data symbols are considered to be mutually independent (due to random interleaving), the joint probability mass function (pmf) of the data can be factorized as

$$\Pr(\boldsymbol{\mathcal{B}}) = \prod_{m=0}^{M-1}\prod_{k=1}^{K}\Pr(b_k[m]).$$

Obviously, for each data block there exist $2^{KM}$ possible hypotheses on transmitted symbols. In order to maximize the function given by (4), it would be necessary to perform a search over all possible sequences and to find the most likely channel realization for each of them. Due to prohibitive complexity of this approach, suboptimum solutions are needed. They are obtained by approximating the true probability mass function of data symbols $\Pr(b_k[m])$, as depicted in Fig. 3 (see [13]).

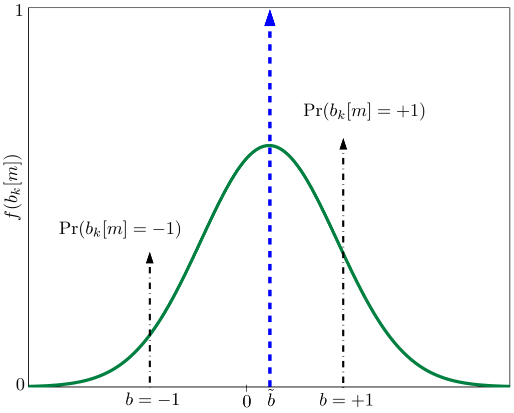

**Figure 3. True and approximated pdf of a data symbol (- - - single mass point and —– Gaussian approximation)**

### *Soft Discrete Symbol Approximation*

In [11] it is proposed to replace $\Pr(b_k[m])$ by a continuous Gaussian pdf of the same mean and variance. Having found this approach very complex to follow, we use a discrete single point approximation of the data symbol pmf (this is the limiting case of the Gaussian pdf with zero variance and the same mean). Thus, we replace each unknown symbol with a soft value equal to the mean $\tilde{b}_k[m]$ calculated as

$$\tilde{b}_k[m] = \sum_{b\in\{+1,-1\}} b\,\Pr(b_k[m]=b), \quad (6)$$

where $\Pr(b_k[m]=b)$ is the a-posteriori probability (APP) of the coded symbol from the output of the $k$-th SISO decoder. If we let $\tilde{\boldsymbol{\mathcal{B}}}$ denote the matrix obtained by replacing the unknown data symbols in matrix $\boldsymbol{\mathcal{B}}$ by their soft mean values, then the function we wish to maximize equals to (5), with $\tilde{\boldsymbol{\mathcal{B}}}$ instead of $\boldsymbol{\mathcal{B}}$. Assuming that the channel coefficients are independent, identically distributed complex Gaussian random variables, with zero mean and unit variance, the MAP–like channel estimator coincides with the linear MMSE estimator, yielding the following result:

$$\begin{aligned}\hat{\boldsymbol{h}}_{\mathrm{MAP}} &= \arg\max_{\boldsymbol{h}}\left\{\ln f(\boldsymbol{y}|\boldsymbol{h}) + \ln f(\boldsymbol{h})\right\}\\ &= \arg\min_{\boldsymbol{h}}\left\{\frac{1}{\sigma_v^2}\left\|\boldsymbol{y}-\tilde{\boldsymbol{\mathcal{B}}}\boldsymbol{h}\right\|^2 + \|\boldsymbol{h}\|^2\right\}\\ &= \left(\tilde{\boldsymbol{\mathcal{B}}}^H\tilde{\boldsymbol{\mathcal{B}}} + \sigma_v^2\boldsymbol{I}\right)^{-1}\tilde{\boldsymbol{\mathcal{B}}}^H\boldsymbol{y}. \quad (7)\end{aligned}$$

This solution requires knowledge of the noise variance at the receiver. Without any prior statistical knowledge, the estimator reduces to the least-squares (LS) solution, which can be decoupled for each antenna, yielding:

$$\hat{\boldsymbol{h}}_{n\,\mathrm{LS}} = \left(\tilde{\boldsymbol{B}}^H\tilde{\boldsymbol{B}}\right)^{-1}\tilde{\boldsymbol{B}}^H\boldsymbol{y}_n, \qquad n=1,\ldots,N. \quad (8)$$

Thus, only one inversion of a $KL\times KL$ matrix $\tilde{\boldsymbol{B}}$ is required for all $n$.

