LDPC Coded MIMO Multiple Access With Iterative Joint Decoding

Amichai Sanderovich, *Student Member, IEEE*, Michael Peleg, *Senior Member, IEEE*, and
Shlomo Shamai (Shiz), *Fellow, IEEE*

Abstract—An efficient scheme for the multiple-access multiple-input multiple-output (MIMO) channel is proposed, which operates well also in the single user regime, as well as in a direct-sequence spread-spectrum (DS-CDMA) setting. The design features scalability and is of limited complexity. The system employs optimized low-density parity-check (LDPC) codes and an efficient iterative (belief propagation—BP) detection which combines linear minimum mean-square error (LMMSE) detection and iterative interference cancellation (IC). This combination is found to be necessary for efficient operation in high system loads $\alpha > 1$. An asymptotic density evolution (DE) is used to optimize the degree polynomials of the underlining LDPC code, and thresholds as close as 0.77 dB to the channel capacity are evident for a system load of 2. Replacing the LMMSE with the complex individually optimal multiuser detector (IO-MUD) further improves the performance up to 0.14 dB from the capacity. Comparing the thresholds of a good single-user LDPC code to the multiuser optimized LDPC code, both over the above multiuser channel, reveals a surprising 8-dB difference, emphasizing thus the necessity of optimizing the code. The asymptotic analysis of the proposed scheme is verified by simulations of finite systems, which reveal meaningful differences between the performances of MIMO systems with single and multiple users and demonstrate performance similar to previously reported techniques, but with higher system loads, and significantly lower receiver complexity.

Index Terms—Code-division multiple access (CDMA), iterative decoding, low-density parity-check (LDPC) code, multiple-input multiple-output (MIMO) channel, multiuser.

I. INTRODUCTION

MANY schemes have been designed to approach the overwhelming multiple-input multiple-output (MIMO) channel capacity [1], in particular, the case of channel state information (CSI) known at the receiver but not at the transmitter is often studied. A detailed description of the popular MIMO settings and their solved or unsolved capacity limits is given in [2]. The multiple-access MIMO capacity is given in [1], where it is shown that since the single-user capacity is achieved with transmissions that are uncoordinated between the antennas, which corresponds to a multiple-access channel (MAC), the sum rate multiple-access capacity over an ergodic channel is equal to the single-user capacity. Code-division multiple access

(CDMA) is closely related to MIMO as a single mathematical model models both, with the MIMO propagation coefficients corresponding to the symbols of the CDMA spreading sequences [3]. This supports the use of the well-studied CDMA-based multiple-access techniques in MIMO channels (with either single or multiple access). Although [1] relates to ergodic MIMO channels, which are not block fading, if one assumes an asymptotically large system, then each block considered is an ergodic and stationary realization of the same process. Thus, the equality of the multiuser and single-user capacities is true only as long as an asymptotic number of antennas is considered, while for a finite number of antennas, similar conclusions require further study of the outage capacity, which is beyond the scope of this paper.

The asymptotic capacity of the CDMA channel is explicitly calculated in [4] for unconstrained signaling, where the number of users and chips is taken to infinity, while their ratio remains fixed. The *Replica method* is used by Tanaka in [5] to extend this result to a binary phase-shift keying (BPSK) signaling and to a quadrature phase-shift keying (QPSK) signaling in [6]. Since the asymptotic analysis assumes randomly generated spreading sequences and since it relies on an eigenvalue distribution, which is robust to the statistics, the analysis remains valid for the MIMO channel (that is, a flat-fading MIMO channel with independent fading coefficients). Many proposed communication schemes are based on the diversity achievable through the MIMO channel. A good overview of such systems is given in [7].

We distinguish between systems which achieve good spectral efficiency [8]–[10] and which maximize the diversity: [11]–[13]. Iterative receivers, in which a detector and decoder exchange extrinsic information, are known to perform well in a variety of communication systems and specifically in multiple-access systems such as CDMA [14], [15]. The asymptotic information-theoretic loss due to separated decoding and detection is given in [16] and strongly justifies the use of iterative schemes. A CDMA multiuser detection technique used over MIMO channels is presented in [9] for lower channel loads. An asymptotic analysis of the linear minimum mean-square error (LMMSE) multiuser detector (MUD) is given in [17]–[19]. Boutros and Caire ([14]) include this analysis in a density evolution (DE) analysis of an iterative receiver for a trellis-coded CDMA. This DE analysis is also used in [20] for analyzing regular low-density parity-check (LDPC) codes over a CDMA channel. The use of an LDPC decoder with an LMMSE multiuser detector enables the elegant DE analysis of the

receiver in the asymptotic realm. DE was originally described by [21] for the analysis of belief propagation (BP) decoders for LDPC codes. These rediscovered codes are known to perform well and many researchers report excellent performance such as a code with a gap to capacity of 0.06 dB for an additive white Gaussian noise (AWGN) channel [22]. Such codes have been included in designs for fading channels [23] and finite MIMO channels [24].

In this paper, we use asymptotic techniques to construct a multiuser MIMO communication system employing LDPC codes and an LMMSE detector. The asymptotic approach enables us to employ exact analytical expressions for the MUD and the LDPC codes which could not be done for finite systems in [20], [25], [24]. We show that the cycle-free assumption holds so that the DE analysis is exact in the asymptotic realm and we also demonstrate its validity for rather large but still finite systems. Using the analysis we optimize the degree distribution pair (DDP) [21] of the underlying LDPC code which is iteratively decoded and detected with the LMMSE detector. Asymptotic techniques [26], [14] are also used for the capacity evaluation so that the quality of the scheme is verified. The LDPC codes are optimized to iterate with LMMSE detectors or alternatively with individually optimal multiuser detectors (IO-MUD). Comparing the thresholds of the resulting schemes to the capacity for high channel load, reveals excellent performance, with a gap to the capacity of 0.77 dB for the practical LMMSE and 0.14 dB for the IO-MUD, the complexity of which is prohibitive. The LDPC codes are optimized with a global search algorithm (differential evolution) to maximize the spectral efficiency. Simulations of finite systems verify the asymptotic expectations of both CDMA and MIMO for 100 antennas/chips, for single- and multiple-user systems with block lengths of 1.5×10^5 and 10^4, respectively. Simulation results of the design exemplify the robustness to the number of users.

This paper further demonstrates via examples and an extrinsic information transfer (EXIT) analysis the importance of code optimization for a multiuser setting where equal-rate equal-power users are enforced. This point is particularly important since in the existing literature there is a sort of dichotomy. On the one hand, results such as [27] and [28] consider the case where near-capacity performance is achieved by successive decoding enabled by allocating different powers to equal-rate users, or by allocating different rates to equal power users. In this case, it is clear that very powerful single-user codes perform well, and that the iterative multiuser decoding scheme serves just as a countermeasure for cleaning up the signal from residual errors. On the other hand, results such as those obtained in [29] for the two-user channel show that by properly designing the user codes, any point on the whole dominant face of the capacity region can be approached without the aid of time sharing. This indicates that by a proper multiuser code design, the iterative BP decoder can indeed approach the performance of a full joint multiuser decoder.

This paper confirms this fact in the large-system limit and MIMO MAC. It shows that codes which are suboptimal in the single-user case are able to achieve performance very close to the theoretical limit even though no shaping of the received powers and/or of the transmitted rates in order to induce implicit successive interference stripping is enforced.

In addition to the advantage of the reduced complexity of the transmitters in the multiuser scheme, the use of a single code is crucial when multiplexing a single long code between the transmitting antennas of a single user working over a finite MIMO channel. The proposed scheme also employs parallel scheduling, which is more practical in systems facing delay constraints.

In the following, we denote vectors by bold-face letters and the elements of the vectors as light-face letters with the proper subscript.

II. System Description

In this paper, we deal with the multiple-access communication channel where each user transmits an independent information via a single antenna over an independent Rayleigh-fading channel to a receiver equipped with M antennas. We assume full synchronization, perfect power control scheme with a single class of K users. A generalization of this work to several classes and/or received powers follows the same lines. The channel load α is defined by $\frac{K}{M}$ in parallel to an equivalent direct-sequence CDMA (DS-CDMA) setting [14]. The receiver, in the proposed scheme, has full knowledge about the CSI, while the transmitters have no CSI available. The MIMO channel coefficients $\{\boldsymbol{h}_{k,\tau}\}_{k=1}^K$ remain constant along an entire transmitted block and are then randomly and independently chosen again for the next transmitted block, so we drop the block index τ in the following. The statistics of this model describes the quasi-static block-fading MIMO channel, while the proposed analysis applies also to the case of a fast fading channel. The model also covers a multiple-access MIMO system with $K_u \leq K$ users, where each transmitter can use a number of $\frac{K}{K_u}$ antennas, not necessarily a single one. We can also consider the case of a single-user MIMO system, as a special case of the general setting above, where there is one user $K_u = 1$ using all $\frac{K}{1} = K$ transmitting antennas. Thanks to the similar asymptotic statistics of MIMO and random CDMA which are both characterized with elements drawn independently and identically distributed (i.i.d.) with vanishing odd-order and finite fourth-order moments [5], [18], the designed scheme is good also for random CDMA.

For example, in multiuser DS-CDMA systems: M is the processing gain, and $\{\boldsymbol{h}_k\}_{k=1}^K$ are the K signatures sequences.

A. The Transmitters and the Channel

The channel (1) is defined along the lines of [14]. The received vector $\boldsymbol{y}_t$ of length M, at symbol time t, consists of a linear superposition of the K transmitted symbols $\{x_{k,t}\}_{k=1}^K$ multiplied by the vectors of channel coefficients $\{\boldsymbol{h}_k\}_{k=1}^K$, by scalar random phases $\{e^{j\theta_{k,t}}\}_{k=1}^K$, and by the received amplitude $\sqrt{\frac{\gamma}{M}}$. The reception suffers from an additive Gaussian

complex noise $\boldsymbol{\nu}_t \sim \mathcal{N}_{\mathbb{C}}(0, \boldsymbol{I})$[1]

$$\boldsymbol{y}_t = \sqrt{\frac{\gamma}{M}} \sum_{k=1}^{K} \boldsymbol{h}_k e^{j\theta_{k,t}} x_{k,t} + \boldsymbol{\nu}_t. \quad (1)$$

In MIMO systems $\boldsymbol{h}_k \sim \mathcal{N}_{\mathbb{C}}(0, \boldsymbol{I})$. We use QPSK signals so

$$x \in \left\{ \pm \frac{1}{\sqrt{2}}, \pm \frac{j}{\sqrt{2}} \right\}.$$

To render the multiple-access interference independent of the transmitted codewords, $e^{j\theta_{k,t}}$ is randomly, independently, and uniformly generated from $\{\pm\frac{1}{\sqrt{2}}, \pm\frac{j}{\sqrt{2}}\}$ [14]. Thus, the performance seen by a user is independent of the symbols transmitted by the other users and thus the analysis can be limited to the case where all users transmit the all-zero codeword without reducing its generality. The receiver has full knowledge of the received powers $\frac{\gamma}{M}$, the channel coefficients $\boldsymbol{h}_k$, and the random phases $e^{j\theta_{k,t}}$. Notice that we normalized the received power with M so that each transmitting antenna is received with a total received power of $\frac{E_s}{N_0} = \gamma$, regardless of the number of receiving antennas, where E_s is the total received energy from one transmitted symbol and $N_0/2$ is the double-sided power spectral density. All the users encode their information bits $\{\boldsymbol{u}_k\}_{k=1}^{K}$ ($\boldsymbol{u}_k \in \{0,1\}^{\lfloor NR \rfloor}$) to $\{\boldsymbol{v}_k\}_{k=1}^{K}$ ($\boldsymbol{v}_k \in \{0,1\}^N$, $v_{k,n} = (\boldsymbol{v}_k)_n$) with the same LDPC code of length N and then apply an independent random bit interleaving ($\{\pi_u\}_{u=1}^{K}$) when the interleaved bits are denoted by $w_{k,\pi_k(n)} = v_{k,n}$. The interleaved vectors $\{w_{k,\pi_k(n)}\}_{k=1,n=1}^{K,N}$ are then modulated into QPSK symbols

$$x_{k,t} = \frac{1}{\sqrt{2}}(1 - 2w_{k,2t-1} + j(1 - 2w_{k,2t})). \quad (2)$$

We use an LDPC block code with length N and rate R and therefore each transmitted block consists of $t = 1, \ldots, \frac{N}{2}$ QPSK symbols and

$$\frac{E_b}{N_0} = \frac{E_s}{N_0} \frac{1}{\log_2(4)R} = \frac{\gamma}{2R}.$$

The MIMO transmitter is depicted in Fig. 1.

