EURASIP Journal on Applied Signal Processing 2005:6, 872–882

Iterative PDF Estimation-Based Multiuser Diversity Detection and Channel Estimation with Unknown Interference

Nenad Veselinovic
Centre for Wireless Communications, University of Oulu, P.O. Box 4500, 90014 Oulu, Finland
Email: nenad.veselinovic@ee.oulu.fi

Tad Matsumoto
Centre for Wireless Communications, University of Oulu, P.O. Box 4500, 90014 Oulu, Finland
Email: tadashi.matsumoto@ee.oulu.fi

Markku Juntti
Centre for Wireless Communications, University of Oulu, P.O. Box 4500, 90014 Oulu, Finland
Email: markku.juntti@ee.oulu.fi

Received 8 October 2003; Revised 14 July 2004

The equivalent diversity order of multiuser detector employing multiple receive antennas and minimum mean squared error (MMSE) processing for frequency-selective channels is decreased if it aims at suppressing unknown cochannel interference (UCCI) while detecting multiple users' signals. This is an unavoidable consequence of linear processing at the receiver. In this paper, we propose a new multiuser signal detection scheme with the aim to preserve the detector's diversity order by taking into account the structure of the UCCI. We use the fact that the structure of the UCCI appears in the probability density function (PDF) of the UCCI plus noise, which can be characterized as multimodal Gaussian. A kernel smoothing PDF estimation-based receiver is derived. The PDF estimation can be based on training symbols only (noniterative PDF estimation) or on training symbols as well as feedback from the decoder (iterative PDF estimation). It is verified through simulations that the proposed receiver significantly outperforms the conventional covariance estimation in channels with low frequency selectivity. The iterative PDF estimation significantly outperforms the noniterative PDF estimation-based receiver with minor training overhead.

Keywords and phrases: turbo equalization, cochannel interference, PDF estimation.

1. INTRODUCTION

The scarcity of the frequency resources and the fact that the frequency spectrum has to be shared by multiple users in future wireless communication systems impose the need for bandwidth-efficient transceiver schemes. A huge volume of research has been done on the development of different techniques for multiple access, the most important examples of which are frequency-division multiple access (FDMA), time-division multiple access (TDMA), and code-division multiple access (CDMA).

The advances in the area of communications using multiple receive antennas have opened a completely new dimension for combating interference, called space-division multiple access (SDMA) [1, 2]. The SDMA concept can be applied to any of the existing multiple-access schemes to further improve the system capacity both in terms of the number of supported users and in terms of supported data rates. Moreover, SDMA can be seen as bandwidth efficient by analogy to CDMA, where the orthogonality between users is maintained by their unique spatial signatures instead of unique spreading waveforms [3]. This, at least in terms of baseband signal processing, offers new possibilities of using large preexisting knowledge of CDMA.

An example of the area where a large experience is present in the research community is multiuser detection for CDMA [4]. It is well known that the maximum-likelihood sequence estimation (MLSE) technique achieves the best performance when detecting the multiple users' transmitted signals. However, its computational complexity, which increases exponentially with the number of users and memory length of the channel, is prohibitive for a practical use. Therefore, a significant amount of research has been conducted to develop suboptimal multiuser receivers [4]. In coded systems,

the complexity of the optimal receiver further increases due to the fact that joint trellis diagram of all the users, their multipath channels, and their channel codes has to be taken into account [5]. In [6], low-complexity receivers that separately perform detection and decoding stages are proposed.

A breakthrough in the development of the suboptimal low-complexity receiver schemes has been initiated by the discovery of turbo codes [7]. The principle of turbo processing [8] has been shown to be extremely powerful in solving the computational complexity problem of the optimal receiver structures. It is based on the concept that different parts of the receiver perform locally optimal signal processing, conditioned on the processing results of the other receiver blocks. By iteratively performing such a processing, near-globally optimal performance can be obtained in various cases. Examples are joint equalization and decoding [9], joint multiuser detection and decoding in CDMA [10, 11, 12], and joint MIMO multiuser detection, equalization, channel estimation, and decoding [13]. The common structure of these receivers consists of the maximum *a posteriori* (MAP) block for multiuser detection and equalization, and a set of soft-input soft-output (SISO) channel decoders, which are separated by interleavers [10, 11]. In [10] and [13], a low-complexity implementation of the iterative receivers is also proposed. It is based on soft interference cancellation and linear minimum mean squared error (SC/MMSE) filtering.

The SC/MMSE equalizer is robust against unknown cochannel interference (UCCI) if the covariance matrix of the UCCI is properly estimated and taken into account in the MMSE filtering [1, 13]. Subspace estimation and projection [14, 15] is another UCCI cancellation technique. However, to suppress UCCI, the both methods unavoidably consume the degrees of freedom (DOFs) provided by spatio-temporal processing in the receiver. This is a consequence of linear signal processing at the receiver that does not take into account the actual structure of the UCCI. This will result in a decrease of the overall diversity order of the receiver [15, 16, 17]. The loss of diversity will be more severe in channels with low frequency selectivity due to the lack of multipath diversity.

If the signal constellation of the UCCI is known at the receiver, the channel of UCCI can be estimated, and the diversity loss of the linear MMSE receiver can be completely recovered by means of joint maximum-likelihood (ML) detection of the desired users and UCCI [17]. This in turn would require either blind channel estimation methods to estimate the channels of UCCIs or blind source separation techniques [18] to estimate jointly the channels and data sequences. However, if the UCCI's channels do not change significantly over the frame, possible states of the interference can be estimated instead of estimating the channel gains themselves. This fact has been used in Bayesian equalization in the presence of UCCI in [19]. In [20], maximum-likelihood sequence estimation (MLSE) equalization was performed in combination with UCCI-plus-noise PDF estimation to combat the impact of UCCI.

In [21], the authors have derived a new receiver that can preserve the diversity gain by estimating the PDF of the UCCI plus noise. The signal processing algorithm shown in [22] is used in the first iteration and the kernel-based PDF estimation [20, 23] is applied for the following iterations. It is shown there that the proposed receiver significantly outperforms the conventional detector of [22] in low frequency-selective channels with relatively small number of UCCIs. There, however, the receiver is restricted to noniterative PDF estimation and it was derived only for binary-phase-shift-keying (BPSK) modulation. In this paper, we generalize the receiver derivation to the multilevel-phase-shift-keying (MPSK) case. Furthermore, an iterative PDF estimation technique using soft feedback is proposed for situations where only short training sequences are available. It is shown that the proposed joint iterative PDF estimation and turbo signal detection technique can significantly improve performance over the noniterative technique, when only short training sequences are available. We restrict the scope of the paper to the multilevel phase-shift-keying (PSK) modulation, but it is straightforward to extend the concept to quadrature amplitude modulation (QAM) cases. The rest of the paper is organized as follows. Section 2 describes system model. Sections 3 and 4 present the conventional and the proposed receivers, respectively. Section 5 presents simulation results, and Section 6 concludes the paper.

2. SYSTEM MODEL

Figure 1 illustrates the system model. Each of $N + N_I$ users encodes information sequence $c_n(i)$, $n = 1,\dots,N + N_I$, $i = 1,\dots,n_0RB$, using convolutional encoder with R, B, and 2^{n_0} being the code rate, frame length in binary symbols, and the number of constellation points in modulation scheme, respectively. The users indexed by $n = 1,\dots,N$ are the desired users and those indexed by $n = N + 1,\dots,N + N_I$ are UCCI. The encoded binary sequences $d_n^{(i)}(k)$, $k = 1,\dots,B$, $i = 1,\dots,n_0$, are interleaved and 2^{n_0}-PSK modulated, resulting in symbol sequences $b_n(k) = \mathcal{M}\{d_n^{(i)}, i = 1,\dots,n_0\} \in \mathcal{Q}$, $k = 1,\dots,B$, where $\mathcal{M}$ is bit-to-symbol mapping function. $\mathcal{Q} = \{\alpha_1,\dots,\alpha_{2^{n_0}}\}$ denotes signal constellation set of the 2^{n_0}-PSK modulation. They are preceded by the user-specific training sequences of length T symbols. The frame structure is presented in Figure 2. The entire frame is transmitted through frequency-selective channel with L paths.