### Soft IC and MMSE Filtering

The detailed algorithm for IC and post-filtering can be found in [12]. We will outline only the main principles.

The inputs to the soft IC stage in one iteration are the received sequence $\boldsymbol{y}$, the channel estimate $\hat{\boldsymbol{h}}$ and the extrinsic (EXT) probabilities $\Pr_{\mathrm{ext}}(b_k[m])$ of coded symbols computed by the SISO decoders in the previous iteration. Soft symbol estimates are computed according to (6), using EXT probabilities instead of APPs. Replicas of MAI and ISI components for each user are calculated and subtracted from the total received signal. After soft IC, the resulting signal contains the symbol of the user of interest and the residual interference due to non-perfect channel and data estimation. Therefore, the signal is further processed by an MMSE filter. The filter output is a soft value $z_k[m]$ closest to the true symbol value $b_k[m]$ in the mean-squared error sense, and it is delivered as a-priori information to the $k$-th SISO decoder.

### SISO Decoding

Single-user SISO decoders for convolutionally encoded symbols are realized using the BCJR algorithm ([15],[2]). In every iteration, based on the outputs of the MMSE filter $\boldsymbol{z}_k$ and knowledge of the code constraints (c.c.), each user's decoder computes the APP for each coded symbol in a block $\mathrm{APP}(b_k[m]{=}b) = \Pr(b_k[m]{=}b\,|\,\boldsymbol{z}_k, \text{c.c.})$. These probabilities are used as feedback information to the channel estimator. For the soft IC, however, it is necessary to use extrinsic information (see [13]). If the a-posteriori and extrinsic information are expressed in terms of log-likelihood ratios (LLR) $\Lambda_{\mathrm{app}}$ and $\Lambda_{\mathrm{ext}}$, respectively, then $\Lambda_{\mathrm{app}} = \Lambda_{\mathrm{aprior}} + \Lambda_{\mathrm{ext}}$, where $\Lambda_{\mathrm{aprior}}$ is the a-priori information that the decoder obtains from the output of the MMSE filter. In the final iteration, the decoder outputs the hard decisions on the information bits.

## SIMULATION RESULTS

We simulated the iterative receiver for a MIMO system with 2 receive antennas and 2-4 users (transmit antennas). The transmitted data blocks were encoded with a convolutional code of rate $1/2$ and constraint length 3. The block length was either 300 (short blocks) or 900 (long blocks) coded symbols. Each block was preceded by a short (15 symbols) or long (50 symbols) pilot sequence. We limited the number of iterations to 4 or 5. As a performance measure for channel estimation we used the average relative estimation error defined as $\delta = \mathsf{E}\left\{||\boldsymbol{h} - \hat{\boldsymbol{h}}||^2/||\boldsymbol{h}||^2\right\}$. Fig. 4 shows the decrease of $\delta$ during iterations for different noise levels. In the first iteration, the channel is estimated over pilot symbols only. In subsequent iterations, the estimate is significantly improved using the soft symbol estimates of the whole block. In Fig. 5 we compare MMSE versus LS channel estimation for short data blocks and different number of pilot symbols. As expected, for short pilot sequences, MMSE benefits from additional knowledge of the noise variance and yields better estimates in the first iteration (pilot-only estimation). After 4 iterations, however, the initial loss of LS estimation is compensated for by the iteration gain and both estimators converge to the same residual error. Thus, we conclude that for iterative estimation it is sufficient to use the simpler LS approach. Note also that for higher SNR reduced pilot sequence length allows sufficiently good channel estimates.