B. The Receiver

The signal ($\boldsymbol{y}_t$) is received by a MUD which produces soft information in the form of log-likelihood ratios (LLR)

$$\text{LLR}_{k,n} = \log \frac{\Pr(v_{k,n} = 0 \mid \boldsymbol{y}_t)}{\Pr(v_{k,n} = 1 \mid \boldsymbol{y}_t)} \quad (3)$$

about the individually coded bits of the K users (Pr or P stands for probability and log is the natural logarithm). These are passed on to the K (or K_u) LDPC decoders, which apply a decoding algorithm that produces extrinsic soft outputs. The soft outputs of the single-user decoders are fed back to the MUD (extrinsic LLR in the case of an optimal detector and extrinsic symbol estimation $\hat{x}$ in the case of a linear detector),
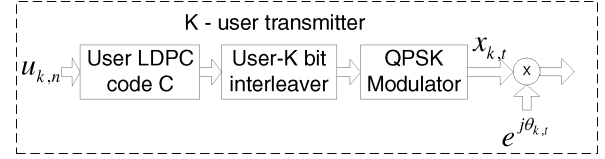
Fig. 1. A transmitter scheme, within a multiple-users system, when the pseudorandom phase is denoted by θ, the LDPC code as C, the information as u, and the modulated symbol as x. The rightmost arrow indicates the signal to be transmitted.

Fig. 2. An iterative receiver scheme. The MUD outputs are connected to the soft inputs of three decoders and the soft output (single-in single-out (SISO) decoders) of these decoders are fed back to the MUD (each arrow here represents N connections). After enough iterations, the outputs are hard-limited and then forwarded as the final decoded word $\{\hat{u}_{i,n}\}_{i=1,n=1}^{3,N}$.

which can then improve its outputs along the iterations, see Fig. 2. The MUD's outputs remain extrinsic with respect to the MUD's inputs, as seen in the sequel.

1) The Multiuser Detector: Recall from (1) that the inputs to the receiver are $\boldsymbol{y}_t, \{\boldsymbol{h}_k\}_{k=1}^{K}, \{\theta_{k,t}\}_{k=1}^{K}, \gamma$, and the inputs to the multiuser detector in the iterative scheme are

$$\boldsymbol{y}_t, \{\boldsymbol{h}_k\}_{k=1}^{K}, \{\theta_{k,t}\}_{k=1}^{K}, \gamma, \{\text{LLR}_{n,k}\}_{k=1,n=1}^{K,N}.$$

a) An individually optimal MUD: The nonlinear and highly complex IO-MUD ([26] and [27]), which ignores any code structure, can be used to calculate the LLR for the nth bit of the kth user. Denote $\boldsymbol{H} = [\boldsymbol{h}_1, \ldots, \boldsymbol{h}_K]$ and $\boldsymbol{\Theta} = \text{diag}(\theta_{1,n}, \ldots, \theta_{K,n})$, then the IO-MUD can be expressed as

$$\begin{aligned} \text{LLR}_{k,n} &= \log \frac{P\{w_{k,\pi_k(n)} = 0 \mid \boldsymbol{y}_t, \overline{\text{LLR}}_{[k,n]}\}}{P\{w_{k,\pi_k(n)} = 1 \mid \boldsymbol{y}_t, \overline{\text{LLR}}_{[k,n]}\}} \\ &= \log \frac{\sum_{\boldsymbol{w}_t \in \mathcal{U}_{k,n}^0} \exp(-\text{D}(\boldsymbol{w}_t) + \text{P}_{k,n}(\boldsymbol{w}_t))}{\sum_{\boldsymbol{w}_t \in \mathcal{U}_{k,n}^1} \exp(-\text{D}(\boldsymbol{w}_t) + \text{P}_{k,n}(\boldsymbol{w}_t))} \quad (4) \end{aligned}$$

where

$$\text{D}(\boldsymbol{w}_t) \triangleq \frac{1}{2}\left|\boldsymbol{y}_t - \sqrt{\frac{\gamma}{M}}\boldsymbol{H}\boldsymbol{\Theta}\boldsymbol{x}(\boldsymbol{w}_t)\right|^2$$

$$\text{P}_{k,n}(\boldsymbol{w}_t) \triangleq \frac{1}{2}(1 - 2\boldsymbol{w}_{[k,\pi_k(n)]})\overline{\text{LLR}}_{[k,\pi_k(n)]}^T,$$

and $\boldsymbol{w}_t \triangleq \{w_{j,2t-1}, w_{j,2t}\}_{j=1}^{K}$ is the vector of coded bits at the inputs of the K modulators at symbol time t, so that $\boldsymbol{w}_t \in \{0,1\}^{2K}$. The vector $\boldsymbol{x}(\boldsymbol{w}_t)$ is the vector of the K symbols that correspond to $\boldsymbol{w}_t$ and is calculated according to (2). $\mathcal{U}_{k,n}^0, \mathcal{U}_{k,n}^1$ are the two vector spaces over $\{0,1\}^{2K}$ that are spanned by all possible $\boldsymbol{w}_t$ with the nth bit of the kth user $w_{k,\pi_k(n)} = v_{k,n}$ set to be 0 and 1, respectively.

$$\boldsymbol{w}_{[k,\pi_k(n)]} \in \{0,1\}^{2K-1} \quad \text{and} \quad \overline{\text{LLR}}_{[k,\pi_k(n)]} \in \mathbb{R}^{2K-1}$$

are the $2K - 1$ interleaved bits and the corresponding LLR (the overline indicates that these LLR correspond to the interleaved bits) that result when excluding $w_{k,\pi_k(n)}$ and $\overline{\mathrm{LLR}}_{k,\pi_k(n)}$ from $\boldsymbol{w}_t$ and $\overline{\mathrm{LLR}}_t$, respectively. We define the function

$$s(k,n) : s(k,n) \in \left\{ 0, \ldots, \frac{N}{2} - 1 \right\}$$

such that $s(k,n)$ represents the interleaving and partitioning of N coded bits into $\frac{N}{2}$ symbols, performed in every transmitter, so that $t = s(k,n)$.

b) An LMMSE: The optimal MUD can be approximated by linear filters combined with interference cancellation (IC). This approach, which ignores any constellations constraints, is widely used and can be implemented with a polynomial complexity in the number of users [14]. A known effective linear multiuser detector is the LMMSE which minimizes the mean-square error (MSE) ξ at its output. In the iterative interference cancellation scheme (LMMSE-IC) we define $\xi_{j,t}$ to be the estimated power of the residue of the canceled j user in symbol time t

$$\xi_{j,t} = \mathrm{E}|x_{j,t} - \hat{x}_{j,t}|^2 \tag{5}$$

where $\mathrm{E}(x_{j,t} - \hat{x}_{j,t}) = 0^2$ and

$$\hat{x}_{j,t} = \frac{1}{\sqrt{2}} \left(\tanh\left(\frac{\overline{\mathrm{LLR}}_{j,2t-1}}{2} \right) + j \tanh\left(\frac{\overline{\mathrm{LLR}}_{j,2t}}{2} \right) \right). \tag{6}$$

We occasionally drop the time index t to allow simpler notations in the sequel. Since the cycle-free assumption holds for a finite number of decoder iterations (see Theorem 1 in Section III) and the users are uncoordinated and uncorrelated, we have

$$\mathrm{E}(x_k - \hat{x}_k)^*(x_j - \hat{x}_j) = 0, \qquad \forall\, j \neq k.$$

If we define $\boldsymbol{\Sigma}_k = \boldsymbol{I} + \gamma \sum_{i \neq k} \xi_i \boldsymbol{h}_i \boldsymbol{h}_i^H$ (the covariance matrix of the multiple-access interference (MAI) plus the noise) and

$$\hat{\boldsymbol{y}}_k = \sqrt{\frac{\gamma}{M}} \sum_{i \neq k} \boldsymbol{h}_i e^{j\theta_i} \hat{x}_i$$

then the estimation of the symbol of the k-user from the LMMSE-IC MUD minimizing ξ_k is [30]

$$z_k = \frac{\sqrt{\gamma} \boldsymbol{h}_k^H \boldsymbol{\Sigma}_k^{-1}(\boldsymbol{y} - \hat{\boldsymbol{y}}_k)}{\gamma \boldsymbol{h}_k^H \boldsymbol{\Sigma}_k^{-1} \boldsymbol{h}_k}. \tag{7}$$

The symbol estimation is converted to the LLRs of the coded bits with a simple QPSK demapper χ for the Gray mapping

$$\begin{aligned} &(\mathrm{LLR}_{k,\pi_k^{-1}(2t-1)}, \mathrm{LLR}_{k,\pi_k^{-1}(2t)}) \\ &= \chi(e^{j\theta_{k,t}} z_{k,t}) = \\ &= 2\sqrt{2}\,\mathrm{SNR}(\Re\{e^{j\theta_{k,t}} z_{k,t}\}, \Im\{e^{j\theta_{k,t}} z_{k,t}\}) \end{aligned}$$

where π_k^{-1} is the inverse interleaver mapping of the bits that correspond to the symbol which is estimated by $z_{k,t}$ and $\mathrm{SNR} = \frac{\gamma}{M} h_k^H \Sigma_{k,t}^{-1} h_k$ is the signal-to-noise ratio (SNR) at the output
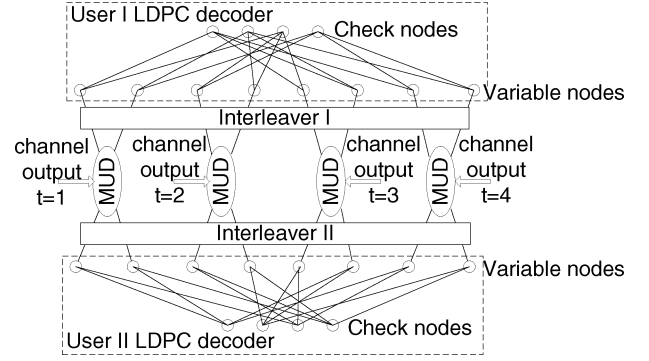
[2]E is the expectation operator.

