Iterative Multiuser Joint Decoding: Unified Framework and Asymptotic Analysis

Joseph Boutros, *Member, IEEE*, and Giuseppe Caire, *Member, IEEE*

***Abstract*—We present a framework for iterative multiuser joint decoding of code-division multiple-access (CDMA) signals, based on the *factor-graph* representation and on the *sum-product* algorithm. In this framework, known parallel and serial, hard and soft interference cancellation algorithms are derived in a unified way. The asymptotic performance of these algorithms in the limit of large code block length can be rigorously analyzed by using *density evolution*. We show that, for random spreading in the *large-system* limit, density evolution is considerably simplified. Moreover, by making a Gaussian approximation of the decoder soft output, we show that the behavior of iterative multiuser joint decoding is approximately characterized by the stable fixed points of a simple one-dimensional nonlinear dynamical system.**

***Index Terms*—Density evolution, interference cancellation, iterative decoding, multiuser detection (MUD).**

I. Introduction

MULTIUSER detection (MUD) has been traditionally regarded as an ensemble of techniques to detect uncoded data in a multiple-access waveform channel (see [1] and references therein).

More recently, research has been focused on the combination and interaction of MUD and channel coding. From an information-theoretic point of view, all points in the capacity region of the Gaussian multiple-access channel are achievable by successive single-user decoding and interference cancellation (IC) [2], [3]. In correlated waveform channels, such as code division multiple access (CDMA), the optimal successive IC decoder takes on the structure of a decision-feedback minimum mean-square error (MMSE) detector, where at each stage the already decoded users are subtracted from the received signal [4].

Information-theoretic results hinge upon the existence of optimal (i.e., capacity-achieving) user codes. A different and perhaps more practical approach to multiuser joint decoding considers a given class of finite-complexity channel codes and investigates the achievable spectral efficiency at given target bit-error rate (BER). The number of works in this direction is overwhelming. Without even trying to be exhaustive, we refer to [5]–[21] and references therein. In these works, joint maximum-likelihood (ML) decoding and some low-complexity iterative IC-based approximations have been investigated, several different algorithms have been proposed, and performance has been evaluated mainly via computer simulation.

This paper contributes to the above stream of work in two ways. First, we provide a unified framework where a large class of previously proposed algorithms can be elegantly derived. Second, we provide an asymptotic performance analysis of iterative IC decoding enabling the quantitative and qualitative evaluation of systems for which simulation would be just impossible.

Our framework is based on the application of the *sum-product algorithm* [22] to the factor-graph representation [22]–[25] of the *a posteriori* joint probability mass function (pmf) of the users information bits. The factor-graph for the problem at hand has cycles if the number of users is larger than one. Therefore, the resulting algorithms are intrinsically iterative. Depending on the algorithm execution scheduling we obtain classical parallel and serial iterative decoding as special cases. By making some simple approximations of the *a posteriori* pmf at the decoders output, we obtain in a direct way several previously proposed iterative IC decoding algorithms, which were motivated mainly by heuristics.

Although formally similar, the algorithms derived from the sum-product algorithm differ by a fundamental detail with respect to their heuristic counterparts: the estimated interference at each iteration is a function of the decoders *extrinsic pmf* (in brief, EXT),[1] rather than of the decoders *a posteriori pmf* (*a posteriori* probability (APP)), as in [5], [7], [9], [12], [19]–[21]. This has two important consequences: from the performance viewpoint, we show that EXT-based iterative algorithms outperform their APP-based counterparts in terms of maximum achievable spectral efficiency at given target BER. From the analysis viewpoint, EXT-based iterative algorithms can be rigorously analyzed by the general technique known as *Density Evolution* (DE) [27], developed to analyze the performance of various families of random-like code ensembles under message-passing decoding (see [28] and references therein).

For the class of iterative multiuser IC decoders derived in this paper and for a class of user channel codes having certain symmetry properties we prove a rigorous concentration result [27] showing that, for large block length, the decoder BER performance converges almost surely to the limit calculated by DE. We show that DE is considerably simplified for random CDMA in the *large-system* regime, i.e., when the number of users and the spreading factor grow without bound while their ratio stays

Manuscript received August 11, 2000; revised December 22, 2001. The material in this paper was presented in part at the IEEE International Symposium on Information Theory, Washington, DC, June 24—29, 2001.

J. Boutros is with Ecole Nationale Supérieure des Télécommunications de Paris, 75634 Paris Cedex, France (e-mail: boutros@com.enst.fr).

G. Caire is with the Institut Eurecom, 06904 Sophia-Antipolis Cedex, France (e-mail: caire@eurecom.fr).

Communicated by F. R. Kschischang, Associate Editor for Coding Theory.

Publisher Item Identifier S 0018-9448(02)05150-7.

[1]In the parlance of "Turbo Coding" literature, this is commonly referred to as decoder "extrinsic information" [26].

constant [29]–[31]. Finally, we provide a *Gaussian approximation* [32]–[34] of the exact DE that yields accurate results and allows the characterization of the decoder performance limits in terms of the stable fixed points of a simple one-dimensional nonlinear dynamical system.

The paper is organized as follows. Section II presents the system model. Section III deals with the factor-graph and the sum-product algorithm and derives iterative joint decoding schemes. In Section III-C, we obtain low-complexity IC decoders. Section IV is dedicated to the asymptotic analysis of IC decoders. Numerical examples are presented in Section V and in Section VI, we point out some recent results originated by this work and some suggestions for further work. All proofs are given in Appendix A, and Appendix B gives the details of the calculation of the Gaussian approximation of DE for the important case of binary convolutional codes and quaternary phase-shift keying (QPSK) modulation.

Notation Definitions: For a matrix $\boldsymbol{A}$, $\boldsymbol{a}^i$, $\boldsymbol{a}_j$, and $a_{i,j}$ denote its ith row, jth column, and (i, j) element, respectively. For a row or column vector $\boldsymbol{v}$, v_i denotes its ith component. $1\{\mathcal{E}\}$ denotes the indicator function of the event $\mathcal{E}$. $\mathcal{N}_{\mathbb{C}}(\boldsymbol{\mu}, \boldsymbol{R})$ denotes the circularly symmetric complex multivariate Gaussian distribution [35] with mean vector $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{R}$. The symbol $\sim$ means "distributed as."

II. Synchronous CDMA System Model

We consider the "canonical" CDMA discrete-time synchronous system (see [1], [29], [31] and references therein) described by

$$\boldsymbol{y}_n = \boldsymbol{S}\boldsymbol{W}_n\boldsymbol{x}_n + \boldsymbol{\nu}_n, \qquad n = 1, \ldots, N \tag{1}$$

where $\boldsymbol{y}_n \in \mathbb{C}^L$ is the signal vector received at time n, $\boldsymbol{\nu}_n \sim \mathcal{N}_{\mathbb{C}}(\boldsymbol{0}, \boldsymbol{I})$ is the corresponding noise vector, $\boldsymbol{x}_n \in \mathbb{C}^K$ is the vector of transmitted user modulation symbols at time n, $\boldsymbol{S} = [\boldsymbol{s}_1, \ldots, \boldsymbol{s}_K] \in \mathbb{C}^{L\times K}$ is the *spreading matrix*, containing the user spreading sequences by columns, and $\boldsymbol{W}_n = \mathrm{diag}(w_{1,n}, \ldots, w_{K,n})$ is a matrix of complex amplitudes. The integers K, L, and N denote the number of users, the *spreading factor* (chips per symbol), and the code block length (symbols per block), respectively.