After coherent demodulation in the receiver, the signals from each of the M receive antennas are matched-filtered and sampled at the symbol rate. The sampled received signal at the antenna m is given by

$$r_m(k) = \sum_{l=0}^{L-1}\sum_{n=1}^{N} h_{m,n}(l)b_n(k-l) + \sum_{l=0}^{L-1}\sum_{n=N+1}^{N+N_I} h_{m,n}(l)b_n(k-l) + v_m(k), \tag{1}$$

where $h_{m,n}(l)$ is a baseband representation of the channel gain between the nth user and mth antenna for the lth path, and $v_m(k)$ is additive white Gaussian noise (AWGN) with

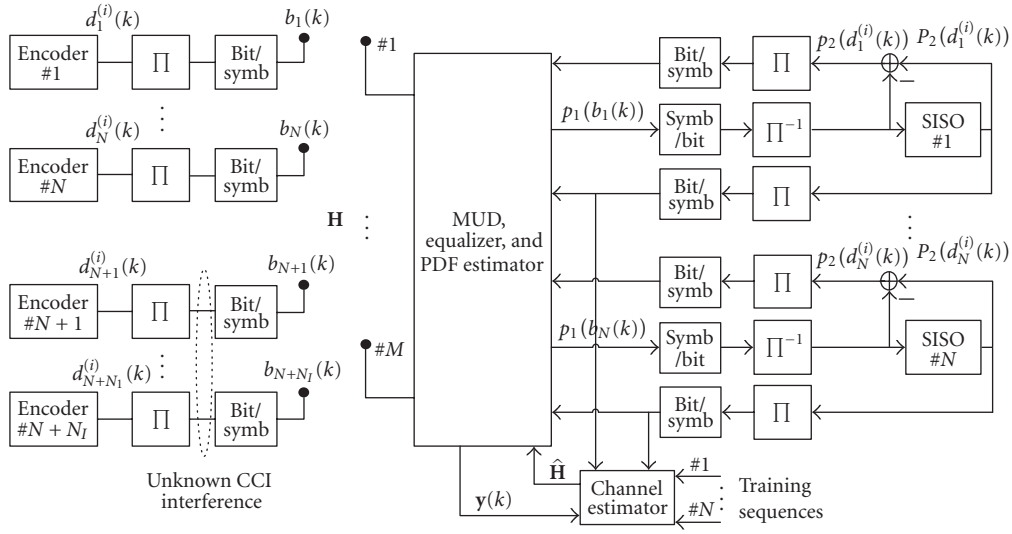

Figure 1: Transmitter and iterative receiver block scheme.

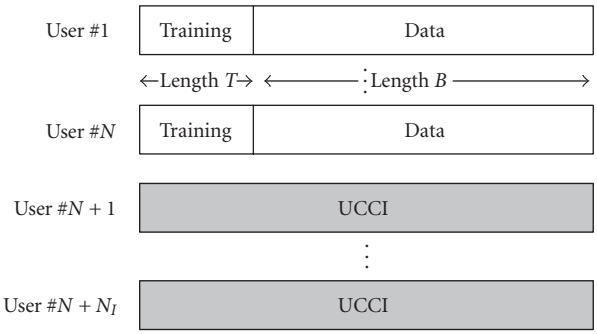

Figure 2: Frame structure.

variance σ^2. After collecting signals from all the antennas into the space vector, we obtain

$$\mathbf{r}(k) = \left[r_1(k) \cdots r_M(k)\right]^T = \sum_{l=0}^{L-1} \mathbf{H}(l)\mathbf{b}(k-l) + \sum_{l=0}^{L-1} \mathbf{H}_I(l)\mathbf{b}_I(k-l) + \mathbf{v}(k), \tag{2}$$

where

$$\mathbf{H}(l) = \begin{bmatrix} h_{1,1}(l) & \cdots & h_{1,N}(l) \\ \vdots & \ddots & \vdots \\ h_{M,1}(l) & \cdots & h_{M,N}(l) \end{bmatrix},$$
$$\mathbf{H}_I(l) = \begin{bmatrix} h_{1,N+1}(l) & \cdots & h_{1,N+N_I}(l) \\ \vdots & \ddots & \vdots \\ h_{M,1}(l) & \cdots & h_{M,N+N_I}(l) \end{bmatrix}, \tag{3}$$
$$\mathbf{b}(k) = \left[b_1(k) \cdots b_N(k)\right]^T,$$
$$\mathbf{b}_I(k) = \left[b_{N+1}(k) \cdots b_{N+N_I}(k)\right]^T,$$
$$\mathbf{v}(k) = \left[v_1(k) \cdots v_M(k)\right]^T.$$

In order to capture the multipath components, the window of received signal samples spanning the time frame of length L at time instant k is collected into the space-time vector $\mathbf{y}(k) \in \mathbb{C}^{LM\times 1}$ given by [13]

$$\begin{aligned} \mathbf{y}(k) &= \left[\mathbf{r}(k+L-1)^T \cdots \mathbf{r}(k)^T\right]^T \\ &= \mathbf{H}\mathbf{u}(k) + \mathbf{H}_I\mathbf{u}_I(k) + \mathbf{n}(k), \end{aligned} \tag{4}$$

where $\mathbf{H} \in \mathbb{C}^{LM\times N(2L-1)}$ and $\mathbf{H}_I \in \mathbb{C}^{LM\times N_I(2L-1)}$ are defined as

$$\mathbf{H} = \begin{bmatrix} \mathbf{H}(0) & \cdots & \mathbf{H}(L-1) & \cdots & \mathbf{0} \\ \vdots & \ddots & & \ddots & \vdots \\ \mathbf{0} & \cdots & \mathbf{H}(0) & \cdots & \mathbf{H}(L-1) \end{bmatrix},$$
$$\mathbf{H}_I = \begin{bmatrix} \mathbf{H}_I(0) & \cdots & \mathbf{H}_I(L-1) & \cdots & \mathbf{0} \\ \vdots & \ddots & & \ddots & \vdots \\ \mathbf{0} & \cdots & \mathbf{H}_I(0) & \cdots & \mathbf{H}_I(L-1) \end{bmatrix}, \tag{5}$$

and $\mathbf{u}(k) \in \mathbb{C}^{N(2L-1)\times 1}$, $\mathbf{u}_I(k) \in \mathbb{C}^{N_I(2L-1)\times 1}$, and $\mathbf{n}(k) \in \mathbb{C}^{LM\times 1}$ are given by

$$\mathbf{u}(k) = \left[\mathbf{b}(k+L-1)^T \cdots \mathbf{b}(k)^T \cdots \mathbf{b}(k-L+1)^T\right]^T,$$
$$\mathbf{u}_I(k) = \left[\mathbf{b}_I(k+L-1)^T \cdots \mathbf{b}_I(k)^T \cdots \mathbf{b}_I(k-L+1)^T\right]^T,$$
$$\mathbf{n}(k) = \left[\mathbf{v}(k+L-1)^T \cdots \mathbf{v}(k)^T\right]^T, \tag{6}$$

with $E\{\mathbf{n}(k)\mathbf{n}^H(k)\} = \sigma^2\mathbf{I}$.

3. CONVENTIONAL ITERATIVE RECEIVER

3.1. SC/MMSE multiuser detector

The conventional receiver for MIMO turbo equalization with UCCI is proposed in [22]. We highlight the main points of

the receiver for convenience. Assume without loss of generality that user 1 is the user of interest. Let

$$\mathbf{x}(k) = \mathbf{H}_I \mathbf{u}_I(k) + \mathbf{n}(k). \tag{7}$$

First iteration

The sample vectors $\mathbf{y}(k), k = 1,\ldots,T$, denoting training sequences, are first directed to the channel estimator to obtain the estimate $\hat{\mathbf{H}}$ of $\mathbf{H}$, and then the samples

$$\hat{\mathbf{x}}(k) = \mathbf{y}(k) - \hat{\mathbf{H}}\mathbf{u}(k), \quad k = 1,\ldots,T, \tag{8}$$

are used to estimate the covariance matrix of the UCCI plus noise using sample average given by

$$\begin{aligned}\mathbf{R}_{\mathbf{xx}} &= E\{\mathbf{x}\mathbf{x}^H\} \\ &\approx \frac{\sum_{k=1}^{T} \hat{\mathbf{x}}(k)\hat{\mathbf{x}}^H(k)}{T} = \hat{\mathbf{R}}_{\mathbf{xx}}.\end{aligned} \tag{9}$$

In order to suppress the known and unknown CCI components as well as the ISI components of the desired signal, a linear filter with weighting vector $\mathbf{w}_1(k)$ is applied to the signal $\mathbf{y}(k)$, $k = T+1,\ldots,B+T$, so as to satisfy the MMSE criterion:

$$\mathbf{w}_1(k) = \arg\min_{\mathbf{w}(k)} \left|\left|\mathbf{w}^H(k)\mathbf{y}(k) - b_1(k)\right|\right|^2, \tag{10}$$

resulting in the optimal weighting vector

$$\mathbf{w}_1(k) = \left[\hat{\mathbf{H}}\hat{\mathbf{H}}^H + \hat{\mathbf{R}}_{\mathbf{xx}}\right]^{-1}\mathbf{h}_1, \tag{11}$$

where

$$\mathbf{h}_1 = \hat{\mathbf{H}}\mathbf{e}_1, \tag{12}$$

and $\mathbf{e}_1 \in \mathbb{R}^{N(2L-1)\times 1}$ is defined as

$$\mathbf{e}_1 = \left[\mathbf{0}_{1\times(L-1)N}\, 1\, \mathbf{0}_{1\times LN-1}\right]^T. \tag{13}$$