Fig. 3. A receiver scheme for QPSK transmissions of two users, each encoding with a regular $(2,4)$ LDPC code and using a random bit interleaver.

of the LMMSE [30]. These equations represent both the conditional and unconditional LMMSE filters [14], [19]. The conditional LMMSE filter minimizes the MSE $\mathrm{E}|x_{k,t} - z_{k,t}|^2$ with the per symbol estimated powers of the interferers $\{\xi_{j,t}\}_{j=1,j\neq k}^K$ conditioned on $\{\hat{x}_{j,t}\}_{j=1,j\neq k}^K$, namely

$$\xi_{j,t} = 1 - |\hat{x}_{j,t}|^2.$$

On the other hand, the unconditional LMMSE assumes these estimated powers remain approximately constant along the transmitted block and estimates them once for the entire block as

$$\xi_{j,t} = \xi_j = 1 - \frac{2}{N} \sum_{t=1}^{\frac{N}{2}} |\hat{x}_{j,t}|^2.$$

Thus, the unconditional LMMSE requires the recalculation of the filter for every user and every iteration, but not for every symbol, which is done with the conditional LMMSE. The latter thus suffers an increase of approximately $N/2$-fold in the MUD complexity. It is seen in the following sections that in terms of complexity–performance tradeoff, the conditional LMMSE offers only a minor performance improvement over the unconditional LMMSE.

2) The Decoder: To decode the LDPC codes, the receiver utilizes the well-known BP decoder [21]. Such decoders are often described using a Tanner graph, which is composed of variable nodes and check nodes and of interconnecting edges. The variable nodes for the standard single-antenna single-user channel are initialized by the channel outputs (which remain constant along the iterations). The overall multiuser receiver graph, however, is composed of three classes of nodes: the $N \times K$ LDPC variable nodes and the $N \times (1 - R) \times K$ check nodes that represent the individual codes' parity-check equations, and the $\frac{N}{2}$ multiuser detector nodes that iteratively improve the LLRs. This is illustrated in Fig. 3, where two users with a regular $(2,4)$ LDPC user code of length 8 and QPSK signaling, are decoded and detected using four multiuser detector nodes and two LDPC Tanner graphs.

3) Iterative Detection and Decoding: The overall receiver steps through four phases in each iteration, as shown in Fig. 4. In the figure, the user interleavers and the LDPC code interleavers are denoted as π_u and π_c, respectively; χ represents the symbol
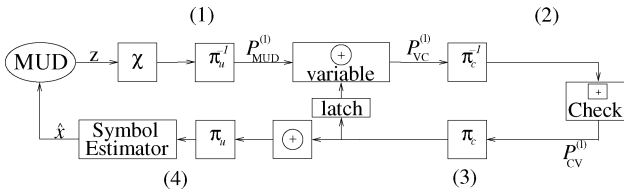
Fig. 4. Illustration of the four phases 1,2,3,4 in each receiver's iteration. The four phases are marked by (#).

to bit estimations conversion. The plus signs represent the check nodes, when they are enclosed by squares, and variable nodes, when they are enclosed by circles. On the first iteration the multiuser detector estimates $K \times \frac{N}{2}$ symbols without any *a priori* information from the SISO decoders, which means that it acts as detector in a separate detection-decoding scheme. During phase 1, all the likelihood ratios, calculated from these symbol estimations are simultaneously forwarded to all the respective variable nodes of all the users' decoders. These variable nodes then send messages to the check nodes, which reply back (phases 2 and 3). The messages from variable to check and from check to variable nodes are denoted as VC and CV, respectively. The replied CV messages are added at each variable node, to provide the multiuser detector with independent estimations of the transmitted symbols for the next iteration (phase 4). These CV messages represent the decoder's current extrinsic estimations of the corresponding bits.

The CV messages are stored during the first and fourth phases, so the variable nodes can sum the extrinsic information from both the MUD and the check nodes.

The message to the MUD (beginning of phase 3) and the message from the MUD (end of first phase) remain independent, since the MUD receives only a sum of extrinsic CV messages, without the initial values of the corresponding variable nodes, originally supplied by the MUD, and also due to the MUD construction (7), (4).

The proposed receiver uses parallel scheduling, so that all the users are simultaneously detected, decoded, and then subtracted in every iteration. It is different from e.g., BLAST (Bell Labs layered space–time) and [28] techniques which sequentially detect, decode, and subtract user after user. In this paper, we search for effective codes, so that the LDPC decoder would iteratively improve the multiuser detector's *a priori* inputs $(\hat{x})$ and reach a low bit-error rate (BER).

A receiver with IO-MUD (see (4)) does not require the symbol-to-bit χ and the bit-to-symbol estimation conversions, since the IO-MUD includes the constellation information and outputs the corresponding LLRs.

III. ASYMPTOTIC ANALYSIS

We follow [14] and [26] and use an analysis of asymptotic systems, where the code length, number of users, and number of antennas is taken to infinity, while the channel load and number of iterations remain fixed. Such analysis, which considers the statistics of the channel, gives rise to a decoder graph with no finite cycles for a trellis code, to concentration theorems, and also to elegant analytical expressions [14] when using the linear MUD. It is known from [26] that in such realm, the outputs

of both the LMMSE and the IO-MUD converge to Gaussian random variables. Thus, the multiuser efficiency η [4], defined as the SNR at the output of the multiuser detector divided by the SNR of a user with no interferers, is sufficient to describe the density.

In order to verify the convergence of the MUD's outputs, so the convergence theorems of [14] hold, in a system that uses LDPC code, first we have to ratify that the empirical distribution of the estimated interfering symbols $\hat{x}$, calculated from the SISO outputs, converges to some limiting distribution for a finite number of iterations. Unlike trellis codes, for which [14] proved the convergence of the SISO decoders' estimations, the outputs of LDPC BP decoder are time dependent, so one might suspect that since $N \rightarrow \infty$, and then $K \rightarrow \infty$, a situation where the SISO's symbol estimation has no limiting density can happen. To show that this is not the case and thus extend the proof of [14] to LDPC codes, first we claim the following.

Theorem 1: The randomly built graph of the entire receiver converges to a graph free of finite cycles as the block length $N \rightarrow \infty$.

Proof: We can use the LDPC single-user decoder cycle-free proof for finite number of iterations of [21] (notice that the cycle free proof of [14] does not hold here verbatim).

Assume a finite K and a maximum variable connectivity of $v_{\max}$. Since each MUD is connected to $2K$ variable nodes, the proof is completed by considering a respective random graph of a single-user LDPC decoder (which is free of finite cycles according to [21]) with a maximum variable degree of $2Kv_{\max}$ and a rate of $1 - \frac{cK}{N}$ (c is the number of check nodes in each block). Notice that we could not avoid the requirement of [14] that first $N \rightarrow \infty$ and only then $K \rightarrow \infty$ which indicates of systems with $K \ll N$, which is the case in many communication systems. □

An asymptotic graph with no finite cycles is referred to, in short, as a cycle-free graph hereafter, since we consider only a finite number of iterations. Now, since the graph indeed converges to a cycle-free graph, the following is straightforward.

Theorem 2: The empirical distribution of the estimated interfering symbols $F(\hat{x})$ which is calculated from the time-dependent SISO outputs $\boldsymbol{v}$, converges to some limiting distribution for finite number of iterations.

Proof: First, recall that a random LDPC graph with independent uniform interleavers $\{\pi_u\}_{u=1}^K$ (or equivalently, independently generated K random LDPC codes) is used.

Now, for any time index t, all the connected decoders' outputs $\{\text{LLR}_{k,n} : \pi_k(n) = 2t \text{ or } 2t - 1\}_{k=1}^K$ are generated according to the same density, since we consider an asymptotic number of users and the LDPC decoder has only a limited connectivity, which is smaller than K.

Formally, since the graph of the receiver is cycle free, the density of the decoders' outputs averaged over the graph ensemble, when $N \rightarrow \infty$ is actually a mixture of the densities of some function of the CV messages calculated by (8) below. This is given the all-zero codeword is transmitted (no loss in generality because of the pseudorandom phase). The asymptotic random building of the LDPC decoder assures the convergence of the

CV messages' densities to the ensemble average, as shown by the concentration theorem proven in [21]. Since the graph is cycle free, they converge to i.i.d. random variables. This leads to the probability law of the LLR

$$\Pr(\text{LLR}_{k,n} \leq V)$$
$$= \sum_{dv=1}^{v_{\max}} \mathrm{P}_{n,dv} \cdot \Pr\left(\tanh\left(\frac{\sum_{j=1}^{dv} \text{CV}_j}{2} \right) \leq V \right) \quad (8)$$

where $\mathrm{P}_{n,dv}$ is the probability that n is a variable node with degree dv, $v_{\max}$ is the maximum variable degree, and CV_j ($0 < j \leq v_{\max}$) is the message from the check node which is connected via the jth edge to the variable node. We recall that $v_{\max} \ll N$, so that in the asymptotic realm, the density of the LLRs converges to some limiting density. Note that this conclusion remains true only as long as $v_{\max}$ is finite, so we use LDPC codes, with finite connectivity.

Since $K \to \infty$, the LMMSE receives an infinite number of independent estimations

$$F_\xi(u) = \frac{1}{K} \sum_{k=1}^{K} F_{\xi_j}(u) = \frac{1}{K} \sum_{k=1}^{K} \Pr(1 - |\hat{x}_k|^2 \leq u)$$
$$= \frac{1}{K} \sum_{k=1}^{K} \Pr\left(\frac{2e^{\overline{\text{LLR}}_{k,2t-1}}}{(1+e^{\overline{\text{LLR}}_{k,2t-1}})^2} + \frac{2e^{\overline{\text{LLR}}_{k,2t}}}{(1+e^{\overline{\text{LLR}}_{k,2t}})^2} \leq u \right) \quad (9)$$

where $\hat{x}_k$ are calculated by (6). Now if $K \gg v_{\max}$ then the LLRs which are the feedback messages from the decoders manifest the probability law of (8). Actually, we see that the infinite number of users guarantees a single distribution at the MUD input. $\square$

This distribution is affecting the MUD in a similar way as a shaped power profile for the users (e.g., [27]), the difference being that this distribution is induced with no need for multiple classes.

The asymptotic analysis of key aspects of the system is presented in the following paragraphs.

A. An Analysis of the LDPC SISO Decoder

The limiting density of the asymptotic LDPC decoder's output $p(v)$ (see (8)) can be determined by the density evolution procedure [21].

In order to get numerical results, we assume without any loss of generality that the transmitted codeword is the all-zero codeword (ensured by the pseudorandom phase θ in (1)) and use the technique that was suggested by [22] to calculate densities that are defined over a quantized and finite space. This way, the calculation of the densities of the CV and VC messages, along the iterations, can be efficiently performed by a finite Fourier transform (FFT) for the VC densities and by a lookup table for the CV densities. The calculated densities are the limiting densities of the messages for $N \to \infty$.