Users send independently encoded information. We assume that the user codewords are aligned in time. Hence, (1) describes the channel during the transmission of one codeword. User codewords are interleaved before transmission. We let $\mathcal{C}_k$ and $\mathcal{A}_k \subset \mathbb{C}$ denote the code and the complex modulation signal set of user k, respectively, and we let $\phi_k\colon \mathbb{F}_2^{m_k} \to \mathcal{A}_k^N$ denote the encoding function for $\mathcal{C}_k$ (including interleaving). The coding rate is $R_k = m_k/N$ bits per symbol. We have

$$\mathcal{C}_k = \left\{\boldsymbol{x} \in \mathbb{C}^N\colon \boldsymbol{x} = \phi_k(\boldsymbol{b}), \qquad \forall \boldsymbol{b} \in \mathbb{F}_2^{m_k}\right\}. \tag{2}$$

Since we put no restrictions on the encoding function ϕ_k, any code over the complex signal set $\mathcal{A}_k$ can be expressed by (2).

The user complex amplitudes are given by $w_{k,n} = \sqrt{\gamma_k}e^{j\theta_{k,n}}$, where the $\theta_{k,n}$'s are random independent and identically distributed (i.i.d.) phases whose distribution is such that $E[e^{j\theta_{k,n}}] = 0$. Phase randomization is artificially introduced by the transmitters in order to make *multiple access interference* (MAI) circularly symmetric with respect to any user.[2] The receiver knows perfectly the CDMA system parameters $\boldsymbol{S}$, $\{\gamma_1, \ldots, \gamma_K\}$, $\boldsymbol{\Theta} \triangleq \{\theta_{k,n}\colon k = 1, \ldots, K, n = 1, \ldots, N\}$, and the user codes $\{\mathcal{C}_1, \ldots, \mathcal{C}_K\}$. In practice, the spreading sequences $\boldsymbol{s}_k$ and the phases $\boldsymbol{\Theta}$ are pseudorandomly generated according to some known algorithm and the γ_k's are imposed by the power control algorithm.

The modulation signal sets $\mathcal{A}_k$ have zero mean and unit average energy, i.e.,

$$\frac{1}{|\mathcal{A}_k|}\sum_{a\in\mathcal{A}_k} a = 0$$

and

$$\frac{1}{|\mathcal{A}_k|}\sum_{a\in\mathcal{A}_k} |a|^2 = 1.$$

Therefore, the average signal-to-noise ratio (SNR) of user k is given by

$$\mathsf{SNR}_k \triangleq \frac{E\left[|\boldsymbol{s}_k|^2|w_{k,n}x_{k,n}|^2\right]}{E\left[|\boldsymbol{s}_k^H\boldsymbol{\nu}_n|^2\right]} = \gamma_k|\boldsymbol{s}_k|^2. \tag{3}$$

Assuming that an ideal Nyquist pulse with zero excess bandwidth [37] is used for modulating the chips, the system spectral efficiency in bits per second per hertz (bit/s/Hz) is given by [31]

$$\rho = \frac{1}{L}\sum_{k=1}^{K} R_k. \tag{4}$$

For an equal-rate system ($R_k = R$ for all users), we get $\rho = \alpha R$, where $\alpha = K/L$ is the number of "users per chip," referred to as the *channel load* [29].

III. Joint Decoding: Graph Representation and Iterative Algorithms

In this section we derive the *factor-graph* representation [22], [23], [25] of the *a posteriori* joint pmf of the user information bits. By applying the *sum-product* algorithm to the resulting factor-graph, we obtain a class of iterative decoding algorithms approximating optimal maximum *a posteriori* (MAP) decoding. In the particular case of parallel and serial *scheduling*, we obtain the algorithms proposed in several works (see, for example, [7], [15], [16], [20]). Finally, by making some simplifications, we derive in a unified and direct way well-known parallel and serial iterative IC decoding algorithms [5], [8], [10], [12], [13], [17], [19]–[21].

A. Factor-Graph Representation

Throughout the paper, we use the proportionality symbol $\propto$ in order to indicate that the quantity in the right-hand side (RHS) is defined up to a multiplicative factor chosen in order to make it a true probability density (or mass) function.

Let $\boldsymbol{b}^k = (b_{k,1}, \ldots, b_{k,m_k}) \in \mathbb{F}_2^{m_k}$ be the vector of information bits of user k, and let $\boldsymbol{Y} = [\boldsymbol{y}_1, \ldots, \boldsymbol{y}_N] \in \mathbb{C}^{L\times N}$ be the received signal and $\boldsymbol{X} \in \mathbb{C}^{K\times N}$ the transmitted code array obtained by stacking by rows the transmitted codewords $\boldsymbol{x}^1, \ldots, \boldsymbol{x}^K$ (the transmitted symbol vector $\boldsymbol{x}_n$ is the nth column of $\boldsymbol{X}$). The multiuser channel with coding is fully de-

[2] In actual CDMA implementations, phase randomization is common practice [36]. In our context, it is mathematically convenient for the analysis of Section IV.

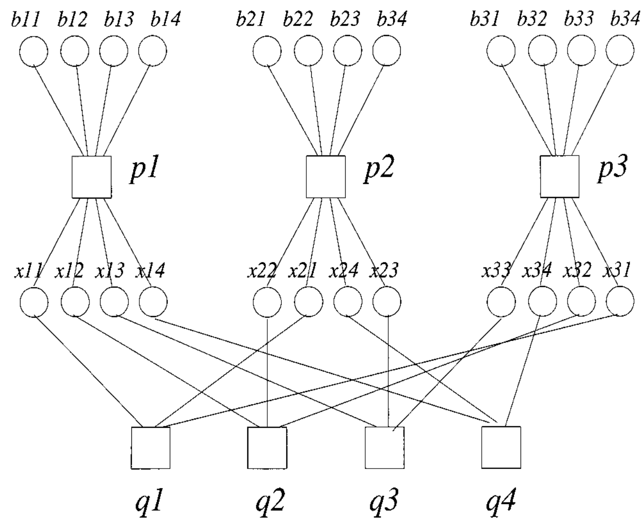

Fig. 1. An instance of the factor-graph for a multiuser coded CDMA system with $K = 3$ users and block length $N = 4$.

scribed by the *a posteriori* pmf of $\{\boldsymbol{b}^1, \ldots, \boldsymbol{b}^K\}$ given $\boldsymbol{Y}$, denoted by $\Omega(\boldsymbol{b}^1, \ldots, \boldsymbol{b}^K|\boldsymbol{Y})$.[3] By using the fact that the vector channel (1) is memoryless, that users codewords are independently generated and that the user information bits have uniform *a priori* probability, we can write

$$\Omega(\boldsymbol{b}^1, \ldots, \boldsymbol{b}^K|\boldsymbol{Y}) \propto \prod_{n=1}^{N} q_n(\boldsymbol{x}_n) \prod_{k=1}^{K} p_k(\boldsymbol{x}^k, \boldsymbol{b}^k) \tag{5}$$

where we define the *code constraint functions*

$$p_k(\boldsymbol{x}, \boldsymbol{b}) \stackrel{\Delta}{=} 1\{\boldsymbol{x} = \phi_k(\boldsymbol{b})\} \tag{6}$$

and the *channel transition functions*

$$q_n(\boldsymbol{x}) \stackrel{\Delta}{=} \exp\left(-|\boldsymbol{y}_n - \boldsymbol{S}\boldsymbol{W}_n\boldsymbol{x}|^2\right). \tag{7}$$

The factor-graph for $\Omega(\boldsymbol{b}^1, \ldots, \boldsymbol{b}^K|\boldsymbol{Y})$ is a representation of the factorization (5) via a bipartite graph $G(\mathcal{V}, \mathcal{F})$, where variables ($b_{k,j}$ and $x_{k,n}$) are represented by "variable nodes" $v \in \mathcal{V}$ and functions (q_k and p_n) by "function nodes" $f \in \mathcal{F}$. Each node f is connected to all nodes v for which the corresponding variables are arguments of the function f. Fig. 1 shows the factor-graph for $\Omega(\boldsymbol{b}^1, \ldots, \boldsymbol{b}^K|\boldsymbol{Y})$ for $K = 3$, $N = 4$, and $m_1 = m_2 = m_3 = 4$.