By approximating the error at the MMSE filter output as Gaussian [13], the extrinsic probabilities to be passed to the decoder are calculated as

$$\begin{aligned}p_1(b_1(k) = \alpha_j) &= \frac{1}{\pi\nu_1^2(k)} e^{-(z_1(k)-\mu_1(k)\alpha_j)^H(z_1(k)-\mu_1(k)\alpha_j)/\nu_1^2(k)}, \\ &\quad j = 1,\ldots,2^{n_0},\end{aligned} \tag{14}$$

where $z_1(k)$ is the MMSE filter output,

$$\begin{aligned}z_1(k) &= \mathbf{w}_1^H(k)\mathbf{y}(k), \\ \mu_1(k) &= \mathbf{w}_1^H(k)\mathbf{h}_1, \\ \nu_1^2(k) &= \mathbf{w}_1^H(k)\left[\hat{\mathbf{H}}\hat{\mathbf{H}}^H + \hat{\mathbf{R}}_{\mathbf{xx}}\right]\mathbf{w}_1(k) - \mu_1(k)\mu_1^*(k).\end{aligned} \tag{15}$$

Subsequent iterations

Let $\tilde{\mathbf{u}}(k)$ denote the training sequence

$$\tilde{\mathbf{u}}(k) = \mathbf{u}(k), \quad k = 1,\ldots,T, \tag{16}$$

or the soft data sequence fed back from the channel decoder:

$$\begin{aligned}\tilde{\mathbf{u}}(k) &= \left[\tilde{\mathbf{b}}^T(k+L-1)\cdots\tilde{\mathbf{b}}^T(k)\cdots\tilde{\mathbf{b}}^T(k-L+1)\right]^T, \\ &\quad k = T+1,\ldots,T+B,\end{aligned} \tag{17}$$

where

$$\tilde{\mathbf{b}}(k) = \left[\tilde{b}_1(k)\cdots\tilde{b}_N(k)\right]^T, \tag{18}$$

$$\tilde{b}_n(k) = \sum_{\alpha_j\in\mathcal{Q}} \alpha_j p_2\{b_n(k) = \alpha_j\}. \tag{19}$$

$p_2\{b_n(k) = \alpha_j\}$ denotes the extrinsic probability obtained by SISO channel decoding (see (34)).

Similarly define

$$\overline{\mathbf{u}}(k) = \mathbf{u}(k), \quad k = 1,\ldots,T, \tag{20}$$

or

$$\begin{aligned}\overline{\mathbf{u}}(k) &= \left[\overline{\mathbf{b}}^T(k+L-1)\cdots\overline{\mathbf{b}}^T(k)\cdots\overline{\mathbf{b}}^T(k-L+1)\right]^T, \\ &\quad k = T+1,\ldots,T+B,\end{aligned} \tag{21}$$

where

$$\overline{\mathbf{b}}(k) = \left[\overline{b}_1(k)\cdots\overline{b}_N(k)\right]^T, \tag{22}$$

$$\overline{b}_n(k) = \sum_{\alpha_j\in\mathcal{Q}} \alpha_j P_2\{b_n(k) = \alpha_j\}. \tag{23}$$

$P_2\{b_n(k) = \alpha_j\}$ denotes the *a posteriori* probability obtained by the SISO decoding (see (35)).

The samples $\overline{\mathbf{u}}(k)$, $k = 1,\ldots,T+B$, are first fed to the channel estimator to reestimate the channel $\mathbf{H}$, and then the samples

$$\hat{\mathbf{x}}(k) = \mathbf{y}(k) - \hat{\mathbf{H}}\overline{\mathbf{u}}(k), \quad k = 1,\ldots,T+B, \tag{24}$$

are used to update the estimate of the covariance matrix $\mathbf{R}_{\mathbf{xx}}$:

$$\hat{\mathbf{R}}_{\mathbf{xx}} = \frac{\sum_{k=1}^{T+B} \hat{\mathbf{x}}(k)\hat{\mathbf{x}}^H(k)}{T+B}. \tag{25}$$

We now denote

$$\hat{\mathbf{u}}_1(k) = \tilde{\mathbf{u}}(k) - \tilde{b}_1(k)\mathbf{e}_1. \tag{26}$$

Soft cancellation of the known CCI components as well as the ISI components of the desired signal is performed, yielding

$$\hat{\mathbf{y}}_1(k) = \mathbf{y}(k) - \hat{\mathbf{H}}\hat{\mathbf{u}}_1(k), \quad k = T+1,\ldots,B+T. \tag{27}$$

After that, a linear filter with weighting vector $\mathbf{w}_1(k)$ is applied to the signal $\hat{\mathbf{y}}_1(k)$ so as to satisfy the MMSE criterion:

$$\mathbf{w}_1(k) = \arg\min_{\mathbf{w}(k)} ||\mathbf{w}^H(k)\hat{\mathbf{y}}_1(k) - b_1(k)||^2, \tag{28}$$

resulting in the optimal weighting vector

$$\mathbf{w}_1(k) = \left[\hat{\mathbf{H}}\boldsymbol{\Lambda}_1(k)\hat{\mathbf{H}}^H + \hat{\mathbf{R}}_{\mathbf{xx}}\right]^{-1}\mathbf{h}_1, \tag{29}$$

where

$$\boldsymbol{\Lambda}_1(k) = \mathbf{I} - E\{\hat{\mathbf{u}}_1(k)\hat{\mathbf{u}}_1^H(k)\}. \tag{30}$$

Note that (30) holds only for multilevel PSK. Note further that the total number of DOF of the iterative linear SC/MMSE receiver after convergence is determined by the product LM. This number is decreased by a factor equal to the rank of the matrix $\hat{\mathbf{R}}_{\mathbf{xx}}$ while cancelling UCCI.

The extrinsic probabilities to be passed to the decoder are calculated as in (14), where the MMSE filter output $z_1(k)$ is now defined as

$$\begin{aligned} z_1(k) &= \mathbf{w}_1^H(k)\hat{\mathbf{y}}_1(k), \\ \mu_1(k) &= \mathbf{w}_1^H(k)\mathbf{h}_1, \\ \nu_1^2(k) &= \mathbf{w}_1^H(k)\left[\hat{\mathbf{H}}\boldsymbol{\Lambda}_1(k)\hat{\mathbf{H}}^H + \hat{\mathbf{R}}_{\mathbf{xx}}\right]\mathbf{w}_1(k) - \mu_1(k)\mu_1^*(k). \end{aligned} \tag{31}$$

3.2. Channel decoder

Each of the single-user channel decoders produces the maximum *a posteriori* (MAP) probability of each binary symbol $d_1^{(i)}(k)$:

$$\begin{aligned} &P_2\left(d_1^{(i)}(k) = \pm 1\right) \\ &\quad = p\left(d_1^{(i)}(k) = \pm 1 | z_1(k),\ k = T+1,\ldots,T+B\right) \\ &\quad = p_2\left(d_1^{(i)}(k) = \pm 1\right) p_1^{(\pi^{-1})}\left(d_1^{(i)}(k) = \pm 1\right), \end{aligned} \tag{32}$$

where $p_1^{(\pi^{-1})}(d_1^{(i)}(k) = \pm 1)$ is the deinterleaved *a priori* information $p_1(d_1^{(i)}(k) = \pm 1)$ obtained from the MMSE detection stage and $p_2(d_1^{(i)}(k) = \pm 1)$ is the decoder extrinsic probability. To obtain $p_1(d_1^{(i)}(k) = \pm 1)$, a symbol-to-bit probability conversion has to be made as follows:

$$p_1\left(d_1^{(i)}(k) = +1\right) = \sum_{\beta \in \mathcal{B}^{+1}} p_1\left(b_1(k) = \beta\right), \tag{33}$$

where $\mathcal{B}^{+1} = \{\beta \in \mathcal{Q} | \beta = \mathcal{M}\{\delta_p, p = 1,\ldots,n_0; \delta_p \in \{+1, -1\}, p \neq i, \delta_i = +1\}\}$, and similarly for $p_1(d_1^{(i)}(k) = -1)$.

The extrinsic probabilities $p_2(d_1(k) = \pm 1)$ are used to make the conversions from bit-to-symbol extrinsic probabilities, yielding

$$\begin{aligned} p_2\left(b_1(k) = \alpha_j\right) &= \prod_{i=1}^{n_0} p_2\left(d_1^{(i)}(k) = \delta_i\right), \\ \alpha_j &= \mathcal{M}\{\delta_i \in \{+1, -1\},\ i = 1,\ldots,n_0\}, \end{aligned} \tag{34}$$

which are delivered to the SC/MMSE receiver through (19). Similarly, the symbol-level *a posteriori* probabilities are calculated as

$$\begin{aligned} P_2\left(b_1(k) = \alpha_j\right) &= \prod_{i=1}^{n_0} P_2\left(d_1^{(i)}(k) = \delta_i\right), \\ \alpha_j &= \mathcal{M}\{\delta_i \in \{+1, -1\},\ i = 1,\ldots,n_0\}, \end{aligned} \tag{35}$$

to be used in (23). After a sufficient number of iterations, when the receiver has converged, the decision on the transmitted binary information symbols $c_n(i)$ is made based on the *a posteriori* log-likelihood ratios for $c_n(i)$, given as

$$\lambda\left(c_n(i)\right) = \frac{p\left(c_n(i) = +1 | z_1(k),\ k = T+1,\ldots,T+B\right)}{p\left(c_n(i) = -1 | z_1(k),\ k = T+1,\ldots,T+B\right)}, \tag{36}$$

for decision making.