B. Optimal Multiuser Detector Analysis

An asymptotic analysis of the IO-MUD detector, and the performance loss of using such detector in a separated scheme, is nicely presented by Verdú and Guo in [26], where it is shown that the asymptotic optimal MUD's output is a Gaussian random

variable (RV) with variance and mean that can be calculated via the replica method. However, [26] refers to the case where no *a priori* information is available to the detector, assumption that is no longer true in iterative schemes. Tanaka, Caire, and Müller in [27], extend the above analysis to the case where the IO-MUD detectors use *a priori* information. Their analysis is used for the limiting density of the IO-MUD's output (which is Gaussian). From [27], the asymptotic achievable multiuser efficiency η of an IO-MUD for real-valued system (such as real-valued CDMA) is the solution of the following fixed point equation:

$$\frac{1}{\eta} = 1 + \alpha\gamma \mathrm{E}_{\hat{x}}[(1 - \hat{x}^2) \int \frac{1 - \tanh(z\sqrt{\eta\gamma} + \eta\gamma)}{1 - \hat{x}^2 \tanh^2(z\sqrt{\eta\gamma} + \eta\gamma)} Dz] \quad (10)$$

where

$$\hat{x} = \tanh\left(\frac{\text{LLR}}{2} \right) \quad \text{and} \quad Dz = \frac{\exp(-\frac{z^2}{2})}{\sqrt{2\pi}} dz.$$

The replica method, used to derive (10), is also used for the capacities with BPSK [5] and QPSK [6] which are the same analytical expression. So we assume that the result (10), originally developed for the real case, extends to the complex channel. This is expected since the $H^{K \times M}$ complex channel has an equivalent representation by a matrix of real Gaussian coefficients $H^{2K \times 2M}$, the only difference being each coefficient appearing twice in the real matrix, the impact of which should be relaxed for asymptotic matrices. Also, the asymptotic equivalence of the real and complex channels is established for the linear MMSE detector [14] and (11) in what follows.

C. Linear Multiuser Detector Analysis

For the LMMSE in IC schemes, we can use the results of Boutros and Caire in [14] which include a closed-form expression for the descriptive statistics of filter's output, as a functional of the density of the estimated symbols. Although the results of [14] were obtained for a multiple-access system, where the users use trellis codes, it carries over to LDPC codes as well due to Theorem 2 (unisotropy degree). These useful results rely on the fact [17], [18] that when dealing with random matrices in the asymptotic regime, the resulting SNR converges to a deterministic limit. Considering Theorem 2, the conditions defined in [17], [18] are fulfilled for both CDMA and MIMO systems and the resulting asymptotic multiuser efficiency η of an LMMSE filter converges (weakly) to a deterministic value as $K \to \infty$. This value is the solution of the following fixed-point equation:

$$\eta = \frac{1}{1 + \alpha \int_0^\infty \frac{\gamma\xi}{1+\gamma\xi\eta} dF_\xi(\xi)} \quad (11)$$

where $F_\xi(\xi)$ is defined by taking the limit $K \to \infty$ of (9). This asymptotic result stems from the asymptotic distribution of the eigenvalues of the channel matrices (specifically, the eigenvalues of $\sum_{k=1}^{K} \boldsymbol{h}_{k,t} \boldsymbol{h}_{k,t}^H$), which is robust to the specific distribution of the matrix elements, and remains true for complex-valued $\boldsymbol{H}$.

Notice that in the asymptotic case, the only difference between the LMMSE-IC and the IO-MUD is in the resulting η, where for the LMMSE it is the solution of (11) and for the optimal detector it is the solution of (10).
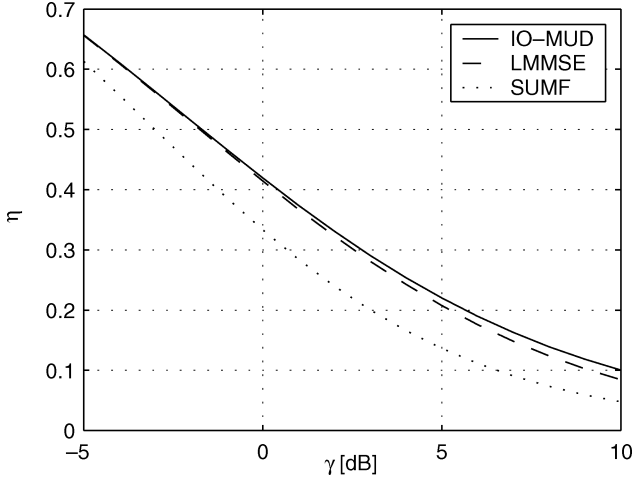
Fig. 5. The asymptotic multiuser efficiencies, for IO-MUD, LMMSE, and SUMF detectors, all for channel load of $\alpha = 2$ and various values of γ. No *a priori* information is assumed.

For the conditional LMMSE $F_\xi(\xi)$ is directly calculated from the output of the SISO decoder, that changes for every iteration, every symbol and every user. For the unconditional LMMSE $F_\xi(\xi) = \boldsymbol{u}(\xi - \bar{\xi})$,[3] where $\bar{\xi}$ is the block-wise average of the residues since the unconditional filter remains constant for all the symbols of the same user in the same iteration. Since the integral over ξ reduces to a simple expression, the fixed point of (11) for the unconditional LMMSE is explicitly written in (12)

$$\eta = \frac{2}{1 + (\alpha - 1)\gamma\bar{\xi} + \sqrt{4\gamma\bar{\xi} + (1 + (\alpha - 1)\gamma\bar{\xi})^2}}. \quad (12)$$

It is noticed that the performance of the conditional LMMSE will never be worse than the performance of the unconditional LMMSE, because of the convexity of $\frac{\xi}{1+\xi\eta}$ in (11).

In a similar manner, using results of [17] and [14], the SNR at the output of a multiuser detector which uses a single-user matched filter (SUMF) with an IC converges to

$$\eta = \frac{1}{1 + \alpha\gamma\bar{\xi}}. \quad (13)$$

Fig. 5 demonstrates the differences between these asymptotic multiuser efficiencies as a function of the SNR γ. It can be seen that the LMMSE's asymptotic multiuser efficiency is significantly better than that of the SUMF and is only slightly degraded compared to that of the IO-MUD.

D. The Gaussian Approximation of the DE

Although the DE predicts the actual performance very well, it also requires a considerable amount of computations, since it uses quantization of real-valued continuous functions. A less complex analysis, which assumes that messages are distributed according to Gaussian densities and therefore iterates only a single parameter, is the Gaussian approximation technique [31]. Since the outputs of the asymptotic LMMSE and the IO-MUD are Gaussian ([18] and [26]), the code optimization process can utilize this approximation and benefit from its low complexity,

---

[3]$\boldsymbol{u}(x)$ is the indicator function, that equals 1 for $x \geq 0$ and 0 otherwise.

leaving the DE for a more accurate evaluation of the optimization products and for verifying the results. All the results presented in the tables and figures in this paper are calculated by the DE to ensure accuracy. Although the Gaussian approximation is motivated by the central limit theorem for the variable nodes and remains inaccurate for the check nodes, it is close enough, in most cases, to the results of the DE. We modified the technique of [31] for the multiple-access receiver, where the intrinsic LLRs originate from the MUD instead of from the memoryless channel and therefore change along the iterations. Since both CV and VC are assumed to be Gaussian, only the mean values m_{CV} and m_{VC} are used by the analysis. If $m_{\mathrm{CV}}^{(l)}$ is the mean of the messages that are sent from the check nodes to the variable nodes at the lth iteration then $m_{\mathrm{CV}}^{(l+1)}$ of the next iteration is calculated by (14). The mean value of the messages from the variable nodes with degree i to the check node in the lth iteration is $m_{V_iC}^{(l)}$. The asymptotic multiuser efficiency of the MUD is represented by $\eta = \psi_{\mathrm{MUD}}$ in (14), where ψ_{MUD} includes the various possibilities (e.g., (8), (10), (12)) for calculating η given the density of CV messages. $\{\rho_j\}_{j=2}^{c_{\max}}$, $\{\lambda_i\}_{i=2}^{v_{\max}}$ are defined in [21]

$$m_{V_iC}^{(l)} = 2\gamma\psi_{\mathrm{MUD}}(m_{\mathrm{CV}}^{(l)}) + (i-1)m_{\mathrm{CV}}^{(l)}$$
$$m_{\mathrm{CV}}^{(l+1)} = \sum_{j=2}^{c_{\max}} \rho_j \phi^{-1}\left(1 - \left[1 - \sum_{i=2}^{v_{\max}} \lambda_i \phi(m_{V_iC}^{(l)})\right]^{j-1}\right) \quad (14)$$

and $\phi(x)$ is defined as in [31] as

$$\phi(x) = \begin{cases} 1 - \frac{1}{\sqrt{4\pi x}} \int_{\mathbb{R}} \tanh(\frac{u}{2}) e^{-\frac{(u-x)^2}{4x}} \, du, & \text{if } x > 0 \\ 1, & \text{if } x = 0. \end{cases}$$

E. Capacity

The capacity for the single-user MIMO channel, when the CSI is known at the receiver but unknown to the transmitter, is achieved [1] when a coded information is sent with the same power and the same code rate from all the transmitting antennas. Thus, this single-user capacity is equal to the MAC's sum rate capacity and if we will find a good multiple-access scheme it can be used for the single-user case as well (as long as we remain in the asymptotic region). The asymptotic sum-rate capacity of a multiuser system constrained to BPSK is calculated in [26]. It was originally derived by Tanaka in [5] assuming the replica method was valid and [26] extended the results for the capacity when using linear detectors (noniteratively). Assuming BPSK symbols with averaged amplitude $\sqrt{\gamma}$ and a channel transfer matrix with i.i.d, zero mean, vanishing odd-order moments and finite fourth-order moments entries (such as the MIMO and CDMA settings), as $M, K \to \infty$, the resulting asymptotic multiuser efficiency η of an individual optimal detector is the solution of the following fixed-point equation:

$$\frac{1}{\eta} = 1 + \alpha\gamma\left[1 - \sqrt{\frac{\eta\gamma}{2\pi}} \int_{-\infty}^{\infty} \tanh(\eta\gamma x) \right.$$
$$\left. \cdot \exp\left(-\frac{\eta\gamma(x-1)^2}{2}\right) dx\right] \quad (15)$$
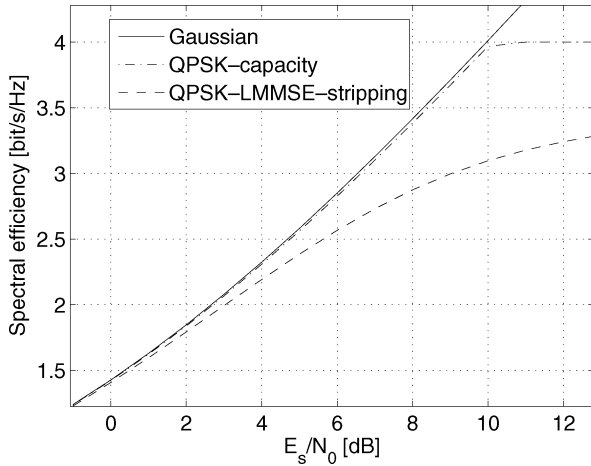
Fig. 6. The Loss due to the use of QPSK instead of an optimal constellation in a MIMO setting, and due to the use of an LMMSE in a stripping receiver, for the asymptotic case with channel load of $\alpha = 2$.

TABLE I
DISTANCES TO CAPACITIES (IN DECIBELS) WITH THE OPTIMIZED CODES AND AN UNCONDITIONAL LMMSE DETECTOR

E_s/N_0	1 dB	2 dB	3 dB	4 dB
$\alpha = 0.2$	0.33	0.31	0.38	0.40
$\alpha = 0.5$	0.44	0.44	0.55	0.63
$\alpha = 1$	0.66	0.69	0.8	1
$\alpha = 1.5$	0.77	1.06	1.09	1.26
$\alpha = 2$	0.77	1.16	1.35	1.69

where the AWGN has variance of 1 and 0 mean as before. The capacity is then calculated with

$$C = \eta\gamma + \frac{\eta - 1 - \log(\eta)}{2\alpha}$$
$$- \sqrt{\frac{\eta\gamma}{2\pi}} \int_{-\infty}^{\infty} \log\Big(\cosh(\eta\gamma x)\Big) \exp\Big(-\frac{\eta\gamma(x-1)^2}{2}\Big) dx. \quad (16)$$

If (15) has more than a single fixed-point solution, the correct one is the one that results with the smallest capacity.