Fig. 6 shows the performance in terms of the bit-error-rate (BER) after first and fourth iteration for the receiver with LS channel estimator that uses soft data feedback based on the decoders' APPs. The iteration gain is large. For comparison, we also plotted the performance of the same receiver performing pilot-only estimation, or using soft symbol estimates computed from EXT information. APP-based iterative channel estimation over the whole data block by far outperforms non-iterative, pilot-only estimation. Using EXT instead of APP for the feedback to the channel estimator significantly degrades performance. Note that this does not hold for feedback from the detectors to the interference canceller (it uses EXT only). This result justifies the iterative structure in Fig. 2. The single-user bound in Fig. 7 represents the receiver's performance when there is only one user in the system (no MAI), and the channel is perfectly known at the receiver. The performance gap with respect to this bound is about 1.4 dB. This result is obtained using short data blocks, which are disadvantageous in our setup. Since fading is assumed block-constant, longer data blocks allow better channel estimation, which yields lower BER for longer blocks. This can be observed by comparing Fig. 6 with Fig. 7, which shows the BER for long blocks, with 15 and 50 pilot symbols. Longer training sequences allow better initial channel estimation which leads to a BER that approaches the BER of 2 users with perfectly known channel. The reduction of pilot sequence length to 15 symbols causes a loss of less than 0.5dB for SNR$\geq 4$ dB.