Iterative channel estimation from [24] is applied. The detailed description is reviewed in Appendix B.

4. ITERATIVE RECEIVER WITH PDF ESTIMATION

4.1. Receiver derivation

Unlike the conventional receiver described in Section 3, which uses MMSE detection after soft cancellation, the proposed receiver uses maximum-likelihood (ML) processing, by making use of the actual structure of the UCCI.

First iteration

We rewrite (4) as

$$\begin{aligned} \mathbf{y}(k) &= \underbrace{\mathbf{h}_1 b_1(k)}_{\text{desired}} + \underbrace{\mathbf{H}_{\text{CISI},1}\mathbf{u}_{\text{CISI},1}(k)}_{\text{CCI+ISI}} + \underbrace{\mathbf{H}_I\mathbf{u}_I(k)}_{\text{UCCI}} + \underbrace{\mathbf{n}(k)}_{\text{noise}} \\ &= \mathbf{h}_1 b_1(k) + \mathbf{x}_1(k), \end{aligned} \tag{37}$$

where $\mathbf{H}_{\text{CISI},1} = \mathbf{H} - [\mathbf{0}_{LM\times(L-1)N}\mathbf{h}_1\mathbf{0}_{LM\times LN-1}]$, $\mathbf{u}_{\text{CISI},1}(k) = \mathbf{u}(k) - b_1(k)\mathbf{e}_1$, and $\mathbf{x}_1(k)$ denotes the total sum of the desired user's intersymbol interference (ISI), known CCI, UCCI, and noise. Since in the first iteration the soft feedback is not available, the ISI and CCI components cannot be cancelled. ML processing requires the PDF of the signal $\mathbf{x}_1(k)$, which is multimodal Gaussian, given by

$$p_{\mathbf{x}_1}\left(\mathbf{x}_1(k)\right) = \frac{1}{2^{D_{\text{tot}}}} \sum_{i=1}^{2^{D_{\text{tot}}}} \frac{1}{\left(\pi\sigma^2\right)^{LM}} e^{-\left(\mathbf{x}_1(k)-\mathbf{t}_{i,1}\right)^H\left(\mathbf{x}_1(k)-\mathbf{t}_{i,1}\right)/\sigma^2}, \tag{38}$$

where $D_{\text{tot}} = n_0[(2L-1)(N+N_I) - N]$, and $\mathbf{t}_{i,1}$ depends on the entries of $\mathbf{H}_I$ and $\mathbf{H}_{\text{CISI},1}$ and the signal constellation of the UCCI. The number of summation terms in (38) increases exponentially with the number of users N which may be large in a practical system. In that case, the samples $\mathbf{x}_1(k)$ will become less structured and their PDF will become more Gaussian-like. To justify the Gaussian approximation, we calculate the Kullback-Leibler (KL) distance (relative entropy)

Table 1: KL distance between the true PDF of (38) and Gaussian approximation of (39), $L = 2$.

$N + N_I$	1	3	5
KL distance	60	25	22.5

[25] between the true PDF given by (38) and the corresponding Gaussian approximation given by

$$p_{\text{Gapp},\mathbf{x}_1}(\mathbf{x}_1(k)) = \frac{1}{\pi^{LM}\det(\mathbf{R}_{\mathbf{x}_1\mathbf{x}_1})} e^{-\mathbf{x}_1(k)^H \mathbf{R}_{\mathbf{x}_1\mathbf{x}_1}^{-1}\mathbf{x}_1(k)}, \quad (39)$$

with

$$\mathbf{R}_{\mathbf{x}_1\mathbf{x}_1} = E\{\mathbf{x}_1(k)\mathbf{x}_1(k)^H\} = \mathbf{H}_{\text{CISI},1}\mathbf{H}_{\text{CISI},1}^H + \mathbf{R}_{\mathbf{xx}}, \quad (40)$$

for several values of $N + N_I$. It can be seen from Table 1 that the KL distance decreases as the $N + N_I$ increases, which means that the true PDF approaches Gaussian. Therefore, by adopting the Gaussian assumption for the $\mathbf{x}_1(k)$, the extrinsic probability to be passed to the first user's SISO decoder can be calculated as

$$p(b_1(k) = \alpha_j) = C_{\text{ML}} e^{-(\mathbf{y}(k)-\mathbf{h}_1\alpha_j)^H \mathbf{R}_{\mathbf{x}_1\mathbf{x}_1}^{-1}(\mathbf{y}(k)-\mathbf{h}_1\alpha_j)}, \quad j = 1,\ldots,2^{n_0}, \quad (41)$$

where $C_{\text{ML}} = 1/\pi^{LM}\det(\mathbf{R}_{\mathbf{x}_1\mathbf{x}_1})$. It can also be shown [26] that the MMSE filter given by (11) can be transformed using the matrix inversion lemma, resulting in

$$\mathbf{w}_1(k) = \frac{\mathbf{R}_{\mathbf{x}_1\mathbf{x}_1}^{-1}\mathbf{h}_1}{1 + \mathbf{h}_1^H \mathbf{R}_{\mathbf{x}_1\mathbf{x}_1}^{-1}\mathbf{h}_1}. \quad (42)$$

Note that the estimates of $\mathbf{H}$ and $\mathbf{R}_{\mathbf{xx}}$ are replaced by their true values in (11) to show that (42) holds. In practice, the estimate of $\mathbf{R}_{\mathbf{x}_1\mathbf{x}_1}$ is obtained from (40) by using estimates $\widehat{\mathbf{H}}$ and $\widehat{\mathbf{R}}_{\mathbf{xx}}$, defined in Section 3. After incorporating (42) into (15), the extrinsic probability at the output of MMSE filter, given by (14), can be represented as

$$p(b_1(k) = \alpha_j) = C_{\text{MMSE}} e^{-(\mathbf{y}(k)-\mathbf{h}_1\alpha_j)^H \mathbf{R}_{\mathbf{x}_1\mathbf{x}_1}^{-1}(\mathbf{y}(k)-\mathbf{h}_1\alpha_j)}, \quad j = 1,\ldots,2^{n_0}, \quad (43)$$

where $C_{\text{MMSE}} = (1 + \mathbf{h}_1^H \mathbf{R}_{\mathbf{x}_1\mathbf{x}_1}^{-1}\mathbf{h}_1)^2/\pi\mathbf{h}_1^H \mathbf{R}_{\mathbf{x}_1\mathbf{x}_1}^{-1}\mathbf{h}_1$. This, however, is just the scaled extrinsic information of (41) obtained by using the ML detector. Since the constants C_{ML} and C_{MMSE} do not have any impact on the receiver performance, in the first iteration the proposed ML receiver is exactly the same as the conventional MMSE receiver presented in Section 3.

Subsequent iterations

Starting from the second iteration, we make use of the soft feedback. Assuming that the soft cancellation in (24) is almost perfect, the ISI components of the desired user and the known CCI components can be cancelled, and the PDF of the signal $\hat{\mathbf{x}}(k)$, given in (24), can be given as

$$p_{\hat{\mathbf{x}}}(\hat{\mathbf{x}}(k)) \approx \frac{1}{2^D}\sum_{i=0}^{2^D-1} \frac{1}{(\pi\sigma^2)^{LM}} e^{-(\hat{\mathbf{x}}(k)-\mathbf{t}_i)^H(\hat{\mathbf{x}}(k)-\mathbf{t}_i)/\sigma^2}, \quad (44)$$

where $D = n_0(2L-1)N_I$, and $\mathbf{t}_i$ depends on the matrices $\mathbf{H}_I$ and the signal constellation of the UCCI. Assuming that the number N_I of UCCIs is relatively small, the structure of the UCCI can be exploited by estimating the PDF of UCCI plus noise given by (44) and applying ML filtering. After the estimate $\widehat{p}_{\hat{\mathbf{x}}}(\hat{\mathbf{x}}(k))$ of $p_{\hat{\mathbf{x}}}(\hat{\mathbf{x}}(k))$ is obtained, the extrinsic probability to be passed to the first user's SISO decoder can be calculated as the output of the single-user ML detector as

$$p_1(b_1(k) = \alpha_j) = \widehat{p}_{\hat{\mathbf{x}}}(\hat{\mathbf{y}}_1(k) - \alpha_j\mathbf{h}_1), \quad k = T+1,\ldots,T+B, \quad j = 1,\ldots,2^{n_0}. \quad (45)$$