Note that this capacity is computed when the signaling is assumed to be BPSK.

This capacity result and the optimal asymptotic MIMO capacity results obtained by Telatar in [1] are used to assess the information-theoretic loss, at least for the asymptotic system, when we constrain ourselves to QPSK signaling. It can be seen that the asymptotic capacity of the complex QPSK system is exactly twice the asymptotic capacity of the BPSK, real-valued channel system [6]. Define the spectral efficiency as $\rho = \alpha C$. The spectral efficiency of a multiuser system with a receiver which serially LMMSE filter, decode, and then cancel each user (stripping LMMSE receiver) is presented in [28] as an integral expression. We compared the spectral efficiencies of these schemes in Fig. 6. This figure depicts the Gaussian optimal MIMO, the QPSK, and the stripping LMMSE asymptotic spectral efficiencies for load of $\alpha = 2$ for various γ values.

Viewing the figure in the lower SNR region, we conclude that the asymptotic capacity using QPSK signaling is very close to the optimal asymptotic capacity and that using a stripping LMMSE receiver causes another small loss here. The inflicted loss due to the use of stripping LMMSE with profiling significantly increases for larger SNR, while the loss due to the QPSK constellation restriction is relatively mild. These observations indicate that the use of QPSK signalling is near optimal up to $\gamma = 9$ dB while the LMMSE stripping is nearly optimal at considerably lower γ. However, this does not guarantees that our approach outperforms the stripping decoder.

Verdú already showed that such QPSK signaling is second-order optimal in [32], further substantiating our conclusion that larger constellations do not provide a meaningful gain at the

lower SNR region. This is also consistent with the intuition presented in [33], where it is expected that as the SNR gets smaller and the channel load α larger, the loss inflicted by the use of a small constellation is reduced. Note that the number of antennas at the receiver and the MUD's complexity are reduced when the channel load is high.

IV. CODE SEARCH AND RESULTS

In this section, we present the results of the search for good LDPC DDPs [21]. The problem of finding good codes is a global optimization problem that maximizes the rate of the code R under the constraint of asymptotic error-free decoding for a given γ and α.

A. Search Algorithm

The search for good codes is performed by a stochastic genetic algorithm known as differential evolution. This algorithm is used for optimizing LDPC codes also by [21] and [23]. Briefly speaking, this algorithm starts off with an initialization stage and then iteratively repeats two stages, until a convergence is achieved [34]. In the initialization stage, an initial DDP's population is randomly chosen according to the uniform distribution over the constrained space. The code rate is used as the parameter to be optimized (unlike in [21], [23], and [25]) instead of the threshold, so reducing the computational load. The following stage consists of randomly altering each DDP in the population by generating random perturbations from randomly selected pairs or quadruplets of DDPs [34]. In the second stage, the population undergos a natural selection process which leaves either the DDPs or their altered versions, depending on their performances. The main idea is that the population gets better and more homogeneous as the iterations progress [34].

B. Search Results

Good pairs are found for five channel loads: $\alpha = 0.2, 0.5, 1, 1.5$ and $\alpha = 2$, each with four threshold constraints: $\gamma = E_s/N_0 = 1, 2, 3,$ and 4 dB. The DDPs were obtained while further limiting the search space by constraining the check and variable degrees in order to reduce computation time.

Table I presents the results of the search process as the gap between the achieved asymptotic thresholds and the asymptotic QPSK capacities (as presented in Section III-E). The loss due to the QPSK restriction can be appreciated by comparing this capacity to the optimal MIMO capacity, as shown in Fig. 6, and we notice a gap of only 0.05 dB for load $\alpha = 2$ and $\gamma = 4$ dB. The scheme reaches up to 0.3 dB from the channel capacity for
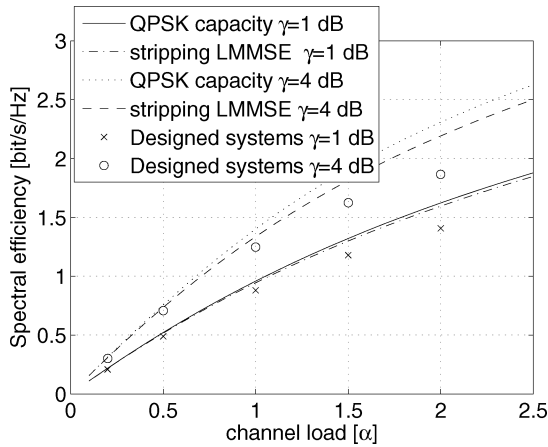
Fig. 7. The spectral efficiencies of the designed systems (markers) with the corresponding QPSK capacity (solid and dotted lines) and QPSK with LMMSE stripping receiver (dashed and dot-dashed lines) bounds. All are presented as a function of the channel load, for $E_s/N_0 = 1$ and 4 dB.

TABLE II
DISTANCES TO CAPACITY OF THE CODE DESIGNED FOR MULTIPLE-ACCESS AND OF A CODE OPTIMIZED FOR THE AWGN OVER BOTH SCENARIOS

	distance to capacity on AWGN channel in [dB]	distance to capacity on multi access channel $\alpha = 2$ in [dB]
MAC-good code R=0.466	3.42	1.64
AWGN-good code R=0.5	0.2	9

$\alpha = 0.2$ and gets as close as 0.77 dB from the QPSK capacity for $\alpha = 2$.

Fig. 7 presents the bounds on the spectral efficiencies derived from the QPSK capacity (16) and the QPSK LMMSE stripping receiver [28]. Those bounds are only about 0.1 dB apart for the γ and α examined here. The spectral efficiencies of the designed systems are presented as markers and exhibit near optimality except when the channel load and SNR are high.

C. The Necessity of Fine Tuning the Code

We compared the code that was designed for a channel load of 2 and threshold value of $\gamma = 4$ dB to an LDPC code that was designed for an AWGN and denoted in the following as AWGN-good code (we use the term AWGN to indicate a single-user additive white Gaussian channel). We choose a code with $dv = 12$ from [21, Table I]. The former code is check-regular and has $dv = 19$ and a rate of $R = 0.466$, while the latter has a rate of $R = 0.5$, so they are quite close in that sense. The differences between the thresholds of these codes and the channel capacity are shown in Table II over both the multiple-access MIMO channel and the single-user, single-antenna channel. Notice that although the AWGN-good code performs very well on the AWGN, its performance on the MAC is poor (even when considering the fact that it has slightly higher rate). In contrast, the code that was designed for the MAC was not as bad on the AWGN channel (a difference of 3.4 dB). This emphasizes the importance of including the multiuser analysis in the search for good codes and not using off-the-shelf AWGN-good codes.

It is of interest to examine whether the performance gap between the AWGN-good and MAC-good codes is caused by some basic characteristic inherent in the codes and independent of the details of the MUD. To this end, we perform an EXIT analysis at the interfaces between the MUD and the LDPC decoders. To examine the characteristics of the LDPC decoder we let it, for the purpose of this analysis only, perform an unlimited number of iterations after each activation of the MUD.

This modification of the scheduling would probably increase the receiver complexity, however, it would not change the final outcome.

This is since as long as an unlimited decoder activations are allowed and the extrinsic information transfer functions are non-decreasing

$$I_{\text{out}}(I_{\text{in}}) \leq I_{\text{out}}(I_{in} + \epsilon) \tag{17}$$

the threshold of an iterative scheme is independent of the iterations scheduling (see, e.g., [35]). I_{in} (or I_{out}) denotes the averaged mutual information between the LLR at the LDPC decoder input (or output) and the transmitted coded bits. It is known ([21]) that the output of an infinite LDPC BP decoder is indeed nondecreasing, and (12) verifies this property also for the multiuser detector. Thus, the threshold value will not change by rescheduling the iterations.

This assures that rescheduling, such that the inner LDPC decoder is infinitely iterated and only then forwarded back to the multiuser detector, does not alter the overall receiver threshold value. The receiver in the modified scheduling does not need to store information in the decoder while the multiuser detector is working (so that the information transfer characteristic of the code is independent from the outer iteration number). This means that the latch in Fig. 4 can be removed for such scheduling. In that figure, this rescheduling means that inner phases 2–3 are repeated infinitely before continuing to outer iterations 1, 4. The extrinsic information characteristics is extracted from the asymptotic analysis (Section III), where the DE of the LDPC decoder is repeated until no further improvement is achieved. Notice that this EXIT analysis is different from that in [24], where the check and the variable nodes were regarded as separate components, while here the analysis reveals the characteristics of the iterative scheme when we consider the LDPC code as a single code without any memory and the MUD as the other component of the iterative scheme. The results are presented in Fig. 8. The scheme with the AWGN-good code from [21] operates with $\gamma = 12$ dB ($\frac{E_b}{N_0} = 15$ dB) and the same scheme, only with the MUD-optimized code operates with $\gamma = 4$ dB ($\frac{E_b}{N_0} = 7.312$ dB) so that both schemes work close to their corresponding thresholds. Notice that the curves start with I_{in} that corresponds to the first MUD detection.

Information-theoretic arguments presented in the Appendix reveal that a good random AWGN code is characterized by a low-valued EXIT curve, until a threshold value of I_{in} (approx-
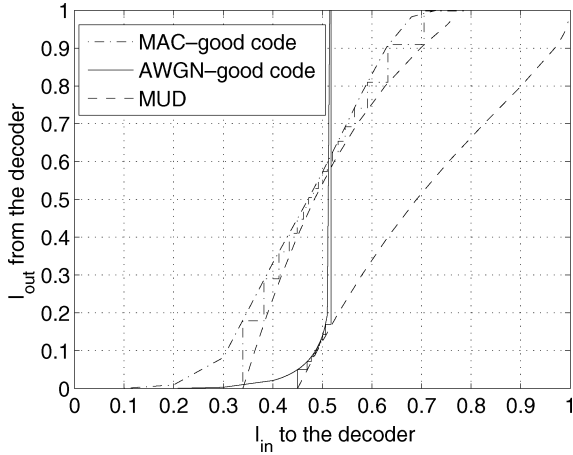
Fig. 8. EXIT charts for system with load of $\alpha = 2$, for both AWGN-good code (solid line) and fine-tuned MAC code (dash-dot line). The AWGN-good code is used with SNR of $\gamma = 12$ dB ($\frac{E_b}{N_0} = 15$ dB) and the MAC-good code with $\gamma = 4$ dB ($\frac{E_b}{N_0} = 7.312$ dB). The actual information transfer trajectories extracted from the DE are the steps-like lines. The EXIT characterization of the MUD for $\gamma = 4$ dB and $\gamma = 12$ are presented as the leftmost and the rightmost dashed line, respectively.

imately the AWGN capacity) is encountered. Such code also achieves an error-free decoding ($I_{\text{out}} = 1$) for I_{in} which is slightly larger than its threshold, resulting in a nearly vertical EXIT curve near the threshold. This curve characterization actually prevents a good random AWGN code from performing well in our setting, since the multiple-access detector does not receive a sufficient feedback from the decoder until the inputs to the decoder are close to the single-user AWGN capacity, and then the decoders can finish decoding without further outer iterations with the MUD detector (a separated scheme).