Remark 1: Each code constraint function block can also be represented as a factor-graph, depending on the code structure. In particular, if code $\mathcal{C}_k$ is a trellis code, turbo code [38], low-density parity check (LDPC) code (see [28] and references therein), repeat–accumulate (RA) code [39], etc., the subgraph formed by the variable nodes $b_{k,1}, \ldots, b_{k,m_k}$, $x_{k,1}, \ldots, x_{k,N}$, and by the function node p_k can be expanded in well-known forms [22]. For example, if all codes in the factor-graph of Fig. 1 are trellis codes of rate 1 bit/symbol, the function nodes p_1, p_2, and p_3 in Fig. 1 can be replaced by the corresponding factor subgraph given in Fig. 2. However, in order to keep our treatment as general as possible, we shall not expand further the factor-graph of Fig. 1. $\diamond$

[3]Conditioning with respect to $\{\boldsymbol{W}_1, \ldots, \boldsymbol{W}_N\}$ and $\boldsymbol{S}$, which are known to the receiver, is omitted for the sake of notational simplicity.

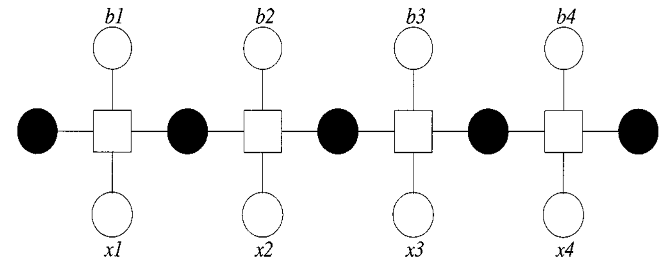

Fig. 2. Factor-graph of a trellis-terminated trellis code of rate 1 bit/symbol and block length $N = 4$. Black circles denote *state variables* and the squares denote the *trellis transition functions* (see [22]).

B. The Sum-Product Algorithm

Let $b_{k,j}$ denote the jth information bit of user k. The optimal MAP detection rule minimizing the average BER for each user is given by

$$\hat{b}_{k,j} = \arg\max_{b \in \mathbb{F}_2} \mathrm{APP}_{k,j}(b), \qquad \text{for all } k, j \tag{8}$$

where $\mathrm{APP}_{k,j}(b)$ denotes the marginal APP of bit $b_{k,j}$, given by

$$\mathrm{APP}_{k,j}(b) = \sum_{\substack{\boldsymbol{b}^1, \ldots, \boldsymbol{b}^K \\ b_{k,j}=b}} \Omega(\boldsymbol{b}^1, \ldots, \boldsymbol{b}^K|\boldsymbol{Y}). \tag{9}$$

Computing the APP of the information bits by brute force has complexity of the order of $\prod_{k=1}^{K} |\mathcal{C}_k|$. Even for small K, this is intractable for practical user code sizes. Also in the case where all user codes are trellis codes, the complexity of joint decoding applied to the Cartesian-product trellis of $\mathcal{C}_1 \times \cdots \times \mathcal{C}_K$ is prohibitive in practice [6], [11].

A general method for approximating (9) consists of applying the sum-product algorithm [22] to the factor-graph. In the sum-product algorithm, adjacent nodes in the factor-graph exchange "messages" in the form of real-valued functions. If $v \in \mathcal{V}$ and $f \in \mathcal{F}$ are connected by the edge $e = (v, f)$, the messages passed along e in either directions are functions of v.

We let $Q_{k,n}(a)$ and $P_{k,n}(a)$ denote the messages calculated at the nth-channel transition function node q_n and at the kth-code constraint function node p_k, respectively, and sent to the variable node $x_{k,n}$, where a denotes a dummy variable taking on values in the symbol alphabet $\mathcal{A}_k$. It is understood that both $Q_{k,n}(a)$ and $P_{k,n}(a)$ are pmf defined over $\mathcal{A}_k$. By following the general sum-product computation rules of [22], we obtain the following.

- Computation at the channel transition function nodes:

$$Q_{k,n}(a) \propto \sum_{\substack{\boldsymbol{a} \in \mathcal{A}_1 \times \cdots \times \mathcal{A}_K \\ a_k = a}} \exp\left(-\left|\boldsymbol{y}_n - \sum_{j=1}^{K} \boldsymbol{s}_j w_{j,n} a_j\right|^2\right) \cdot \prod_{j \neq k} P_{j,n}(a_j), \qquad \text{for } a \in \mathcal{A}_k. \tag{10}$$

- Computation at the code constraint function nodes:

$$P_{k,n}(a) \propto \sum_{\substack{\boldsymbol{a} \in \mathcal{C}_k \\ a_n = a}} \prod_{j \neq n} Q_{k,j}(a_j), \qquad \text{for } a \in \mathcal{A}_k. \tag{11}$$

- Computation at the variable nodes $x_{k,n}$: since these nodes have degree 2, they act as relays, i.e., $x_{k,n}$ sends $Q_{k,n}(a)$ to p_n and $P_{k,n}(a)$ to q_n.
- Computation at the variable nodes $b_{k,j}$: since these nodes have degree 1, they just receive a message from p_k. This is given by

$$\widetilde{\mathrm{APP}}_{k,j}(b) \propto \sum_{\substack{\boldsymbol{a}=\phi_k(\boldsymbol{b})\\ b_{k,j}=b}} \prod_{n=1}^{N} Q_{k,n}(a_n), \quad \text{for } b \in \mathbb{F}_2. \tag{12}$$

The information bit detection is obtained by using the above *approximated* APP in (8).

Remark 2 (Extrinsic pmfs, Soft-In–Soft-Out (SISO) Decoding, and Optimal MUD): A fundamental characteristic of the sum-product algorithm is that " …*a message sent from a node along an adjacent edge cannot depend on the message previously received along the same edge*" [27]. As observed in [27], it is precisely this restriction that enables the analysis of the iterative decoder. We notice that $Q_{k,n}(a)$ and $P_{k,n}(a)$ defined in (10) and (11) fulfill the above requirement.

The quantity defined in (11) is the "extrinsic pmf" (EXT) of the decoder [26], [38], [40]. The APP of $x_{k,n}$, given the *a priori* marginal pmfs $\{Q_{k,j}(a)\colon\ j = 1, \ldots, N\}$, is given by $\propto Q_{k,n}(a)P_{k,n}(a)$. The calculation of the EXT in (11) is often referred to as *soft-in soft-out* (SISO) decoding [26]. If the user codes admit a trellis representation (e.g., they are trellis-terminated convolutional codes), the SISO computation (11) can be carried out by the forward-backward Bahl–Cocke–Jelinek–Raviv (BCJR) algorithm [41] with linear complexity in N.

The quantity defined in (10) is the EXT for Verdú's optimal symbol-by-symbol multiuser detector [1] with *a priori* marginal pmfs $\{P_{j,n}(a)\colon\ j = 1, \ldots, K\}$. Unfortunately, the CDMA vector channel (1) has no particular structure enabling efficient evaluation of (10). Therefore, the computation of (10) has complexity of the order of $\prod_{k=1}^{K} |\mathcal{A}_k|$, i.e., exponential in K. For large K, the iterative sum-product algorithm is still too complex for practical implementations. ◊

Scheduling: It can be shown that the sum-product algorithm is able to compute exactly the marginals of the underlying multivariate function in a finite number of steps if the corresponding factor-graph is cycle-free (see [22], [24] and reference therein). It is apparent from direct inspection of Fig. 1 that our problem yields a factor-graph with cycles unless $K = 1$ or the users codes are trivial (i.e., $\mathcal{C}_k = \mathcal{A}_k^N$ for all k). The consequences of cycles are: 1) detection based on (12) is in general suboptimal; 2) the result is sensitive to the order in which computation is carried out through the nodes (scheduling); 3) different scheduling yields generally nonequivalent iterative algorithms.