The PDF estimation procedure is described in the sequel. First, the channel is reestimated based on $\overline{\mathbf{u}}_1(k)$, $k = 1,\ldots,B+T$, as in Section 3. Then, the samples $\hat{\mathbf{x}}(k)$, $k = 1,\ldots,T+B$, are used to make the estimate of the UCCI-plus-noise PDF. Note that by using the samples indexed by $k = 1,\ldots,B+T$, we perform iterative PDF estimation. In the noniterative PDF estimation, only first T samples, $\hat{\mathbf{x}}(k)$, $k = 1,\ldots,T$, corresponding to the training sequence, would be used. In order to perform the PDF estimation, either parametric [19] or nonparametric [23] approach can be used. The former one estimates the parameters D and $\mathbf{t}_i$ based on the samples $\hat{\mathbf{x}}(k)$. These estimates are then used in (44). On the other hand, the nonparametric approach estimates PDF *directly*, where each sample $\hat{\mathbf{x}}(k)$ contributes to the total estimate through a weighting function. For example, for an arbitrary $\mathbf{a} = [a_1,\ldots,a_{LM}]^T \in \mathbb{C}^{LM\times 1}$, the nonparametric multidimensional kernel-based PDF estimator [23] estimates the $p_{\hat{\mathbf{x}}}(\mathbf{a})$ as

$$\widehat{p}_{\hat{\mathbf{x}}}(\mathbf{a}) = \frac{1}{T+B}\sum_{k=1}^{T+B} \frac{K_1((\hat{\mathbf{x}}(k)-\mathbf{a})/\sigma_0)}{\sigma_0^{2LM}}, \quad (46)$$

where $K_1(\mathbf{a}) = 1/(2\pi)^{LM} e^{-\mathbf{a}^H\mathbf{a}/2}$ is a Gaussian kernel weighting function and σ_0 is a smoothing parameter. Although other kernel functions can be used [23], it will be shown that this choice gives an asymptotically unbiased and consistent PDF estimator. The estimation accuracy is controlled by the smoothing parameter σ_0. The larger value of σ_0 results in the smoother but less accurate PDF estimate, and vice versa. In order to find the optimal value for σ_0, one approach is to minimize the mean integrated square error (MISE) [23] between the true PDF and its estimate, as defined by

$$\text{MISE}(\widehat{p}_{\hat{\mathbf{x}}}) = \int_{\mathbb{R}^{2LM}} \{p_{\hat{\mathbf{x}}}(\mathbf{a}) - \widehat{p}_{\hat{\mathbf{x}}}(\mathbf{a})\}^2 d\mathbf{a}, \quad (47)$$

where $d\mathbf{a} = d\Re a_1 d\Im a_1 \cdots d\Re a_{LM} d\Im a_{LM}$. It is shown in Appendix A that the optimal smoothing parameter $\sigma_{0,\text{opt}}$ can

be lower bounded as follows:

$$\sigma_{0,\mathrm{opt}} \geq \left(\frac{2}{(T+B)(LM+1)}\right)^{1/(2LM+4)}, \quad \sigma = \gamma(LM). \quad (48)$$

A similar result was obtained in [20] for the univariate case. It is a special case of (48) for $LM = 1$. Furthermore, the estimate

$$\hat{\sigma}_0 = k_0 \gamma(LM) \quad (49)$$

of $\sigma_{0,\mathrm{opt}}$ satisfies the sufficient conditions for consistency and asymptotic unbiasedness. These conditions are given as $\lim_{(T+B)\to\infty} \hat{\sigma}_0 = 0$ and $\lim_{(T+B)\to\infty}(T+B)\hat{\sigma}_0 = \infty$, and they are satisfied if the parameter $k_0 \in \mathbb{R}$ is chosen to be $k_0 \geq 1$ [20]. Thereby, the estimator dependence on D and $\mathbf{t}_i$ reflects only through the constant k_0 since $\gamma(LM)$ is independent of these parameters. The bit error rate (BER) performance versus parameter k_0 with different numbers of users and different numbers of multipaths as parameters is shown in Figure 3. Interestingly, the optimal value of k_0 that minimizes BER is shown to be rather *insensitive* to the change of the other parameters. Moreover, it is shown in [20] that for $LM = 1$, the optimal parameter k_0 also does not depend on the signal-to-noise ratio [21]. From Figure 3, it can be seen that $k_0 \approx 2$ is a good choice for a wide range of situations. This indicates that, in practice, the knowledge about the parameters N_I, L, and, correspondingly, D is *not needed*. If these parameters are known to the receiver, they could be used to access a lookup table in which the optimal values of k_0 for different combinations of parameters can be stored *a priori*. The same procedure is performed for the rest of desired users to obtain the soft estimates $\widetilde{b}_n(k)$ and $\overline{b}_n(k)$ for the next iteration.

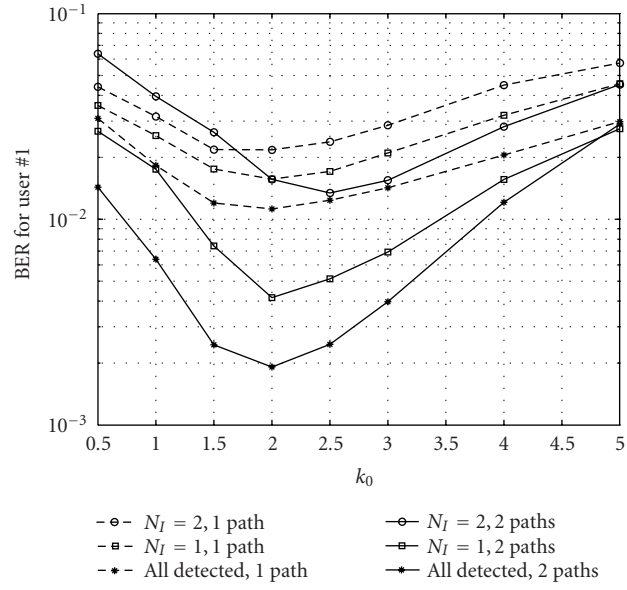

FIGURE 3: BER versus k_0 performance, noniterative PDF estimation ($N + N_I = 3$, $M = 3$, $T = 100, B = 900$, LS channel estimation, 4 iterations, $E_b/N_0 = 2$ dB).

4.2. Symmetrizing

If the UCCI signal constellation is known to the receiver, the symmetry of the constellation set can be utilized to increase the number of samples that can be used for PDF estimation. In case of 2^{n_0}-PSK modulation, a 2^{n_0}-fold increase of the number of samples can be achieved by using the fact that $p(\mathbf{a}) = p(\mathbf{a}e^{-j2\pi k/2^{n_0}})$, $k = 1, \ldots, 2^{n_0} - 1$.

4.3. Computational complexity

Since (45) contains the sums of exponentials, it can be approximated using the Jacobian algorithm [27]. The complexity per symbol of the proposed method is roughly $O\{(T + B)LM\}$ or $O\{TLM\}$, depending on whether we use soft feedback for PDF estimation or not, respectively. The conventional SC/MMSE receiver's complexity is $O\{L^3M^3\}$.

5. NUMERICAL EXAMPLES

The performance of the proposed receiver was tested through simulations. The training sequence lengths of $T = 100$, 20, and 10, data sequence length of $B = 900$, and BPSK modulation were used. The channel gains for each path of each user were assumed to have equal average powers, with Rayleigh-distributed amplitudes. They are constant over each transmitted frame, and they change independently from a frame to another frame. The rate $R = 1/2$ convolutional code with the generator polynomials $(5, 7)_8$ and the MAP decoder [10] were used for all MIMO users. User-specific random interleavers were assumed. A lower-complexity least-square (LS) channel estimation (see [24]) was used, since it is shown in [28] that the more complex MMSE channel estimation (see Appendix B) does not offer significant performance benefits unless the power ratio between UCCI and desired signals is very strong.

In Figures 4 and 5, BER versus per-antenna E_b/N_0 is presented for $L = 1$ and $L = 2$ cases, respectively. The noniterative PDF estimation is used in these examples, since long overhead ($T = 100$) was used. In both cases, the proposed receiver significantly outperforms the conventional one in the case where one or two out of three users are UCCI. This is the consequence of the linear processing of the conventional receiver of [22] that does not take into account the actual structure of the UCCI plus noise. Performance curve when all the users are to be detected is shown for comparison (indicated by "all known").

The performance is closer to the "all known" case for $L = 1$ (frequency flat fading) than for $L = 2$, and for $N_I = 1$ than $N_I = 2$. This is because the PDF of (44) becomes more scattered in the LM-dimensional space with increased L and N_I. It means that fewer samples $\hat{\mathbf{x}}$ (out of T available) effectively contribute to the estimate $\hat{p}_{\hat{\mathbf{x}}}(\mathbf{a})$ of $p_{\hat{\mathbf{x}}}(\mathbf{a})$ in (46), which decreases the PDF estimation accuracy. The increased M with fixed T also reduces the estimation accuracy due to the increased dimensionality of $\hat{\mathbf{x}}$ [23]. Its impact can, however, be compensated for in part by (48) with an appropriate choice of optimal k_0.