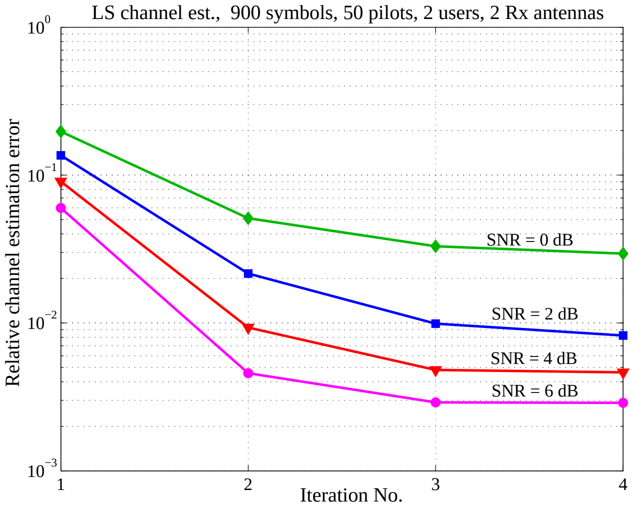

**Figure 4. Reduction of channel estimation error by iterative estimation**

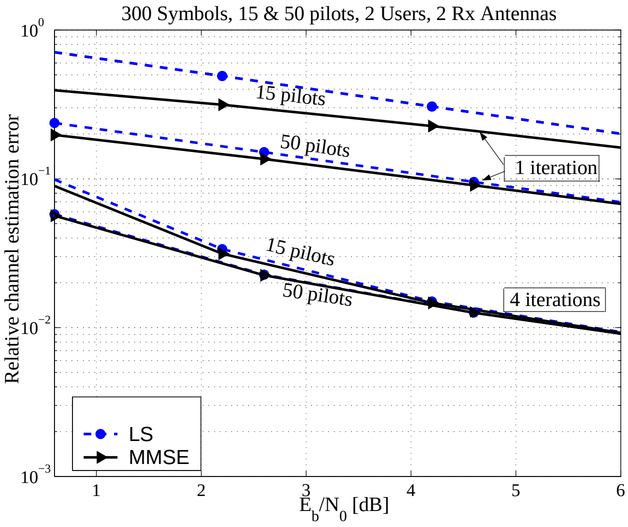

**Figure 5. Channel estimation error for LS and MMSE estimator**

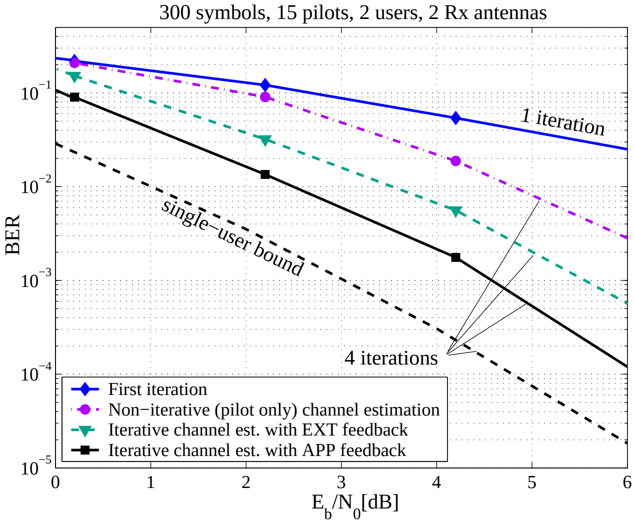

**Figure 6. BER for different channel estimation approaches**

Fig. 8 depicts the receiver's performance when we increase the number of users $K$, while keeping the number of receive antennas $N = 2$. The number of iterations is 5. Iterative receiver can successfully separate 3 users, with a BER that is very close to the 2-user case. The performance degrades for $K = 4$ (i.e. $K/N = 2$). We also used this setup with 4 users to compare the performance of soft decision feedback (from the decoders to the channel estimator) with hard decision feedback of "reliable" symbols only, as proposed in [12]. The symbols are identified as reliable by comparing their $\Lambda_{\text{app}}$ in every iteration with a certain threshold value. We used the value of 0.5. Estimating the channel over tentative soft symbol estimates yields much better results compared to the case of using hard decisions. This difference is more pronounced for lower SNR, in larger systems and in case of transmission with shorter pilot sequences.