A scheduling is defined by a sequence of node subsets to be activated. Nodes in the same subset can be activated in any arbitrary order, as the value of their output messages does not depend on the activation order within the subset. When a node is activated, all its output messages are calculated by using the current value of its input messages.

In our case, the simplest and most intuitive schedulings are *parallel* and *serial*. In parallel scheduling, one iteration is given by the sequence

$$\begin{aligned}&\{q_n\colon n = 1, \ldots, N\}\\ &\quad\to \{x_{k,n}\colon k = 1, \ldots, K,\ n = 1, \ldots, N\}\\ &\quad\to \{p_k\colon k = 1, \ldots, K\}\\ &\quad\to \{x_{k,n}\colon k = 1, \ldots, K,\ n = 1, \ldots, N\}.\end{aligned} \tag{13}$$

In serial scheduling, users are considered in a given cyclic order (without loss of generality, we consider the natural ordering $k = 1, 2, \ldots, K$). One iteration is given by the sequence

$$\begin{aligned}&\{q_n\colon n = 1, \ldots, N\}\\ &\quad\to \{x_{1,n}\colon n = 1, \ldots, N\} \to \{p_1\}\\ &\quad\to \{x_{1,n}\colon n = 1, \ldots, N\} \to \{q_n\colon n = 1, \ldots, N\}\\ &\quad\to \{x_{2,n}\colon n = 1, \ldots, N\} \to \{p_2\}\\ &\quad\to \{x_{2,n}\colon n = 1, \ldots, N\}\\ &\qquad\qquad\vdots\\ &\quad\to \{q_n\colon n = 1, \ldots, N\} \to \{x_{K,n}\colon n = 1, \ldots, N\}\\ &\quad\to \{p_K\} \to \{x_{K,n}\colon n = 1, \ldots, N\}.\end{aligned} \tag{14}$$

In both cases, the algorithm is initialized by the uniform pmf $P_{k,n}(a) = 1/|\mathcal{A}_k|$ for all $a \in \mathcal{A}_k$, $k = 1, \ldots, K$ and $n = 1, \ldots, N$.

C. Low-Complexity Approximations: IC Schemes

We notice that (10) consists of computing the *a posteriori* pmf of $x_{k,n}$ given the observation $\boldsymbol{y}_n$, assuming that the interfering symbols $x_{j,n}$, $j \neq k$ are statistically independent with marginal pmf $P_{j,n}(a)$. The exponential complexity of (10) is due to the fact that the symbols $x_{j,n}$ take on values in the discrete set $\mathcal{A}_j$. By artificially modifying the marginal pmfs of the interfering symbols, several low-complexity algorithms can be derived in a unified way.

Hard IC: By replacing $P_{j,n}(a)$ with its single mass point approximation

$$\tilde{P}_{j,n}(a) = 1\left\{a = \arg\max_{a' \in \mathcal{A}} P_{j,n}(a')\right\} \tag{15}$$

(10) reduces to

$$Q_{k,n}(a) \propto \exp\left(-\beta_{k,n}\,|z_{k,n} - a|^2\right) \tag{16}$$

where

$$\beta_{k,n} = \gamma_k |\boldsymbol{s}_k|^2 \tag{17}$$

and where

$$z_{k,n} = \frac{1}{w_{k,n}|\boldsymbol{s}_k|^2}\,\boldsymbol{s}_k^H\left(\boldsymbol{y}_n - \tilde{\boldsymbol{y}}_{k,n}\right) \tag{18}$$

with

$$\tilde{\boldsymbol{y}}_{k,n} = \sum_{j \neq k} \boldsymbol{s}_j w_{j,n} \tilde{x}_{j,n}$$

and

$$\tilde{x}_{j,n} = \arg\max_{a \in \mathcal{A}_j} P_{j,n}(a).$$

In the following, we refer to the subtraction of the MAI estimate from the received signal followed by filtering (as in (18)) as "IC-MUD." In the case of (18), the MAI estimate $\tilde{\boldsymbol{y}}_{k,n}$ is formed by using the symbol-by-symbol hard decisions based on the SISO decoders EXT pmfs, and the filter is the conventional single-user matched filter (SUMF). Depending on the scheduling, we obtain the well-known serial and parallel hard IC schemes.

The coefficient $\beta_{k,n}$ defined in (17) can be interpreted as the *nominal* signal-to-interference-plus-noise ratio (SINR) at the output of the IC-MUD. In fact, by letting $z_{k,n} = x_{k,n} + \zeta_{k,n}$, if all symbol-by-symbol hard decision were correct (i.e., $\tilde{x}_{j,n} = x_{j,n}$ for all $j \neq k$), then the variance of the residual interference-plus-noise term $\zeta_{k,n}$ would be $E[|\zeta_{k,n}|^2|] = 1/\beta_{k,n}$.

Linear MMSE (LMMSE)-Based Soft IC: By replacing $P_{j,n}(a)$ with a complex circularly symmetric Gaussian probability density function (pdf) with the same mean and variance [9], given by

$$\tilde{P}_{j,n}(a) = \mathcal{N}_{\mathbb{C}}(\tilde{x}_{j,n}, \xi_{j,n}) \tag{19}$$

where

$$\begin{aligned}\tilde{x}_{j,n} &= \sum_{a \in \mathcal{A}_j} a P_{j,n}(a) \\ \xi_{j,n} &= \sum_{a \in \mathcal{A}_j} |a - \tilde{x}_{j,n}|^2 P_{j,n}(a)\end{aligned} \tag{20}$$

the MAI can be treated as a Gaussian vector with covariance matrix

$$\tilde{\boldsymbol{\Sigma}}_{k,n} = \sum_{j \neq k} \boldsymbol{s}_j \boldsymbol{s}_j^H \gamma_j \xi_{j,n} \tag{21}$$

and mean

$$\tilde{\boldsymbol{y}}_{k,n} = \sum_{j \neq k} \boldsymbol{s}_j w_{j,n} \tilde{x}_{j,n}. \tag{22}$$

With the above assumption, $\boldsymbol{y}_n$ and $\{x_{j,n}: j \neq k\}$ are conditionally jointly Gaussian given $x_{k,n} = a$, and (10) corresponds to the marginalization of a multivariate Gaussian pdf. After some algebra, (10) can be put in the form (16) with nominal SINR

$$\beta_{k,n} = \gamma_k \boldsymbol{s}_k^H \tilde{\boldsymbol{\Sigma}}_{k,n}^{-1} \boldsymbol{s}_k \tag{23}$$

and with

$$z_{k,n} = \frac{w_{k,n}^*}{\beta_{k,n}} \boldsymbol{s}_k^H \tilde{\boldsymbol{\Sigma}}_{k,n}^{-1} \left(\boldsymbol{y}_n - \tilde{\boldsymbol{y}}_{k,n}\right). \tag{24}$$

In this case, the IC-MUD output $z_{k,n}$ is obtained by subtracting from $\boldsymbol{y}_n$ the MAI estimate $\tilde{\boldsymbol{y}}_{k,n}$ formed by using the soft symbol-by-symbol estimates (20) based on the SISO decoders EXT pmfs, and the filter is the *nominal* unbiased linear MMSE (LMMSE) filter defined by the vector of filter coefficients

$$\boldsymbol{h}_{k,n} = \frac{w_{k,n}}{\beta_{k,n}} \tilde{\boldsymbol{\Sigma}}_{k,n}^{-1} \boldsymbol{s}_k. \tag{25}$$

Depending on the scheduling, we refer to this algorithm as the LMMSE-based serial or parallel soft IC scheme.