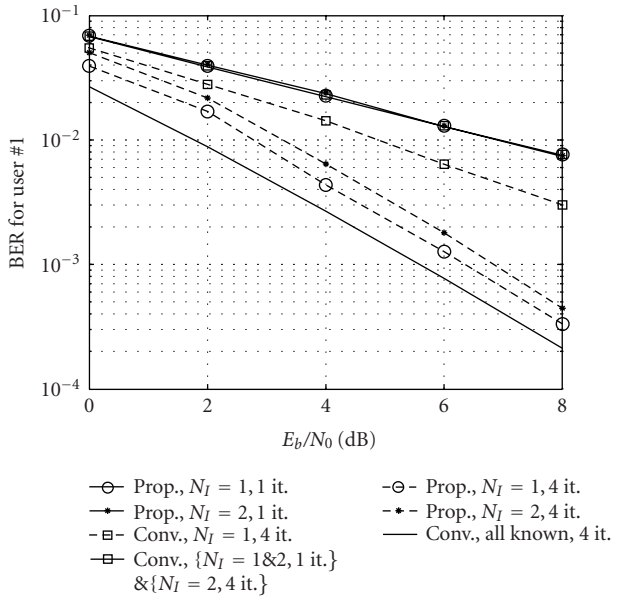

FIGURE 4: BER versus E_b/N_0 performance, frequency flat fading, noniterative PDF estimation ($N+N_I = 3, M = 3, T = 100, B = 900, k_0 = 2$).

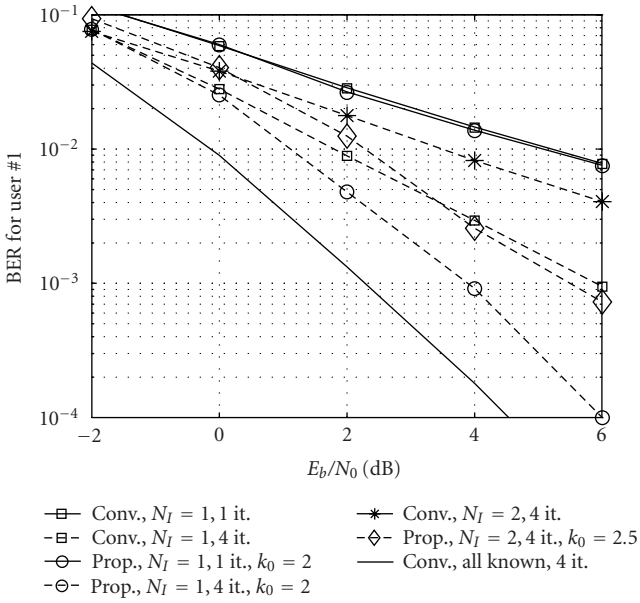

FIGURE 5: BER versus E_b/N_0 performance, 2-path fading, noniterative PDF estimation ($N + N_I = 3, M = 3, T = 100, B = 900$).

In Figure 6, BER performance of iterative and noniterative PDF estimation is presented. The abbreviations FB, no FB, conv., and prop. stand for the iterative PDF estimation (feedback), noniterative PDF estimation (no feedback), and conventional and proposed receivers, respectively. It can be found from Figure 6 that the iterative PDF estimation-based receiver with a short ($T = 10$ and 20) training sequence can

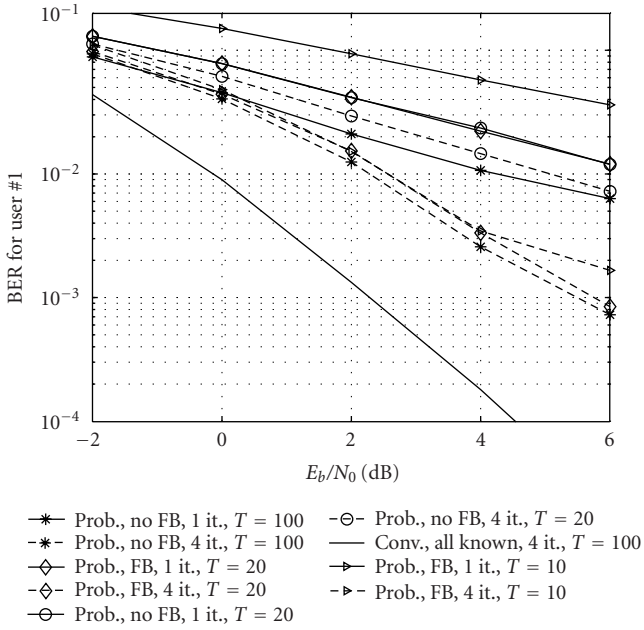

FIGURE 6: BER versus E_b/N_0 performance, 2-path fading, iterative versus noniterative PDF estimation ($N = 1, N_I = 2, M = 3, B = 900$, LS channel estimation).

achieve almost the same performance as the noniterative receiver with long ($T = 100$) training sequence. It should be emphasized that the reduction in overhead due to training when using iterative PDF estimation is rather significant.

6. CONCLUSIONS

A kernel smoothing PDF estimation-based receiver was derived to preserve the diversity order of iterative SC/MMSE receivers for multiuser detection in frequency-selective channels in the presence of unknown cochannel interference. The PDF estimation can be based on training symbols only (noniterative PDF estimation) or on training symbols as well as feedback from the decoder (iterative PDF estimation). It was verified through simulations that the proposed receiver significantly outperforms the conventional covariance estimation in channels with low frequency-selectivity, where the degradation is more severe due to the lack of multipath diversity. In higher frequency-selectivity channels, the PDF estimation accuracy will decrease, since the UCCI-plus-noise components will be more scattered in the multidimensional data space. Fortunately, the need for diversity is less stringent therein. The proposed receiver with iterative PDF estimation can significantly outperform both the conventional and noniterative PDF estimation-based receiver with minor training overhead. Moreover, its performance has been shown to be very close to that of noniterative PDF estimation with a long overhead. Thus, the proposed receiver provides significant potential both for bandwidth-efficiency improvement and for system capacity increase in multiuser communications in flat and moderately frequency-selective channels. Potential application areas may be in cellular systems, where there

are usually only a few dominant other-cell interferers or high data-rate users, which can be suppressed by the method presented here. The receiver may also serve as a basis for a random access scheme where short bursts are transmitted in an asynchronous mode. Assuming that the collisions are not too numerous, they could be handled by the proposed method.

APPENDICES

A. DERIVATION OF LOWER BOUND ON $\sigma_{0,\text{opt}}$

A reasonable approximation of (47) can be done by using its Taylor series expansion [23], with which

$$\text{MISE}(\hat{p}_{\hat{\mathbf{x}}}) \approx \frac{\sigma_0^4 \alpha^2 \Gamma(p_{\hat{\mathbf{x}}})}{4} + \frac{\int_{\mathbb{R}^{2LM}} K_1^2(\mathbf{a})\, d\mathbf{a}}{T\sigma_0^{2LM}}, \tag{A.1}$$

where $\alpha = \int_{\mathbb{R}^{2LM}} (\Re a_1)^2 K_1(\mathbf{a}) d\mathbf{a} = 1$ and

$$\Gamma(p_{\hat{\mathbf{x}}}) = \int_{\mathbb{R}^{2LM}} \left[\sum_{i=1}^{LM} \left(\frac{\partial^2 p_{\hat{\mathbf{x}}}(\mathbf{a})}{\partial(\Re a_i)^2} + \frac{\partial^2 p_{\hat{\mathbf{x}}}(\mathbf{a})}{\partial(\Im a_i)^2} \right) \right]^2 d\mathbf{a}. \tag{A.2}$$

From (A.1), the optimal smoothing parameter $\sigma_{0,\text{opt}}$ is found to be

$$\sigma_{0,\text{opt}} = \left(\frac{2LM \int_{\mathbb{R}^{2LM}} K_1^2(\mathbf{a})\, d\mathbf{a}}{(T+B)\Gamma(p_{\hat{\mathbf{x}}})} \right)^{1/(2LM+4)}, \tag{A.3}$$

with

$$\int_{\mathbb{R}^{2LM}} K_1^2(\mathbf{a})\, d\mathbf{a} = \frac{1}{(2\pi)^{LM}} \int_{\mathbb{R}^{2LM}} e^{-\mathbf{a}^H\mathbf{a}/2}\, d\mathbf{a} = \frac{1}{(4\pi)^{LM}}. \tag{A.4}$$

In general, the function $\Gamma(p_{\hat{\mathbf{x}}})$ is dependent on D and $\mathbf{t}_i$, $i = 1, \ldots, D$. However, it is shown in [20] for the univariate case that the upper bound on $\Gamma(p_{\hat{\mathbf{x}}})$ obtained using Cauchy's inequality is dependent neither on $\mathbf{t}_i$ nor on D. Adopting the same approach in the sequel, we generalize the upper bound derivation to the multivariate case. We denote

$$p_G(\mathbf{a}) = \frac{1}{(\sigma^2 2\pi)^{LM}} e^{-\mathbf{a}^H\mathbf{a}/2\sigma^2}. \tag{A.5}$$