Actually, even the slope of the simple repetition code better fits the EXIT curve of the MUD than that of the AWGN-good code.

Since we used a code which better fits the MUD, it produces suitable lower error probabilities (higher I_{out}) for code rates that exceed the AWGN capacity. This difference is well observed in Fig. 8 for $I_{\text{in}} = 0.45$, where the MAC-good code demonstrates $I_{\text{out}}(0.45) = 0.45$ whereas the AWGN-good code achieves only $I_{\text{out}}(0.5) = 0.05$.

The EXIT curves of the MUD for $\gamma = 4$ dB and $\gamma = 12$ dB are presented as the leftmost and the rightmost dashed line, respectively. It is seen that as γ increases, the line indeed becomes steeper. Notice that the optimized code has an EXIT curve which fits quite well to the EXIT curve of the MUD.

Observing the DDP [21] of both codes

$$
\begin{aligned}
\lambda_{\text{AWGN}}(x) =\ & 0.2442x + 0.2591x^2 + 0.0105x^3 + \\
& + 0.0551x^4 + 0.0146x^7 + 0.0128x^9 \\
& + 0.4037x^{11} \\
\rho_{\text{AWGN}}(x) =\ & 0.2547x^6 + 0.7344x^7 + 0.0109x^8 \\
\lambda_{\text{MAC}}(x) =\ & 0.7202x + 0.2798x^{18} \\
\rho_{\text{MAC}}(x) =\ & x^4
\end{aligned}
$$

we can calculate that the AWGN-good code has a connectivity [36] of 7.7 per information bit whereas the MAC-good code has a connectivity of 5.7 per information bit. This further indicates

TABLE III
DISTANCE TO CAPACITY IN (DECIBELS), AS ACHIEVED WITH IO-MUD, CONDITIONAL LMMSE, AND UNCONDITIONAL LMMSE, FOR $\alpha = 2$ AND THRESHOLD OF $\gamma = 1$ dB

	unconditional LMMSE	conditional LMMSE	IO-MUD
distance [dB]	0.77	0.67	0.14

that the AWGN-good code is indeed better for approaching the AWGN capacity [36] and thus performs poorly in the joint iterative scheme.

D. On the Degradation Due to the Use of a Linear MUD

The gaps to the channel capacity for high channel loads, as appear in Table I, are mainly due to the limitations of the linear MUD. To confirm this conclusion, we searched for good DDPs for iterative schemes which include the conditional LMMSE-IC and the IO-MUD. The distances between the thresholds of these systems to the capacity, with channel load $\alpha = 2$ are presented in Table III. A small improvement of 0.1 dB for the conditional LMMSE over the unconditional LMMSE is noticed compared with the more significant 0.63-dB improvement with the IO-MUD. This quantifies the limitations of the linear detectors, especially for high channel loads.

V. SIMULATION RESULTS

The system was simulated to verify the analysis and to assess the performance with finite number of users. All the simulations were performed with a channel load of $\alpha = 2$ except for systems with $K = 15$ which were simulated with $\alpha = 1.875$. All the LDPC codes were constructed according to the DDP that was found by the asymptotic analysis for a threshold of $E_s/N_0 = 1$ dB and for a channel load of $\alpha = 2$. The resulting code rate is $R = 0.352$.

A. Comparison of the Asymptotic Analysis to the Simulations of Finite Systems

The BER predicted by the asymptotic analysis and simulation results are drawn in Fig. 9 as a function of the iterations. These simulations were done for a multiuser MIMO system with $K = 200$ users ($\alpha = 2$) employing an LDPC code of length $N = 10^4$ and $\gamma = E_s/N_0 = 1.45$ dB ($E_b/N_0 = 2.98$ dB). It can be seen that the DE predicted the performance of the system well, until about the 20th iteration. The simulations will probably agree with the DE prediction along more iterations if the codeword length N will be increased.

B. Finite LMMSE

Separated simulations of a finite LMMSE filter, designed for either a CDMA with QPSK spreading sequences ($h_k \sim \frac{1}{\sqrt{2}}\{\pm 1 \pm j\}^M$) or for a MIMO ($h_k \sim \mathcal{N}_{\mathbb{C}}(0, I)$) channel can assess the difference between the asymptotic and the average MSE at the output of a finite filter. It also indicates how many users such system should include to approach its asymptotic performance. These simulations test the transmission of an uncoded symbols through MIMO or nonfading CDMA channels. The receiver uses an LMMSE filter where
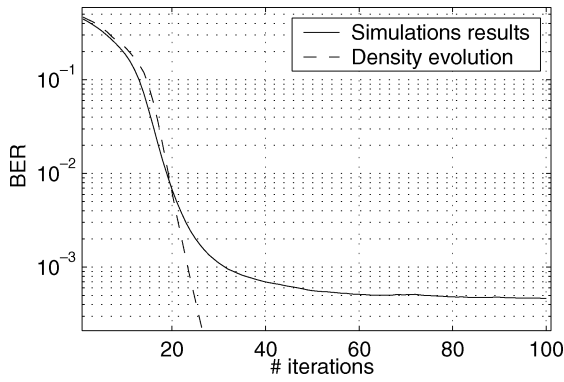
Fig. 9. Simulation results versus DE of the average of the check to variable messages of the LDPC decoders, in a MIMO system with $K = 200$ users, that uses an LDPC code with $R = 0.352$ and $N = 10^4$, and a channel load $\alpha = 2$, $E_s/N_0 = 1.45$ dB ($E_b/N_0 = 2.98$ dB), as a function of number of iterations.
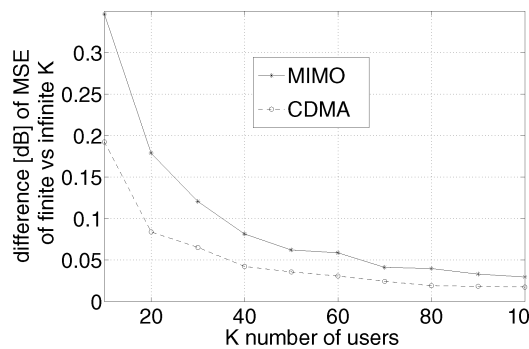

Fig. 10. Differences, in decibels, of the MSE at the output of finite LMMSE filters (averaged over enough blocks) to the MSE of an asymptotic LMMSE filter, as calculated by (12). The differences are plotted for $E_s/N_0 = 1.2$ dB and $\alpha = 2$, with both non fading CDMA ($h_k \sim \frac{1}{\sqrt{2}}\{\pm 1 \pm j\}^M$, dashed line) and MIMO ($h_k \sim \mathcal{N}_\mathbb{C}(0, I)$, solid line) systems as a function of the number of users K.

both the power of the users and the channel transfer coefficients are known. In the following, we compare the average MSE of the LMMSE output to the asymptotic MSE as calculated by (12) for different K and channel statistics. Notice that according to [18], the asymptotic behavior of these two channels is identical but the convergence rates depend on higher order statistics. Fig. 10 plots these differences in decibels, for systems with channel load $\alpha = 2$ and $\gamma = 1.2$ dB as a function of K. It is noticed that the difference to the asymptotic MSE is about 0.35 dB, for the MIMO system with as few as ten users (five receive antennas) and 0.2 dB for the nonfading CDMA system with ten users and five chips. This difference further decreases with the number of antennas. As expected, the MSE for both MIMO and nonfading CDMA finite filters converges to the asymptotic MSE with different rates, although these rates are quite close. These results demonstrate the effect of the finiteness of the filter on the performance, excluding any effect of feedback from the decoders.

C. MIMO Versus Nonfading CDMA Systems

Fig. 11 shows the performance of a multiuser MIMO system versus a nonfading CDMA system for both $K = 15$ (squares) and (circles) $K = 200$ users. It presents the averaged BER from simulations of these systems, for various E_b/N_0, for channel loads $\alpha = 2$ ($K = 200$) and $\alpha = 1.875$ ($K = 15$) and for the
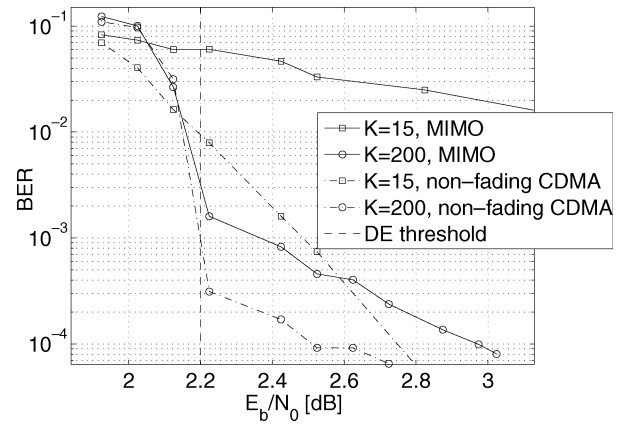

Fig. 11. Simulation results of a multiuser MIMO ($h_k \sim \mathcal{N}_\mathbb{C}(0, I)$, solid lines) versus a nonfading CDMA ($h_k \sim \frac{1}{\sqrt{2}}\{\pm 1 \pm j\}^M$, dash-dot lines) systems where both systems use the same code with block length $N = 10^4$. The comparison is done for $K = 15$ (squares) and 200 (circles) users, with channel loads $\alpha = 1.875$ and $\alpha = 2$, respectively. The results are presented as a function of E_b/N_0 after sufficient number of iterations at the receiver. The theoretical threshold is presented, for comparison as a dashed line.

same code with length $N = 10^4$. In these simulations, the receiver iterated the decoding-detection cycle until no further improvements were noticed, so that the graph represents the performance after enough receiver iterations. We can see that in both MIMO and CDMA systems, the more dimensions we operate over (diversity), the better the BER performance we get. So a system with only $K = 15$ users is worse than that of $K = 200$ users, in the lower BER region ($E_b/N_0 > 2.12$ dB). We note that reducing the number of users in the very low SNR region helps to improve the BER performance. Notice, also, that while the nonfading CDMA system performs quite well, even with only $K = 15$ users, the performance of the MIMO system with the same number of users and the same channel load is severely degraded. When increasing the dimensionality to $K = 200$ users, the MIMO and the nonfading CDMA performances are close, especially in the waterfall region. The difference between them is in the higher SNR region, where the nonfading CDMA system performs better. The fading effect, as was seen in Section V-B, is evident for the iterative receiver in Fig. 11, when comparing the multiuser nonfading CDMA system curve to the MIMO curve, both with $K = 15$ and $K = 200$. The very poor performance of the multiuser MIMO system with $K = 15$ is attributed to the outage probability of the MIMO channel and with $K = 200$ users, the spacial diversity overcomes the fading effect. Since we deal with independent codes on a block-fading channel, a severely faded block can impair the user's code ability to correct any errors and therefore also denies any possibility for reducing the MAI of other users with possibly better channels. This situation is more rare in the CDMA system, since no fading is considered and the channel is impaired only through the correlations between the spreading sequences.