SUMF-Based Soft IC: If, in addition to the above conditional Gaussian assumption, $\boldsymbol{y}_n$ is assumed to be conditionally white with the same MAI+noise total power, the assumed conditional covariance matrix for $\boldsymbol{y}_n$ is

$$\mathrm{diag}\left(\frac{1}{L}\mathrm{tr}(\tilde{\boldsymbol{\Sigma}}_{k,n}), \ldots, \frac{1}{L}\mathrm{tr}(\tilde{\boldsymbol{\Sigma}}_{k,n})\right)$$

and (10) takes on again the form (16) with nominal SINR

$$\beta_{k,n} = \frac{\gamma_k |\boldsymbol{s}_k|^2}{1 + \frac{1}{L} \sum\limits_{j \neq k} \gamma_j |\boldsymbol{s}_j|^2 \xi_{j,n}} \tag{26}$$

and

$$z_{k,n} = \frac{1}{w_k |\boldsymbol{s}_k|^2} \boldsymbol{s}_k^H \left(\boldsymbol{y}_n - \tilde{\boldsymbol{y}}_{k,n}\right) \tag{27}$$

with $\tilde{\boldsymbol{y}}_{k,n}$ given by (22). In this case, the IC-MUD output $z_{k,n}$ is obtained by passing the same difference signal as in (24) through the SUMF. Depending on the scheduling, we refer to this algorithm as the SUMF-based serial or parallel soft IC scheme.

Remark 3 (Using APPs Instead of EXTs): In several papers (e.g., [5], [8], [10], [12], [13], [17], [19]–[21], [42], [43]), $\tilde{x}_{j,n}$ in (20) is calculated from the APP at the SISO output instead of the EXT $P_{j,n}(a)$, thus violating the basic principle of the sum-product algorithm. As a consequence, the residual interference plus noise at the IC-MUD output $\zeta_{k,n} = z_{k,n} - x_{k,n}$ (where $z_{k,n}$ is given by (18), (24) or (27)), is biased conditionally on $x_{k,n}$, i.e., $E[\zeta_{k,n}|x_{k,n}] \neq 0$. Moreover, it can be shown that the conditional bias *reduces* the useful signal component in $z_{k,n}$ [44], [45].

A consequence of this bias effect is that the APP-based IC algorithms cannot be analyzed by simply tracking the *variance* of the residual MAI plus noise through the decoder iterations, as done in in [5], [46], [47]. As a matter of fact, the "variance" analysis applied to APP-based algorithms yields too optimistic results. $\Diamond$

Remark 4 (Relation With Uncoded Iterative IC-MUD): A large number of recent works proposed and analyzed iterative IC-MUD for uncoded CDMA (see [44], [48]–[54] and references therein). In particular, with *linear feedback* these algorithms can be seen as matrix-polynomial approximations of *linear* multiuser detectors (decorrelator and LMMSE [1]) while, with nonlinear feedback, they can be seen as low-complexity approximations of the optimal multiuser detector [1] obtained via the *Expectation Maximization* (EM) or the *Subspace Alternating Generalized EM* (SAGE) approaches [55]. Finally, both linear and nonlinear feedback iterative IC-MUD with parallel or serial scheduling can be derived in a unified way as the iterative solution of a constrained ML detection problem [52].

It is natural to ask whether the iterative multiuser IC decoding algorithms presented in this paper are related to IC-MUD algorithms for uncoded CDMA. A closer look at the factor-graph of Fig. 1 reveals that, despite obvious analogies, these two classes of algorithms are qualitatively very different. In fact, the factor-graph for an uncoded system is obtained by removing the code constraint function nodes from the graph of Fig. 1. The resulting graph is a *forest*, i.e., a collection of mutually disconnected trees $T_n = \{x_{1,n}, \ldots, x_{K,n}, q_n\}$, for $n = 1, \ldots, N$. The sum-product algorithm applied to each tree T_n yields a noniterative algorithm which (essentially) coincides with the optimal symbol-by-symbol multiuser detector. Therefore, no estimated MAI signal can be iteratively produced by the sum-product algorithm applied to the uncoded case.

In conclusions, we claim that while iterative MUD for uncoded CDMA finds its natural explanation in the framework of iterative solution of linear systems of equations [44], [51], of multistage Wiener filtering [50], [53], [54], and of iterative solution of (constrained) ML estimation problems [52], [55], the natural framework for iterative multiuser IC decoding is the factor-graph and sum-product algorithm approach, as illustrated in this paper. ◊

Remark 5 (An Alternative Interpretation of the LMMSE-Based IC Algorithms): In [12], [20], [21], the LMMSE-based parallel and serial schemes are derived from a *conditional* MMSE argument. Namely, $z_{k,n}$ in (24) is obtained as the conditional LMMSE estimate of $x_{k,n}$ given that the symbols $x_{j,n}$ for $j \neq k$ are distributed according to $P_{j,n}(a)$. Indeed, assuming that $\{P_{j,n}(a)\colon j \neq k\}$ are statistically independent of $x_{k,n}$ and of the noise $\boldsymbol{\nu}_n$, $z_{k,n}$ has the structure of a Wiener MMSE linear estimator for $x_{k,n}$ with observation $\boldsymbol{y}_n$, where the observation mean value

$$\tilde{\boldsymbol{y}}_{k,n} = E\left[\boldsymbol{y} | \{P_{j,n}(a)\colon j \neq k\}\right] \tag{28}$$

is subtracted from the observation and the difference is filtered by $\boldsymbol{h}_{k,n}$ given in (25), solution of the conditional unbiased MMSE problem

$$\begin{cases} \min_{\boldsymbol{h}} & E\left[\left|x_{k,n} - \boldsymbol{h}^H(\boldsymbol{y}_n - \tilde{\boldsymbol{y}}_{k,n})\right|^2 \Big| \{P_{j,n}(a)\colon j \neq k\}\right] \\ \text{subject to} & \boldsymbol{h}^H \boldsymbol{s}_k w_{k,n} = 1. \end{cases} \tag{29}$$

For finite N, this interpretation is not exact, since messages $\{P_{j,n}(a)\colon j \neq k\}$ are statistically *dependent* on $x_{k,n}$ and on $\boldsymbol{\nu}_n$ and neither (22) is obtained from (28) nor (25) is the solution of (29). However, under certain assumptions, $\{P_{j,n}(a)\colon j \neq k\}$ are asymptotically statistically independent of $x_{k,n}$ and of $\boldsymbol{\nu}_n$ for $N \to \infty$, and the above "conditional MMSE" interpretation holds (see Remark 8 in Section IV).

Interestingly, if the APPs instead of EXTs are fed back from the SISO decoders (as in [12], [20], [21]) the above interpretation is never exact, even in the limit for large N and random interleaving, since the APP for $x_{j,n}$ depends on $\boldsymbol{y}_n$, which contains $x_{k,n}$ and $\boldsymbol{\nu}_n$, no matter how large N is. ◊

Remark 6 (Implementation Issues): The nominal LMMSE filter (25) must be recalculated for every user, every decoder iteration, and every symbol interval. Therefore, the LMMSE-based algorithms are much more complex than the SUMF-based (or hard-decision based) algorithms. As an alternative, the LMMSE filters can be calculated adaptively, as proposed in [14], [42]. Another approach for low-complexity implementation of the LMMSE-based algorithms is proposed in [56], where the term $\xi_{n,k}$ in (20) is replaced by its ensemble average

$$E[\xi_{k,n}] = 1 - E[|\tilde{x}_{k,n}|^2]$$

which can be approximated by

$$E[\xi_{k,n}] \approx 1 - \frac{1}{N}\sum_{n=1}^{N} |\tilde{x}_{k,n}|^2$$

and can be easily computed from the kth SISO decoder output. In this way, the LMMSE filter (25) must be recalculated for every user and every iteration, but its is constant for all symbol intervals. This heuristic simplification of LMMSE-based algorithms is referred to as *unconditional LMMSE* in [56], since the resulting filter can be interpreted as the solution of the unconditional unbiased MMSE problem

$$\begin{cases} \min_{\boldsymbol{h}} & E\left[\left|x_{k,n} - \boldsymbol{h}^H(\boldsymbol{y}_n - \tilde{\boldsymbol{y}}_{k,n})\right|^2\right] \\ \text{subject to} & \boldsymbol{h}^H \boldsymbol{s}_k w_{k,n} = 1 \end{cases} \tag{30}$$

under the assumption that $\{P_{j,n}(a)\colon j \neq k\}$ are statistically independent of $x_{k,n}$ and of $\boldsymbol{\nu}_n$, which holds asymptotically for $N \to \infty$.