Equation (44) denoting the exact PDF of the UCCI plus noise can be rewritten as

$$p_{\hat{\mathbf{x}}}(\mathbf{x}(k)) = \frac{1}{2^D} \sum_{i=0}^{2^D-1} p_G(\mathbf{x}(k) - \mathbf{t}_i). \tag{A.6}$$

With (A.5) the expression for $\Gamma(p_{\hat{\mathbf{x}}})$ can be rewritten as

$$\Gamma(p_{\hat{\mathbf{x}}}) = \int_{\mathbb{R}^{2LM}} \left[\frac{1}{2^D} \sum_{k=0}^{2^D-1} \Upsilon_k(p_G) \right]^2 d\mathbf{a}, \tag{A.7}$$

where

$$\Upsilon_k(p_G) = \sum_{i=1}^{LM} \left(\frac{\partial^2 p_G(\mathbf{a}-\mathbf{t}_k)}{\partial(\Re a_i)^2} + \frac{\partial^2 p_G(\mathbf{a}-\mathbf{t}_k)}{\partial(\Im a_i)^2} \right). \tag{A.8}$$

Applying Cauchy's inequality

$$\left[\sum \nu_1 \nu_2 \right]^2 \le \sum \nu_1^2 \sum \nu_2^2 \tag{A.9}$$

to (A.7) with $\nu_1 = 1$ and $\nu_2 = \Upsilon_k(p_G)$, we obtain

$$\Gamma(p_{\hat{\mathbf{x}}}) \le \frac{1}{2^D} \sum_{k=0}^{2^D-1} \Gamma_k(p_G), \tag{A.10}$$

where

$$\Gamma_k(p_G) = \int_{\mathbb{R}^{2LM}} [\Upsilon_k(p_G)]^2\, d\mathbf{a}. \tag{A.11}$$

It can be shown that

$$\Gamma_k(p_G) = \frac{1}{\sigma^4} \int_{\mathbb{R}^{2LM}} \left[\sum_{i=1}^{LM} \left[\frac{(\Re a_i)^2}{\sigma^2} + \frac{(\Im a_i)^2}{\sigma^2} \right] - 2LM \right]^2 p_G^2(\mathbf{a})\, d\mathbf{a}, \tag{A.12}$$

which is *independent* on k. Furthermore, it can be shown that

$$\Gamma(p_G) = I_1 + I_2 + I_3 + I_4, \tag{A.13}$$

where

$$\begin{aligned}
I_1 &= \frac{1}{\sigma^4} \int_{-\infty}^{\infty} \sum_{i=1}^{LM} \left[\frac{(\Re a_i)^4}{\sigma^2} + \frac{(\Im a_i)^4}{\sigma^2} \right] p_G^2(\mathbf{a})\, d\mathbf{a} \\
&= \frac{3LM}{4\sqrt{\pi}\sigma^5} \left(\frac{1}{2\sigma\sqrt{\pi}} \right)^{2LM-1}, \\
I_2 &= \frac{4(LM)^2}{\sigma^4} \int_{-\infty}^{\infty} p_G^2(\mathbf{a})\, d\mathbf{a} \\
&= \frac{4(LM)^2}{\sigma^4} \left(\frac{1}{2\sigma\sqrt{\pi}} \right)^{2LM}, \\
I_3 &= -\frac{2LM}{\sigma^4} \int_{-\infty}^{\infty} \sum_{i=1}^{LM} \left[\frac{(\Re a_i)^2}{\sigma^2} + \frac{(\Im a_i)^2}{\sigma^2} \right] p_G^2(\mathbf{a})\, d\mathbf{a} \\
&= -\frac{(LM)^2}{\sigma^5\sqrt{\pi}} \left(\frac{1}{2\sigma\sqrt{\pi}} \right)^{2LM-1}, \\
I_4 &= \frac{1}{\sigma^4} \int_{-\infty}^{\infty} \sum_{q_i, q_j \in \mathbb{A}} \frac{q_i^2 q_j^2}{\sigma^2} p_G^2(\mathbf{a})\, d\mathbf{a} \\
&= \frac{2LM(2LM-1)}{16\sigma^6\sqrt{\pi}} \left(\frac{1}{2\sigma\sqrt{\pi}} \right)^{2LM-2}, \\
\mathbb{A} &= \{\Re a_i, \Re a_j, \Im a_i, \Im a_j | i \ne j; i,\ j = 1, \ldots, LM\}.
\end{aligned} \tag{A.14}$$

From (A.13) and (A.14), it follows that

$$\Gamma_k(p_G) = \frac{LM(LM+1)}{(4\pi)^{LM}\sigma^{2LM+4}}. \tag{A.15}$$

Finally, by replacing (A.15) in (A.10) and (A.10) in (A.3), the lower bound of (48) directly follows.

B. ITERATIVE CHANNEL ESTIMATION

For the purpose of channel estimation, we will introduce a different system model notation than in the main body of the paper, following [24] for convenience of notation. Starting from (1), we collect the received signal samples at the mth receive antenna into the vector $\mathbf{q}_m \in \mathbb{C}^{(T+\Delta+L-1)\times 1}$ given by

$$\begin{aligned}\mathbf{q}_m &= \left[r_m(1),\ldots,r_m(T+\Delta+L-1)\right]^T \\ &= \overline{\mathbf{B}}\mathbf{g}_m + \overline{\mathbf{B}}_I\mathbf{g}_{m,I} + \nu_m, \\ \overline{\mathbf{B}} &= \left[\overline{\mathbf{B}}_1,\ldots,\overline{\mathbf{B}}_N\right], \\ \overline{\mathbf{B}}_I &= \left[\overline{\mathbf{B}}_{N+1},\ldots,\overline{\mathbf{B}}_{N+N_I}\right], \\ \mathbf{g}_m &= \left[\mathbf{g}_{m,1}^T,\ldots,\mathbf{g}_{m,N}^T\right]^T, \\ \mathbf{g}_{m,I} &= \left[\mathbf{g}_{m,N+1}^T,\ldots,\mathbf{g}_{m,N+N_I}^T\right]^T.\end{aligned} \tag{B.16}$$

In the first iteration $\Delta = 0$ (no soft feedback is available), in the subsequent iterations $\Delta = B$, and

$$\overline{\mathbf{B}}_n = \begin{bmatrix} \mathbf{s}_n & 0 & \cdots & 0 \\ 0 & \mathbf{s}_n & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \mathbf{s}_n \end{bmatrix} \in \mathbb{C}^{(T+\Delta+L-1)\times L},$$

$$\mathbf{s}_n = \left[b_n(1),\ldots,b_n(T),\overline{b}_n(T+1),\ldots,\overline{b}_n(T+\Delta)\right]^T \in \mathbb{C}^{(T+\Delta)\times 1}, \tag{B.17}$$

and $\mathbf{g}_{m,n} \in \mathbb{C}^{L\times 1}$ and $\nu_m \in \mathbb{C}^{(T+\Delta+L-1)\times 1}$ are defined as

$$\begin{aligned}\mathbf{g}_{m,n} &= \left[h_{m,n}(0),\ldots,h_{m,n}(L-1)\right]^T, \\ \nu_m &= \left[\nu_m(1),\ldots,\nu_m(T+\Delta+L-1)\right]^T.\end{aligned} \tag{B.18}$$

The channel estimate for the mth receive antenna is obtained using the least-squares (LS) criterion, expressed by

$$\hat{\mathbf{g}}_m = \arg\min_{\mathbf{g}_m} \left|\left|\mathbf{q}_m - \overline{\mathbf{B}}\mathbf{g}_m\right|\right|^2, \tag{B.19}$$

resulting in [29]

$$\hat{\mathbf{g}}_m = \left(\overline{\mathbf{B}}^H\overline{\mathbf{B}}\right)^{-1}\overline{\mathbf{B}}^H\mathbf{q}_m. \tag{B.20}$$

If the knowledge about the second-order statistics of the UCCI and noise $\overline{\mathbf{B}}_I^H\overline{\mathbf{B}}_I + \sigma^2\mathbf{I}$ is available, the MMSE channel estimation would result in the following estimate [29]:

$$\hat{\mathbf{g}}_m = \left(\overline{\mathbf{B}}^H\overline{\mathbf{B}} + \overline{\mathbf{B}}_I^H\overline{\mathbf{B}}_I + \sigma^2\mathbf{I}\right)^{-1}\overline{\mathbf{B}}^H\mathbf{q}_m. \tag{B.21}$$

The elements of vectors $\hat{\mathbf{g}}_m$ are used to form the matrix $\hat{\mathbf{H}}$.