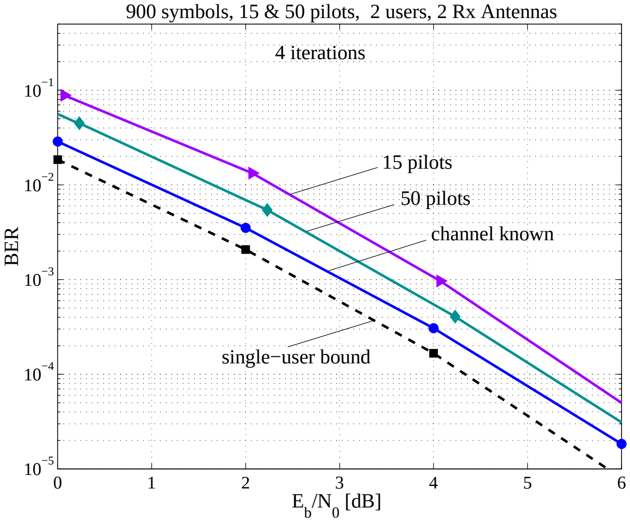

**Figure 7. BER for estimated and perfectly known channel**

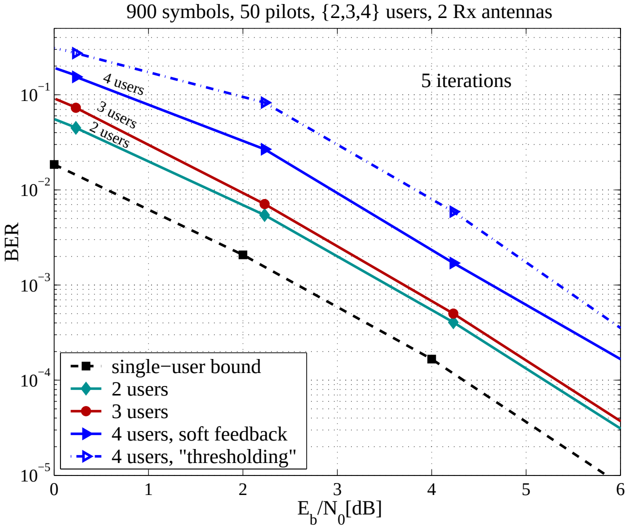

**Figure 8. BER for different number of users in the system**

## SUMMARY AND CONCLUSIONS

An iterative receiver is proposed for MIMO systems in frequency-selective fading channels. The receiver performs iterative channel estimation, soft interference cancellation with MMSE post-filtering and SISO decoding. Modified MMSE and LS channel estimators, which use the whole block of received data for estimating the channel, are derived. The unknown data symbols are replaced by their soft estimates. Soft-symbol estimates used in channel estimation should be based on APPs, while the IC block uses EXT-based estimates. Simulations show that the LS and MMSE estimators perform identically after few iterations. The iterative receiver allows accommodating a higher number of users without significant performance loss. The soft feedback approach outperforms the hard decision feedback, particularly for increased number of users and shorter pilot sequences.

## REFERENCES

[1] F. Tarköy, "Iterative multi-user decoding for asynchronous users," in Proc. IEEE ISIT, 1997, p. 30.

[2] X. Wang and V. H. Poor, "Iterative (turbo) soft interference cancellation and decoding for coded CDMA," IEEE Trans. Comm., vol. 47, no. 7, pp. 1046–1061, July 1999.

[3] M. C. Reed and P. D. Alexander, "Iterative mulituser detection using antenna arrays and FEC on multipath channels," IEEE J. Sel. Ar. Comm., vol. 17, no. 12, pp. 2082–2089, Dec. 1999.

[4] D. Reynolds and X. Wang, "Low-complexity turbo-equalization for diversity channels," Signal Processing, Elsevier Science Publishers, vol. 81, no. 5, pp. 989–995, May 2000.

[5] P. Alexander and A. Grant, "Iterative channel and information sequence estimation in CDMA," in IEEE Int. Symp. on Spread-Spectrum Tech. & Appli., NJIT, 2000, pp. 593–597.

[6] C. Grant, A. Schlegel, "Differential turbo space-time coding," in Proc. of Info. Theory Workshop, 2001, pp. 120–122.

[7] Q. Li, N. Georghiades, and X. Wang, "An iterative receiver for turbo-coded pilot-assisted modulation in fading channels," IEEE Comm. Letters, vol. 5, no. 4, pp. 145–147, Apr. 2001.

[8] M. Kobayashi, J. Boutros, and G. Caire, "Successive interference cancellation with SISO decoding and EM channel estimation," IEEE J. Sel. Ar. Comm., vol. 19, no. 8, Aug. 2001.

[9] Z. Shi and C. Schlegel, "Joint iterative decoding of serially concatenated error control coded CDMA," IEEE J. Sel. Ar. Comm., vol. 19, no. 8, pp. 1646–1653, Aug. 2001.

[10] M. Valenti and B. Woerner, "Iterative channel estimation and decoding of pilot symbol assisted turbo codes over flat-fading channels," IEEE J. Sel. Ar. Comm., vol. 19, no. 9, pp. 1697–1705, Sept. 2001.

[11] G. Caire, A. Tulino, and E. Biglieri, "Iterative multiuser joint detection and parameter estimation: A factor-graph approach," in IEEE Proc. Information Theory Workshop, Cairns, Australia, Sept. 2001.

[12] T. Abe and T. Matsumoto, "Space–time turbo equalization and symbol detection for frequency selective MIMO channels," in Proc. of VTC 2001 Fall, Oct. 2001, vol. 2, pp. 1230–1234.

[13] J. Boutros and C. Caire, "Iterative multiuser joint decoding: Unified framework and asymptotic analysis," IEEE Trans. on Inform. Theory, vol. 48, no. 7, pp. 1772–1793, July 2002.

[14] J. Wehinger, R. Müller, M. Lončar, and C. Mecklenbräuker, "Performance of iterative CDMA receiver with channel estimation in multipath environments," submitted to Asilomar Conference on Signals, Systems and Computers, 2002.

[15] L.R. Bahl, J. Cocke, F. Jelinek, and J. Raviv, "Optimal decoding of linear codes for minimizing symbol error rate," IEEE Trans. Inf. Theor., vol. 20, pp. 284–287, Mar. 1974.