For practically relevant system size (say, $\min\{K, L\} \geq 10$), the LMMSE-based algorithm previously derived is prohibitively complex, while its unconditional LMMSE approximation is easily implementable. ◊

IV. Asymptotic Performance Analysis for M-PSK Trellis Codes

In this section, we provide a rigorous BER analysis (asymptotically for large block length N) of the iterative multiuser IC decoders derived in Section III-C in the special case where the user codes are derived from the same trellis code over M-PSK. Our analysis is based on the DE method, introduced in [27] to study LDPC codes on binary-input symmetric-output memoryless channels, and in [57] for LDPC codes on binary-input time-invariant intersymbol interference (ISI) Gaussian channels with finite memory.

The main principle underlying DE is that, as $N \to \infty$, the pdf of the messages generated by the iterative decoder after any fixed number of iterations ℓ *concentrates* around the expected pdf resulting from a random cycle-free graph. Furthermore, we show that for random spreading sequences and in the limit for large K and L, with fixed channel load $K/L = \alpha$, DE for parallel scheduling is greatly simplified and, for this case, we provide an accurate and computationally very efficient *Gaussian approximation* (GA) of the exact DE (GA-DE).

A. Assumptions and Definitions

User Codes and Phase Randomization: We assume that all user codes $\mathcal{C}_k$ are derived by the same "basic" code $\mathcal{C}$, and differ only by the interleaver, that is randomly and independently chosen with uniform probability in the set of all permutations of $\{1, \ldots, N\}$. $\mathcal{C}$ is a block code obtained by trellis termination of a time-invariant trellis code with s code symbols (encoder outputs) per trellis section. We assume that $\mathcal{C}$ is a geometrically uniform code [58] over the M-PSK signal set $\mathcal{A}_M \triangleq \{e^{j2\pi m/M}\colon m = 0, \ldots, M-1\}$ with $M > 2$, obtained from a linear trellis code over the ring $\mathbb{Z}_M$ via the natural mapping

μ: $\mathbb{Z}_M \to \mathcal{A}_M$ such that $\mu(m) = e^{j2\pi m/M}$. Furthermore, we assume that the phase randomization sequences

$$\boldsymbol{\Theta} = \{\theta_{k,n}: k = 1, \ldots, K, n = 1, \ldots, N\}$$

have i.i.d. elements uniformly distributed on $\{2\pi m/M: m = 0, \ldots, M-1\}$. We need the following.

Definition 1: The additive-noise memoryless channel $y = x + \nu$ is said to be $\mathcal{A}_M$-invariant if $p_{a\nu}(z) = p_\nu(z)$ for any $a \in \mathcal{A}_M$. $\Diamond$

It follows immediately that, for a fixed sequence of phases $\{\theta_n\}$ known to the receiver with elements in $\{2\pi m/M: m = 0, \ldots, M-1\}$, the statistics of the SISO EXT pmfs $\{P_n(a): n = 1, \ldots, N\}$ given that $\boldsymbol{x} \in \mathcal{C}$ is transmitted over the $\mathcal{A}_M$-invariant additive-noise memoryless channel $y_n = e^{j\theta_n} x_n + \nu_n$, does not depend on $\{\theta_n\}$. Hence, without loss of generality, in a $\mathcal{A}_M$-invariant channel we can always assume that the all-one codeword belongs to the code $\mathcal{C}$.

SISO Decoders: The message $P_{k,n}(a)$ at the output of the kth SISO decoder is computed by the "windowed" BCJR algorithm [59], i.e., by the forward–backward recursion [41] applied to a window including W trellis sections to the right and W trellis sections to the left of the trellis section containing the coded symbol transmitted, after interleaving, at time n.[4] The forward and backward recursions on each window are initialized by uniform probabilities of the trellis states. It is well known that, as W increases, the windowed BCJR algorithm converges to the performance of the full BCJR algorithm [59].

Computation Flowgraph and Oriented Neighborhoods: From the factor-graph, we derive a *computation flowgraph* containing only the nodes where computation actually takes place: the IC-MUD nodes and the SISO decoder nodes (corresponding to the nodes q_n and p_k in the original factor-graph of Fig. 1). Moreover, the SISO decoder nodes are expanded into (generally overlapping) windows of $2W+1$ trellis sections each, corresponding to the computation carried out by the windowed BCJR algorithm. We need the following definition.

Definition 2: Consider the oriented edge $\vec{e}_{k,n}$ between the nth IC-MUD node and the kth SISO decoder node in the computation flowgraph. The *oriented neighborhood* $\mathcal{N}_{k,n}^{(\ell)}$ of $\vec{e}_{k,n}$ is defined as the collection of all paths of length 2ℓ starting at the nth IC-MUD node and not containing $\vec{e}_{k,n}$ [27].[5] $\Diamond$

For example, Fig. 3 shows an instance of $\mathcal{N}_{k,n}^{(\ell)}$ with $k = m = \ell = 1$ in a computation flowgraph with $K = 3$ users, $s = 1$ symbols per trellis section, and decoder trellis windows of width $W = 1$. The edges in $\mathcal{N}_{k,n}^{(\ell)}$ are oriented according to the computation message flow. In particular, each IC-MUD node has $K-1$ *input* edges and one *output* edge.

The neighborhood *support* $\mathcal{S}(\mathcal{N}_{k,n}^{(\ell)})$ is defined as the collection of all (user $= k$, time $= n$) index pairs such that the nth IC-MUD node has an input edge coming from a trellis window of the kth user SISO decoder.

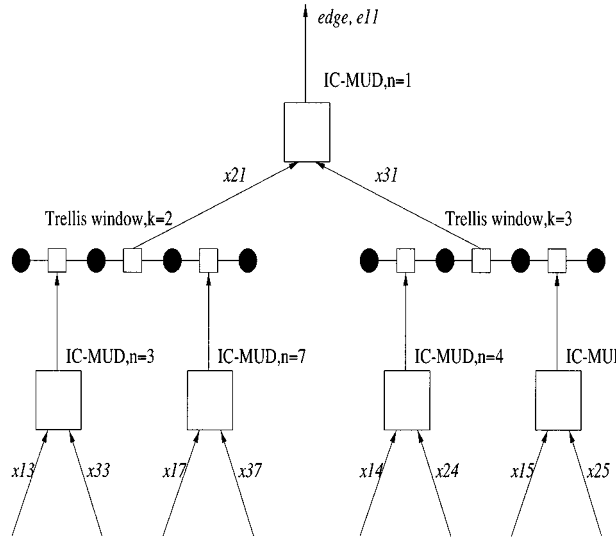

Fig. 3. An instance of the oriented neighborhood $\mathcal{N}_{1,1}^{(1)}$ in a computation flowgraph with $K = 3$ users and trellis window width $W = 1$.