D. Single-User Versus Multiple-User MIMO Block Fading

The simulations of multiple-access MIMO systems approached the prediction of the asymptotic analysis with $K = 200$ users (although sufficient performance was achieved with only $K = 50$ users). That means using $M = 100$ antennas
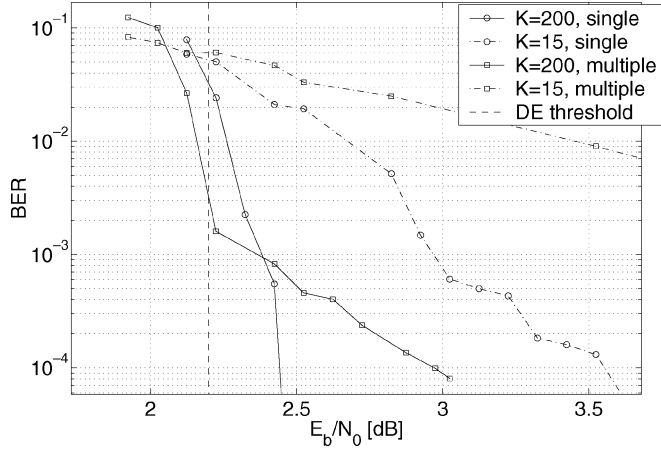
Fig. 12.   Simulation results of a multiuser scheme with a block length of $N = 10^4$ (squares) versus a single-user MIMO system with a block length of $N = 15 \times 10^4$ (circles). Both codes have the same rate $R = 0.352$ and the schemes were simulated with $K = 200$ (solid lines) and $K = 15$ (dashed lines) users with channel loads $\alpha = 2$ and $\alpha = 1.875$, respectively. The BER presented was measured after a sufficient number of iterations, as a function of E_b/N_0. DE threshold is also presented as a dashed line for comparison.

Fig. 13.   Simulation results of a MIMO receiver with channel coefficients estimation errors with a variance of ϵ, ($K = 100$ users, channel load is $\alpha = 2$, and $\gamma = 2$ dB) is drawn as a solid line. The used code is the same code used in the previous section with length $N = 10^4$. Simulation of a system with perfect CSI, but with a modified signal power $\gamma \leftarrow \frac{\gamma}{1+\epsilon\gamma}$ is drawn by a dashed line, for comparison.

in the receiver. Significantly less users, such as $K = 15$ (with a lower channel load $\alpha = 1.875$) results in a severe degradation of the performance. A single-user MIMO system, however, can utilize a single long LDPC code along with a multiplexer to transmit from all the antennas. The corresponding receiver uses a single decoder which will benefit from a much more diverse channel (factor of K) than the ones the multiple-access decoders faced. Thus, now the event of a deep fade in a few of the propagation coefficients can be mitigated by the LDPC code directly and not by the interactions with the MUD as in the multiple-access case. This effect can be seen in Fig. 12, where we simulated a single-user (circles) and a multiple-user (squares) MIMO systems with code lengths of $N = 15 \times 10^4$ and $N = 10^4$, respectively, and with $K = 15$ (dashed lines) and $K = 200$ (solid lines) transmitting antennas. We used the same codes (of length $N = 15 \times 10^4$) for the single-user MIMO system with both $K = 15$ and $K = 200$ due to the complexity of constructing a code of length $N = 2 \times 10^6$ and we believe that these results suffice to demonstrate the difference between the single-user MIMO and the multiple-user MIMO systems in a finite setting. It is clear that the single-user systems outperform the respective multiuser systems in most scenarios. A single-user MIMO system with $K = 200$ reaches BER of 10^{-4} within 0.2 dB from the predicted DE threshold, and within 1 dB from the channel capacity. It is comparable to the excellent performance reported in [9] for a system which includes a sphere decoder, except that our proposed system operates at higher channel load and requires a reduced complexity receiver.

We believe that the 0.1-dB difference in the waterfall region for $K = 200$ in favor of the multiple-access system is related to the fact that $N = 10^4$ in the MAC system and $N = 15 \times 10^4$ in the single-user system. Such difference in the block lengths suggests that the single-user code is slightly better in terms of the threshold value (which is indeed seen in Fig. 12). However, looking at an EXIT analysis as in Section IV-C, we know that since the MUD curve is identical for both single- and multiple-access schemes and since the code of the multiple-access
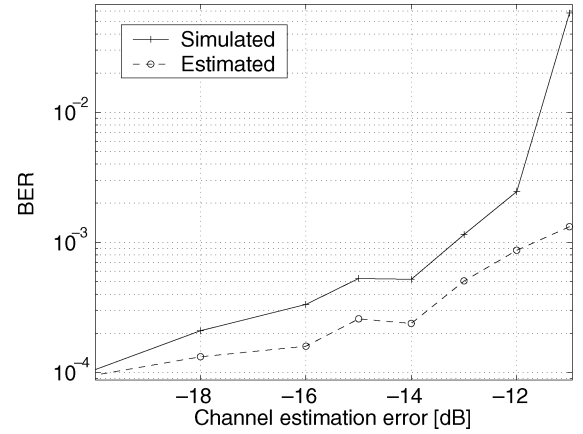
scheme is slightly worse, the weeker LDPC code would give better performance for SNR which is smaller than the threshold. Notice that this also accounts for the slower descent of the multiple-access system when $\frac{E_b}{N_0} > 2.2$ dB, which cannot be considered an error floor.

It is evident that in our setting more users are required to achieve the asymptotic multiple-access MIMO performance than the number of antennas required to approach the ergodic single-user MIMO asymptotic performance [1]. This is a direct consequence of our block-fading model which prevents coding over differently faded symbols, thus causing erroneous blocks for a small number of users (so they are noise limited). In fact, the multiple-access outage capacity is strictly lower than the single-user outage capacity for a finite block length over a block-fading channel. This effect is avoided if the considered model is a fast-fading model (but still the CSI is assumed to be known to the receiver) and not a quasi-static block-fading model.

E. Effect of Coefficients Estimation Error

The preceding simulation results are obtained when the receiver has full knowledge of the channel coefficients. In real communication systems, such perfect knowledge is rare and a more realistic model should consider some imperfections in the estimations of the coefficients. These imperfections can be modeled by an independent white additive noise vector $e_{t,k}$, so that $(y_t, \{h_k + \sqrt{\epsilon}e_{t,k}\}_{k=1}^K)$ is known to the receiver and used instead of the exact value h_k, where $\{e_{t,k}\}_{k=1}^K, \{h_k\}_{k=1}^K \sim \mathcal{N}(0, I)$, ϵ is some constant, and where y still follows (1). The simulation results of this scenario are drawn in Fig. 13, for a MIMO system with $K = 100$ users, $\alpha = 2$, $\gamma = 2$ dB, and the same code used in the previous section with block length $N = 10^4$. Notice that this system is quite sensitive to these estimation errors and less than -10 dB ($\epsilon \mid_{dB} < -10$) is needed for satisfactory performance. This result is understandable since the MUDs are not designed for such estimation errors and by including such imperfections in the asymptotic analysis one might improve its resilience. Nevertheless, if the additive estimation

error $\sqrt{\frac{\gamma\epsilon}{M}}\boldsymbol{e_k}$ is modeled as an equivalent decrease in the signal power $\gamma \leftarrow \frac{\gamma}{1+\epsilon\gamma}$, simulations with perfect coefficients' knowledge can be performed which would reflect the resilience of designs made for lower thresholds against these imperfections (The thresholds are expressed in terms of the power γ). Results of such simulations are also plotted in Fig. 13. It is clear that as long as the estimation errors are low, the two models agree, but higher estimation errors seem to cause a substantial difference between them, due to the use of an unaware LMMSE in the receiver.

VI. CONCLUSION

In this paper, we described an efficient MIMO communication scheme for both single and multiple users employing LDPC codes, LMMSE-based MUD, and a fully iterative receiver. An asymptotic analysis, which is a variation of the one in [14], was used. The LDPC code was optimized for this unique setting, achieving, at a channel load of 2, a large improvement over good LDPC codes optimized for the AWGN single-user channel, revealing the necessity of the modified search. Further EXIT analysis explained the severe degradation due to the use of AWGN-good codes. The necessity of optimizing the code is established, provided that other methods of matching such as user classes and/or power profiling are not used [28], [27]. The resulting low-complexity scheme performs well for a large number of users, and for channel load of $\alpha = 2$, the asymptotic threshold of the system is only 0.77 dB away from the corresponding channel capacity (evaluated with the replica method). Capacity computations showed the loss due to the QPSK restriction is negligible for the examined conditions (channel load $\in [0, 2]$ and SNR $\in [1, 4]$ dB). The lower complexity unconditional LMMSE scheme is compared to both conditional LMMSE and IO-MUD schemes, for the assessment of the performance loss of using a linear filter, computed only once per iteration. This loss turns out to be less than 0.6 dB with respect to the very complex IO-MUD and only 0.1 dB with respect to the conditional LMMSE. The asymptotic analysis is verified by simulations of finite systems which demonstrate the expected differences between single- and multiple-user systems and between fading MIMO and nonfading CDMA. It is seen that the single-user system does well enough with only 15 antennas whereas the multiuser system requires considerably more. These differences stem from the nonergodic nature of the block-fading channel with a finite number of antennas and can be reduced by using more than one transmitting antenna for each user, by using several frequencies (in wide-band communications), or by adopting faster fading model, so that each decoder experiences several attenuations in a single block. Future work may include a generalization of the model and the code optimization to the cases where no synchronization can be assumed and where several classes of users exist. Another interesting aspect is when some feedback channel exists between the receiver and the transmitter. This feedback can substantially improve the overall achievable spectral efficiency of such system by some power profiling technique. Broader complexity–performance tradeoffs can be achieved by incorporating more than one kind of MUD. Such scheme can include a sphere decoder and an LMMSE or an SUMF.

APPENDIX
ON EXIT CHARTS OF RANDOM CODES

We desire to transmit information $\boldsymbol{U}$. We do so in the standard way by producing a codeword $\boldsymbol{x}$, a vector of N channel symbols x, belonging to an asymptotically long random code X of rate R. We transmit $\boldsymbol{x}$ over a memoryless channel producing the vector $\boldsymbol{y}$.

We are interested in a symbol x_j, that is, a symbol at the jth place in $\boldsymbol{x}$. We define $\boldsymbol{x}_{[j]}$ as $\boldsymbol{x}$ with x_j excluded and correspondingly $\boldsymbol{y}_{[j]}$ as $\boldsymbol{y}$ with y_j excluded. We define X′ as a code obtained from X with the symbol x_j excluded from each codeword. The extrinsic information is then defined as $I_{[j]} \triangleq I(x_j; \boldsymbol{y}_{[j]})$. To streamline the presentation we use binary x_j.

When the code rate R is below the capacity C we have perfect decoding for asymptotically long random code, even if the single symbol y_j is removed (erased) before the decoding, thus we have $I_{[j]} = 1$.

When the code rate R is at or above capacity it is well known [37] that

$$NH^M \geq H(\boldsymbol{y}) \geq NH^M - N\epsilon \qquad (18)$$

where H^M is the entropy of a single received symbol. Now

$$I_{[j]} = I(x_j; \boldsymbol{y}_{[j]}) = H(\boldsymbol{y}_{[j]}) - H(\boldsymbol{y}_{[j]} \mid x_j). \qquad (19)$$

Following (18), for random codes $H(\boldsymbol{y}_{[j]}) \cong (N-1)H^M$ because $\boldsymbol{y}_{[j]}$ is the outcome of transmitting codewords selected uniformly from the code X′, which is of a rate similar to that of X for large N. Also $H(\boldsymbol{y}_{[j]} \mid x_j) \cong (N-1)H^M$ because now $\boldsymbol{y}_{[j]}$, when x_j is given, is the outcome of transmitting codewords selected uniformly from a subcode of X′, comprising approximately half the codewords. Thus, $I_{[j]}$ is a difference between very similar values. To upper-bound it, we shall construct a tree of subcodes of X′ as follows: X′ is the root of the tree, that is, the single code at level 0 of the tree. Level 1 of the tree comprises X_0^1 and X_1^1 which are subcodes of X′ associated with $x_j = 0$ and $x_j = 1$, correspondingly. Level j of the tree is generated by randomly partitioning each code in the $j-1$ level into two codes with an equal number of codewords. The relation between the partitioned code and the pair of the resulting subcodes is denoted parent–child. Enough levels of the tree will be generated until the last $j = K$ level contains codes of rate equal to capacity. Clearly, all the codes on the tree are random codes. Moving down one level in the tree adds one information bit to the codes, thus slightly increasing their rate. If the rate of X′ is (even slightly) above C, then K will be proportional to $N-1$.