ACKNOWLEDGMENTS

The authors gratefully acknowledge Mr. Kari Horneman and Mr. Jyri Hämäläinen of Nokia Networks for their fruitful comments and suggestions. This work was supported by Nokia, Elektrobit, Finnish Air Forces, National Technology Agency of Finland (Tekes), Instrumentointi, Nokia Foundation, and Infotech Oulu Graduate School.

REFERENCES

[1] J. Winters, "Optimum combining in digital mobile radio with cochannel interference," *IEEE J. Select. Areas Commun.*, vol. 2, no. 4, pp. 528–539, 1984.

[2] J. H. Winters, J. Salz, and R. D. Gitlin, "The impact of antenna diversity on the capacity of wireless communication systems," *IEEE Trans. Commun.*, vol. 42, no. 2/3/4, pp. 1740–1751, 1994.

[3] J. Thomas and E. Geraniotis, "Space-time iterative receivers for narrowband multichannel networks," *IEEE Trans. Commun.*, vol. 50, no. 7, pp. 1049–1054, 2002.

[4] S. Verdú, *Multiuser Detection*, Cambridge University Press, Cambridge, UK, 1998.

[5] T. R. Giallorenzi and S. G. Wilson, "Multiuser ML sequence estimator for convolutionally coded asynchronous DS-CDMA systems," *IEEE Trans. Commun.*, vol. 44, no. 8, pp. 997–1008, 1996.

[6] T. R. Giallorenzi and S. G. Wilson, "Suboptimum multiuser receivers for convolutionally coded asynchronous DS-CDMA systems," *IEEE Trans. Commun.*, vol. 44, no. 9, pp. 1183–1196, 1996.

[7] C. Berrou, A. Glavieux, and P. Thitimajshima, "Near Shannon limit error-correcting coding and decoding: Turbo-codes. 1," in *Proc. IEEE International Conference on Communications (ICC '93)*, vol. 2, pp. 1064–1070, Geneva, Switzerland, 1993.

[8] L. Hanzo, T. H. Liew, and B. L. Yeap, *Turbo Coding, Turbo Equalization and Space-Time Coding for Transmission Over Fading Channels*, John Wiley & Sons, Chichester, UK, 2003.

[9] M. Tuchler, R. Koetter, and A. C. Singer, "Turbo equalization: principles and new results," *IEEE Trans. Commun.*, vol. 50, no. 5, pp. 754–767, 2002.

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

[11] M. C. Reed and P. D. Alexander, "Iterative multiuser detection using antenna arrays and FEC on multipath channels," *IEEE J. Select. Areas Commun.*, vol. 17, no. 12, pp. 2082–2089, 1999.

[12] P. D. Alexander, M. C. Reed, J. A. Asenstorfer, and C. B. Schlegel, "Iterative multiuser interference reduction: Turbo CDMA," *IEEE Trans. Commun.*, vol. 47, no. 7, pp. 1008–1014, 1999.

[13] T. Abe and T. Matsumoto, "Space-time turbo equalization in frequency-selective MIMO channels," *IEEE Trans. Veh. Technol.*, vol. 52, no. 3, pp. 469–475, 2003.

[14] D. Reynolds and X. Wang, "Turbo multiuser detection with unknown interferers," *IEEE Trans. Commun.*, vol. 50, no. 4, pp. 616–622, 2002.

[15] B. Lu and X. Wang, "Iterative receivers for multiuser space-time coding systems," *IEEE J. Select. Areas Commun.*, vol. 18, no. 11, pp. 2322–2335, 2000.

[16] C. Schlegel, S. Roy, P. D. Alexander, and Z.-J. Xiang, "Multiuser projection receivers," *IEEE J. Select. Areas Commun.*, vol. 14, no. 8, pp. 1610–1618, 1996.

[17] S. J. Grant and J. K. Cavers, "Further analytical results on the joint detection of cochannel signals using diversity arrays," *IEEE Trans. Commun.*, vol. 48, no. 11, pp. 1788–1792, 2000.

[18] S. Talwar, M. Viberg, and A. Paulraj, "Blind separation of synchronous co-channel digital signals using an antenna array. I. Algorithms," *IEEE Trans. Signal Processing*, vol. 44, no. 5, pp. 1184–1197, 1996.
[19] S. Chen, S. McLaughlin, B. Mulgrew, and P. M. Grant, "Bayesian decision feedback equaliser for overcoming co-channel interference," *IEE Proceedings–Communications*, vol. 143, no. 4, pp. 219–225, 1996.
[20] C. Luschi and B. Mulgrew, "Nonparametric trellis equalization in the presence of non-Gaussian interference," *IEEE Trans. Commun.*, vol. 51, no. 2, pp. 229–239, 2003.
[21] N. R. Veselinovic, T. Matsumoto, and M. J. Juntti, "A PDF estimation-based iterative MIMO signal detection with unknown interference," IEEE Commun. Lett., submitted, 2003.
[22] T. Abe, S. Tomisato, and T. Matsumoto, "A MIMO turbo equalizer for frequency-selective channels with unknown interference," *IEEE Trans. Veh. Technol.*, vol. 52, no. 3, pp. 476–482, 2003.
[23] B. W. Silverman, *Density Estimation for Statistics and Data Analysis*, Chapman and Hall, New York, NY, USA, 1986.
[24] M. Loncar, R. Muller, J. Wehinger, and T. Abe, "Iterative joint detection, decoding, and channel estimation for dual antenna arrays in frequency selective fading," in *Proc. 5th International Symposium on Wireless Personal Multimedia Communications*, vol. 1, pp. 125–129, Honolulu, Hawaii, USA, 2002.
[25] T. Cover and J. Thomas, *Elements of Information Theory*, John Willey & Sons, New York, NY, US, 1991.
[26] A. F. Naguib, *Adaptive Antennas for CDMA Wireless Networks*, Ph.D. thesis, Standford University, Standford, Calif, USA, 1996.
[27] B. Vucetic and J. Yuan, *Turbo Codes: Principles and Applications*, Kluwer Academic Publishers, London, UK, 2000.
[28] N. Veselinovic and T. Matsumoto, "Iterative signal detection in frequency selective MIMO channels with unknown cochannel interference," in *Proc. COST 273 Workshop on Broadband Wireless Local Access*, Paris, France, 2003.
[29] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*, Prentice Hall, New York, NY, US, 1993.

Nenad Veselinovic was born in Valjevo, Serbia and Montenegro, in 1975. He received his M.S. and Ph.D. degrees from University of Belgrade, Belgrade, Serbia and Montenegro, in 1999, and University of Oulu, Finland, in 2004, respectively. In 2000, he joined the Centre for Wireless Communications, University of Oulu, Oulu, Finland, where he is currently working as a Research Scientist. His main research interests are in statistical signal processing and receiver design for broadband wireless communications. He is a Member of IEEE.

Tad Matsumoto received his B.S., M.S., and Ph.D. degrees in electrical engineering from Keio University, Yokohama-shi, Japan, in 1978, 1980, and 1991, respectively. He joined Nippon Telegraph and Telephone Corporation (NTT) in April 1980. From April 1980 to Januray 1991, he researched signal transmission techniques, such as modulation/demodulation, error control, and radio link design schemes for 1st and 2nd -generation mobile communications systems. In July 1992, he transferred to NTT DoCoMo, where he researched code-division-multiple-access (CDMA) techniques. From 1992 to 1994, he served as a part-time lecturer at Keio University. In April 1994, he transfered to NTT America, where he served as a Senior Technical Advisor in the NTT-NEXTEL communications joint project. In March 1996, he returned to NTT DoCoMo, and he was appointed a head of radio signal processing laboratory at NTT DoCoMo, where he researched adaptive signal processing, MIMO Turbo signal detection, interference cancellation, and space-time coding techniques for broadband mobile communications. In May 2002, he moved to university of Oulu, Finland, where he is a Professor at the Center for Wireless Communications. Presently, he is serving as a Board-of-Governor of the IEEE VT Society for a term from January 2002 to December 2004.

Markku Juntti received his M.S. and Dr.S. degrees in electrical engineering from the University of Oulu, Oulu, Finland, in 1993 and 1997, respectively. Dr. Juntti has been with the University of Oulu since 1992. In 1994–1995, he was a Visiting Research Scientist at Rice University, Houston, Tex. In 1999–2000, he was with Nokia Networks as a Senior Specialist. Dr. Juntti has been a Professor of telecommunications at the University of Oulu since 2000. His research interests include communication and information theory and signal processing for wireless communication systems as well as their application in wireless communication system design. He is an author or coauthor in some 120 papers published in international journals and conference records as well as in the book *WCDMA for UMTS* published by Wiley. Dr. Juntti is a Senior Member of IEEE and an Associate Editor for IEEE Transaction on Vechicular Technology. He was the Secretary of IEEE Communication Society Finland Chapter in 1996–1997 and the Chairman for years 2000–2001. He has been the Secretary of the Technical Program Committee of the 2001 IEEE International Conference on Communication (ICC'01) and is the Cochair of the Technical Program Committee of 2004 Nordic Radio Symposium.