The *type* $\mathcal{T}(\mathcal{N}_{k,n}^{(\ell)})$ of $\mathcal{N}_{k,n}^{(\ell)}$ is defined as the *restriction* of the transmitted codewords $\boldsymbol{X}$ onto the support $\mathcal{S}(\mathcal{N}_{k,n}^{(\ell)})$, i.e.,

$$\mathcal{T}\left(\mathcal{N}_{k,n}^{(\ell)}\right) = \left\{x_{i,j}: (i,j) \in \mathcal{S}\left(\mathcal{N}_{k,n}^{(\ell)}\right)\right\}. \tag{31}$$

We have

$$n_t(\ell) \triangleq \left|\mathcal{T}\left(\mathcal{N}_{k,n}^{(\ell)}\right)\right| = (K-1)\frac{\mathcal{K}^{\ell+1}-1}{\mathcal{K}-1} \tag{32}$$

where $\mathcal{K} = (K-1)((2W+1)s-1)$. In the example of Fig. 3, we have $n_t(1) = 10$ and

$$\begin{aligned}\mathcal{S}\left(\mathcal{N}_{1,1}^{(1)}\right) &= \{(2,1), (3,1), (1,3), (3,3), (1,7), (3,7),\\ &\qquad (1,4), (2,4), (1,5), (2,5)\}\\ \mathcal{T}\left(\mathcal{N}_{1,1}^{(1)}\right) &= \{x_{2,1}, x_{3,1}, x_{1,3}, x_{3,3}, x_{1,7}, x_{3,7}, x_{1,4},\\ &\qquad x_{2,4}, x_{1,5}, x_{2,5}\}.\end{aligned}$$

Random Graphs Ensemble: We consider the ensemble of all computation flowgraphs $(\mathcal{G}, \boldsymbol{\Theta})$ defined by the edge connections in the system factor-graph $\mathcal{G}$ (which, in its turn, is defined by the random interleavers) and by the random phase sequences $\boldsymbol{\Theta}$, while the spreading sequences $\boldsymbol{S}$ and the basic code $\mathcal{C}$ are fixed.

The message $Q_{k,n}^{(\ell)}(a)$ sent over edge $\vec{e}_{k,n}$, is a function of the IC-MUD output variable $z_{k,n}^{(\ell)} = x_{k,n} + \zeta_{k,n}^{(\ell)}$, where the superscript (ℓ) indicates the decoder iteration. The empirical *cumulative distribution function* (cdf) of $\zeta_{k,n}^{(\ell)}$, for a given instance of $(\mathcal{G}, \boldsymbol{\Theta})$ and of the channel noise $\boldsymbol{N} = [\boldsymbol{\nu}_1, \ldots, \boldsymbol{\nu}_N]$, and for given transmitted codewords $\boldsymbol{X} = [\boldsymbol{x}_1, \ldots, \boldsymbol{x}_N]$, is defined by

$$F_k^{(\ell,N)}(z|\boldsymbol{X}) \triangleq \frac{1}{N}\sum_{n=1}^{N} 1\left\{\mathrm{Re}\left\{\zeta_{k,n}^{(\ell)}\right\} \le \mathrm{Re}\{z\},\right.$$
$$\left.\mathrm{Im}\left\{\zeta_{k,n}^{(\ell)}\right\} \le \mathrm{Im}\{z\}\right\}. \tag{33}$$

For fixed $z \in \mathbb{C}$, $F_k^{(\ell,N)}(z|\boldsymbol{X})$ is a function of $(\mathcal{G}, \boldsymbol{\Theta}, \boldsymbol{N})$ and of $\boldsymbol{X}$. Hence, for random $(\mathcal{G}, \boldsymbol{\Theta}, \boldsymbol{N})$ (defined over a common

[4]Because of trellis termination, the computation window is truncated to the left (resp., to the right) for symbols in trellis sections $1, \ldots, W$ (resp., $N/s - W + 1, \ldots, N/s$).

[5]We define a neighborhood as a collection of paths, and not as a subgraph. Then, a neighborhood can contain repeated nodes and edges. In general, we use the term "collection" to denote a multiset, where elements might be repeated.

probability space), $F_k^{(\ell,N)}(z|\boldsymbol{X})$ is a random variable taking on values in $[0, 1]$.

We define also the (nonrandom) cdf

$$\overline{F}_k^{(\ell)}(z|\boldsymbol{X}) \triangleq E\left[1\left\{\mathrm{Re}\left\{\zeta_{k,n}^{(\ell)}\right\} \le \mathrm{Re}\{z\},\right.\right. \\ \left.\left.\mathrm{Im}\left\{\zeta_{k,n}^{(\ell)}\right\} \le \mathrm{Im}\{z\}\right\}\middle| \boldsymbol{X}, \mathcal{N}_{k,n}^{(\ell)} \text{ is cycle-free}\right] \quad (34)$$

where expectation is with respect to $(\mathcal{G}, \boldsymbol{\Theta}, \boldsymbol{N})$. It is evident by symmetry that $\overline{F}_k^{(\ell)}(z|\boldsymbol{X})$ does not depend on n.

B. Concentration and Density Evolution

The results of this section show that, for every $k = 1, \ldots, K$ and a given finite ℓ, as $N \to \infty$ the random empirical cdf $F_k^{(\ell,N)}(z|\boldsymbol{X})$ converges almost surely to the nonrandom cdf $\overline{F}_k^{(\ell)}(z|\boldsymbol{X})$ and that the latter does not depend on $\boldsymbol{X}$. It follows that the limit for large block length of the residual interference cdf for user k at decoder iteration ℓ can be computed by propagating the expected "densities" of the messages through a cycle-free random neighborhood. The corresponding density propagation algorithm is referred to as DE [27]. We need the following preliminary results, proved in Appendix A.

Lemma 1: For given k, n, ℓ, and the ensemble of random computation flow graphs defined in Section IV-A

$$\Pr\left(\mathcal{N}_{k,n}^{(\ell)} \text{ is not cycle-free}\right) \le \frac{c_1}{N} \quad (35)$$

where c_1 depends on ℓ, K, W, and s, but it is independent of N. □

Lemma 2: Let $\mathcal{C}$ be a geometrically uniform code over $\mathcal{A}_M$ as defined in Section IV-A and consider the transmission of $\mathcal{C}$ over a memoryless $\mathcal{A}_M$-invariant channel with noise pdf $p_\nu(z)$. Let $\{P_n(a)\colon a \in \mathcal{A}_M\}$ be the nth output message produced by the SISO decoder (11) when the transmitted codeword is $\hat{\boldsymbol{x}}$, the channel observation is $\boldsymbol{y} = \hat{\boldsymbol{x}} + \boldsymbol{\nu}$, and the corresponding input messages are $\{Q_j(a) = p_\nu(y_j - a), j = 1, \ldots, N\}$. Let $\hat{x}_n = b$, for some $b \in \mathcal{A}_M$ and $f_{P_n|\hat{\boldsymbol{x}}}(p_1, \ldots, p_M)$, be the pdf of the output message $\{P_n(a)\colon a \in \mathcal{A}_M\}$. Then, for any $\hat{\boldsymbol{x}}' \in \mathcal{C}$ with $\hat{x}_n' = b'$, there exists a permutation π of $\{1, \ldots, M\}$ that depends only on b and b' such that

$$f_{P_n|\hat{\boldsymbol{x}}'}(p_1, \ldots, p_M) = f_{P_n|\hat{\boldsymbol{x}}}(p_{\pi(1)}, \ldots, p_{\pi(M)}) \quad \square$$

Corollary 1: Given that $\mathcal{N}_{k,n}^{(\ell)}$ is cycle-free

$$\boldsymbol{X} \to \mathcal{T}(\mathcal{N}_{k,n}^{(\ell)}) \to \zeta_{k,n}^{(\ell)}$$

is a Markov chain, i.e., $\zeta_{k,n}^{(\ell)}$ is independent of $\boldsymbol{X}$ given the neighborhood type. □

Corollary 2: For the ensemble of computation flowgraphs and channel noise $\{(\mathcal{G}, \boldsymbol{\Theta}, \boldsymbol{N})\}$ and for a user basic code $\mathcal{C}$ in the class defined in Section IV-A, $\overline{F}_k^{(\ell)}(z|\boldsymbol{X}) = \overline{F}_k^{(\ell)}(z)$, independent of $\boldsymbol{X}$. □

Then, our main concentration result is as follows.