We shall define $H^X \triangleq H(\boldsymbol{y})$ as the entropy of $\boldsymbol{y}$ when X is used. Each parent code X^a can be obtained from its child subcodes X^b and X^c by combining. By the convexity of entropy, we have $H^{X^a} \geq 0.5(H^{X^b} + H^{X^c})$, thus, the average of the entropies $H^{X_j^i}$ in each level is increased or unchanged when the level in the tree is decreased. We shall denote this average entropy in level j as H^j. The sequence $H^K, \ldots, H^1, H^0$ is monotonically increasing. Using (18) we have

$$(N-1)H^M - (N-1)\epsilon \leq H^K \leq H^0 \leq (N-1)H^M.$$

Thus, the differences between all adjacent terms in the last sequence are nonnegative and their sum is upper-bounded by $(N-1)\epsilon$.

Since the number of terms in the sequence of differences is proportional to $N-1$, their average is bounded by $A\epsilon$ where A is the finite constant of proportion determined by the code rate. Only $1/F$ part of the difference terms can be more than F times the average, thus, for most of codes and for most of code rates, the differences are smaller than $A\epsilon F$, so as small as desired for large N.

The value of $H(\boldsymbol{y}_{[j]}) - H(\boldsymbol{y}_{[j]} \mid x_j)$ in (19) is just the last term in the aforementioned sequence of differences. Its statistical distribution is similar to the previous terms while the sequence of differences tends to be decreasing due to the convexity of entropy and to the increasing similarity of the codes occupying the lower levels of the tree.

So the EXIT chart of most good random codes is a step function, 0 at rates above capacity and 1 at rates below capacity, the 0 value above capacity being a direct result of the 1 above the capacity.

It may be expected that the same conclusions will hold for a wider class of capacity-achieving codes since all capacity-achieving codes must reflect a random-like behavior, as shown, for example, in [38], and since (18) applies to all capacity-achieving codes [37].

ACKNOWLEDGMENT

REFERENCES

[1] E. Telatar, "Capacity of Multi-Antenna Gaussian Channels," Bell-Labs, Lucent Technologies, Tech. Rep., 1999.

[2] A. Goldsmith, S. A. Jafar, N. Jindal, and S. Vishwanath, "Capacity limits of MIMO channels," *IEEE J. Sel. Areas Commun.*, vol. 21, no. 5, pp. 684–702, Jun. 2003.

[3] E. Biglieri, G. Caire, and G. Taricco, "Limiting performance of block-fading channels with multiple antennas," *IEEE Trans. Inf. Theory*, vol. 47, no. 4, May 2001.

[4] S. Verdú and S. Shamai (Shitz), "Spectral efficiency of CDMA with random spreading," *IEEE Trans. Inf. Theory*, vol. 45, no. 2, pp. 622–640, Mar. 1999.

[5] T. Tanaka, "A statistical-mechanics approach to large-system analysis of CDMA multiuser detectors," *IEEE Trans. Inf. Theory*, vol. 48, no. 11, pp. 2888–2910, Nov. 2002.

[6] ——, "Performance analysis of optimum multiuser detector under phase mismatch," in *Proc. IEEE Int. Symp. Information Theory (ISIT2004)*, Chicago, IL, Jun/Jul. 2004, p. 317.

[7] D. Gesbert, M. Shafi, D. Shiu, P. J. Smith, and A. Naguib, "From theory to practice an overview of MIMO space-time coded wireless systems," *IEEE J. Sel. Areas Commun.*, vol. 21, no. 3, pp. 281–302, Apr. 2003.

[8] P. W. Wolniansky, G. J. Foschini, G. D. Golden, and R. A. Valenzuela, "V-BLAST: An architecture for realizing very high data rates over the rich-scattering wireless channel," in *Proc. Int. Symp. Signals, Systems and Electronics (ISSSE-98)*, Pisa, Italy, Sep. 1998, pp. 295–300.

[9] B. M. Hochwald and S. ten Brink, "Achieving near-capacity on a multiple-antenna channel," *IEEE Trans. Commun.*, vol. 51, no. 3, pp. 389–399, Mar. 2003.

[10] L. Ping, K. Y. Wu, L. Liu, and W. Leung, "A simple unified approach to nearly optimal multiuser detection and space-time coding," in *Proc. Information Theory Workshop (ITW2002)*, Bangalore, India, Oct. 20–25, 2002, pp. 53–56.

[11] N. Seshadri, V. Tarokh, and A. R. Calderbank, "Space time codes for high data rate wireless communication performance criterion and code construction," *IEEE Trans. Inf. Theory*, vol. 44, no. 2, pp. 744–765, Mar. 1998.

[12] S. M. Alamouti, "A simple transmit diversity technique for wireless communications," *IEEE J. Sel. Areas Commun.*, vol. 16, no. 8, pp. 1451–1458, Oct. 1998.

[13] B. Hassibi and B. M. Hochwald, "High-rate codes that are linear in space and time," *IEEE Trans. Inf. Theory*, vol. 48, no. 7, pp. 1804–1824, Jul. 2002.

[14] J. Boutros and G. Caire, "Iterative multiuser joint decoding: Unified framework and asymptotic analysis," *IEEE Trans. Inf. Theory*, vol. 48, no. 7, pp. 1772–1793, Jul. 2002.

[15] Z. Qin, K. C. Teh, and E. Gunawan, "Iterative multiuser detection for asynchronous CDMA with concatenated convolutional coding," *IEEE J. Sel. Areas Commun.*, vol. 19, no. 9, pp. 1784–1792, Sep. 2001.

[16] R. R. Müller and W. Gerstacker, "On the capacity loss due to separation of detection and decoding," *IEEE Trans. Inf. Theory*, vol. 50, no. 8, pp. 1769–1778, Aug. 2004.

[17] J. Zhang, E. K. P. Chong, and D. N. C. Tse, "Output MAI distributions of linear MMSE multiuser receivers in DS-CDMA systems," *IEEE Trans. Inf. Theory*, vol. 47, no. 3, pp. 1128–1144, Mar. 2001.

[18] D. N. C. Tse and O. Zeitouni, "Linear multiuser receivers in random environments," *IEEE Trans. Inf. Theory*, vol. 46, no. 1, pp. 171–188, Jan. 2000.

[19] D. N. C. Tse and S. Hanly, "Linear multiuser receivers: Effective interference, effective bandwidth and user capacity," *IEEE Trans. Inf. Theory*, vol. 45, no. 2, pp. 641–657, Mar. 1999.

[20] G. Yue and X. Wang, "Coding-spreading optimization for turbo multiuser detector in LDPC-coded CDMA," in *Proc. 3rd Int. Symp. Turbo Codes and Related Topics*, Brest, France, Sep. 2003, pp. 181–184.

[21] T. Richardson, A. Shokrollahi, and R. Urbanke, "Design of capacity-approaching irregular low-density parity-check codes," *IEEE Trans. Inf. Theory*, vol. 47, no. 2, pp. 619–637, Feb. 2001.

[22] S. Y. Chung, G. D. Forney, T. J. Richardson, and R. Urbanke, "On the design of low-density parity-check codes within 0.0045 dB of the Shannon limit," *IEEE Commun. Lett.*, vol. 5, no. 2, pp. 58–60, Feb. 2001.

[23] J. Hou, P. H. Siegel, and L. B. Milstein, "Performance analysis and code optimization of low density parity-check codes on Rayleigh fading," *IEEE J. Sel. Areas Commun.*, vol. 19, no. 5, pp. 924–934, May 2001.

[24] S. ten Brink, G. Kramer, and A. Ashikhmin, "Design of low density parity-check codes for multi-antenna modulation and detection," Bell-Labs, Lucent Technologies, Tech. Memo., 2002.

[25] B. Lu, G. Yue, and X. Wang, "Performance analysis and design optimization of LDPC-coded MIMO OFDM systems," *IEEE Trans. Signal Process.*, vol. 52, no. 2, pp. 348–361, Feb. 2004.

[26] S. Verdú and D. Guo, "Randomly spread CDMA: Asymptotics via statistical physics," *IEEE Trans. Inf. Theory*, submitted for publication.

[27] T. Tanaka, G. Caire, and R. Müller, "Density evolution and power profile optimization for iterative multiuser decoders based on individually optimum multiuser detectors," in *Proc. 40th Annu. Allerton Conf. Communication, Control and Computing*, Monticello, IL, Oct. 2002, pp. 417–427.

[28] G. Caire, S. Guemghar, A. Roumy, and S. Verdú, "Maximizing the spectral efficiency of coded CDMA under successive decoding," *IEEE Trans. Inf. Theory*, vol. 50, no. 1, pp. 152–164, Jan. 2004.

[29] A. Amraoui, S. Dusad, and R. Urbanke, "Achieving general points in the 2-user Gaussian MAC without time-sharing or rate-splitting by means of iterative coding," in *Proc. IEEE Int. Symp. Information Theory*, Lausanne, Switzerland, Jun./Jul. 30–5, 2002, p. 334.

[30] S. Verdú, *Multiuser Detection*. Cambridge, U.K.: Cambridge Univ. Press, 1998.

[31] S. Y. Chung, T. J. Richardson, and R. L. Urbanke, "Analysis of sum-product decoding of low-desity parity-check codes using a Gaussian approximation," *IEEE Trans. Inf. Theory*, vol. 47, no. 2, pp. 657–670, Feb. 2001.

[32] S. Verdú, "The information-theoretic optimality of QPSK," in *Proc. 4th Int. ITG Conf. Source and Channel Coding*, Berlin, Germany, Jan. 28–30, 2002, pp. 167–174.

[33] E. Biglieri, J. Proakis, and S. Shamai (Shitz), "Fading channels: Information-theoretic and communications aspects," *IEEE Trans. Inf. Theory*, vol. 44, no. 6, pp. 2619–2692, Oct. 1998.

[34] R. Storn and K. Price, "Differential evaluation – A simple and efficient heuristic for global optimization over continuous spaces," *J. Global Optimiz.*, vol. 11, pp. 341–359.

[35] F. Brännström, L. K. Rasmussen, and A. Grant, "Optimal scheduling for multiple serially concatenated codes," in *Proc. 3rd Int. Symp. Turbo Codes and Related Topics*, Brest, France, Sep. 2003, pp. 383–386.

[36] I. Sason and R. Urbanke, "Parity-check density versus performance of binary linear block codes over memoryless symmetric channels," *IEEE Trans. Inf. Theory*, vol. 49, no. 7, pp. 1611–1635, Jul. 2003.

[37] T. S. Han and S. Verdú, "Approximation theory of output statistics," *IEEE Trans. Inform. Theory*, vol. 39, no. 3, pp. 752–772, May 1993.

[38] S. Shamai (Shitz) and S. Verdú, "The empirical distribution of good codes," *IEEE Trans. Inf. Theory*, vol. 43, no. 3, pp. 836–846, May 1997.