Proposition 1: For all $k = 1, \ldots, K$ and for any $\epsilon > 0$, there exists $c_2 > 0$ independent of N and k such that, if $N > 2c_1/\epsilon$, then

$$\Pr\left(\left|F_k^{(\ell,N)}(z|\boldsymbol{X}) - \overline{F}_k^{(\ell)}(z)\right| \ge \epsilon\right) \le 2e^{-c_2\epsilon^2 N}. \quad (36)$$

Proof: See Appendix A. □

Remark 8: An immediate consequence of Proposition 1 is that the EXT messages from SISO decoders $j \neq k$ sent to the nth IC-MUD node are asymptotically statistically independent of $x_{k,n}$ and of the noise $\boldsymbol{\nu}_n$. In fact, in a cycle-free neighborhood these EXT messages depend only on the noise and symbols corresponding to IC-MUD nodes that are children of the nth IC-MUD node. Because of the cycle-freeness, these nodes correspond to time indexes $j \neq n$. ◊

From Proposition 1 and the first Borel–Cantelli lemma [60, p. 53], since ϵ is arbitrary we get that $F_k^{(\ell,N)}(z|\boldsymbol{X}) \to \overline{F}_k^{(\ell)}(z)$ holds almost surely. Hence, for $N \to \infty$, the residual interference variables contributing to same trellis window of the SISO decoder of any user k at iteration ℓ is almost surely i.i.d., $\sim\overline{F}_k^{(\ell)}(z)$. It follows that the BER of user k at iteration ℓ is the same as if $\mathcal{C}$ were transmitted over a memoryless time-invariant additive-noise channel with noise cdf $\overline{F}_k^{(\ell)}(z)$. Since the BER is a continuous functional of the noise cdf, we conclude that the BER for any user k after ℓ decoder iterations can be calculated from $\overline{F}_k^{(\ell)}(z)$. The iterative computation of $\overline{F}_k^{(\ell)}(z)$ is obtained by the following DE algorithm.

Density Evolution: We let $\overline{f}_k^{(\ell)}(z)$ be the pdf corresponding to $\overline{F}_k^{(\ell)}(z)$. Because of the phase randomization and since the noise is circularly symmetric, it is easy to see that the residual interference variable $\zeta_{k,n}^{(\ell)} \sim \overline{F}_k^{(\ell)}(z)$ is $\mathcal{A}_M$-invariant. Let $\overline{g}_k^{(\ell)}(z|x)$ be the pdf of the symbol estimate $\tilde{x}_{k,n}^{(\ell)}$ conditioned on $x_{k,n} = x$, at the ℓth DE iteration. Lemma 2 implies that

$$\overline{g}_k^{(\ell)}(z|x') = \overline{g}_k^{(\ell)}(x(x')^*z|x).$$

By exploiting this symmetry, only the evaluation of the conditional pdf $\overline{g}_k^{(\ell)}(z|x=1)$ is actually required and, without loss of generality, DE can be run by assuming that the all-one codeword is transmitted by all users. For the sake of notational simplicity, we indicate $\overline{g}_k^{(\ell)}(z|x=1)$ simply by $\overline{g}_k^{(\ell)}(z)$.

IC-MUD and SISO decoding define the (random) mappings

$$\{\tilde{x}_{j,n}\colon j \neq k\} \mapsto \zeta_{k,n} \quad (37)$$

and

$$\{\zeta_{k,j}\colon j \neq n\} \mapsto \tilde{x}_{k,n} \quad (38)$$

respectively, where, because of what was said earlier, it is understood that $x_{k,n} = 1$ is the transmitted symbol, and where randomness comes from the phases $\boldsymbol{\Theta}$ and the channel noise $\boldsymbol{N}$.

We define the pdf mappings $f_k = \mathcal{F}_k(\{g_j\colon j \neq k\})$ and $g_k = \mathcal{G}_k(f_k)$ corresponding to (37) and to (38), respectively, under the assumption of statistically independent inputs. This means that, if

$$\{\tilde{x}_{j,n}\colon j \neq k\} \sim \prod_{j \neq k} g_j(z_j)$$

then $\zeta_{k,n} \sim f_k = \mathcal{F}_k(\{g_j: j \neq k\})$, and if

$$\{\zeta_{k,j}: j \neq n\} \sim \prod_{j\neq n} f_k(z_j)$$

then $\tilde{x}_{k,n} \sim g_k = \mathcal{G}_k(f_k)$.

The estimated modulation symbols are initialized as $\tilde{x}_{k,n}^{(0)} = 0$, yielding $\overline{g}_k^{(0)} = \delta$, where δ denotes a Dirac distribution centered at zero. DE for parallel scheduling can be concisely expressed as follows. Let $\overline{f}_k^{(1)} = \mathcal{F}_k(\{g_j = \delta: j \neq k\})$ for $k = 1, \ldots, K$. Then, for $\ell = 2, 3, \ldots$, let

$$\overline{f}_k^{(\ell)} = \mathcal{F}_k\left(\left\{\mathcal{G}_j\left(\overline{f}_j^{(\ell-1)}\right): j \neq k\right\}\right), \qquad k = 1, \ldots, K. \tag{39}$$

Similarly, DE for serial scheduling is given by

$$\overline{f}_k^{(1)} = \mathcal{F}_k\left(\left\{\mathcal{G}_j\left(\overline{f}_j^{(1)}\right): j < k\right\}, \{g_j = \delta: j > k\}\right), \quad k = 1, \ldots, K$$

and, for $\ell = 2, 3, \ldots$, by

$$\overline{f}_k^{(\ell)} = \mathcal{F}_k\left(\left\{\mathcal{G}_j\left(\overline{f}_j^{(\ell)}\right): j < k\right\}, \left\{\mathcal{G}_j\left(\overline{f}_j^{(\ell-1)}\right): j > k\right\}\right), \quad k = 1, \ldots, K. \tag{40}$$

By inspection of (39) and (40) we see that $\{f_k : k = 1, \ldots, K\}$ is a fixed point of the parallel DE, i.e., it satisfies

$$f_k = \mathcal{F}_k(\{\mathcal{G}_j(f_j): j \neq k\}), \qquad \text{for all } k$$

if and only if it is a fixed point of the serial DE. Therefore, serial and parallel scheduling have (almost surely, as $N \to \infty$) the same set of fixed points. If $K = 2$, the two schedulings yield identical DE and, therefore, yield the same BER. For $K > 2$, it is difficult to say if they converge to the same stable fixed point for any choice of the system parameters.

The DE algorithms (39) and (40) are easily implemented by Monte Carlo simulation of the empirical histograms corresponding to the pdfs $\overline{f}_k^{(\ell)}(z)$ and $\overline{g}_k^{(\ell)}(z)$. The BER after a given number of iterations can be measured from the output of the SISO decoder, as in standard Monte Carlo simulation. A faster approximated method for BER evaluation consists of making the Gaussian approximation

$$\overline{f}_k^{(\ell)}(z) \approx \mathcal{N}_{\mathbb{C}}\left(0, E\left[\left|\zeta_{k,n}^{(\ell)}\right|^2\right]\right). \tag{41}$$

Let BER $= e_b(\mathsf{SNR})$ be the BER versus SNR function for the basic code $\mathcal{C}$ in additive white Gaussian noise (AWGN).[6] Then, the kth user BER at iteration ℓ can be approximated as $\text{BER}_k^{(\ell)} \approx e_b(1/E[|\zeta_{k,n}^{(\ell)}|^2])$. In the next section, we shall see that this approximation holds exactly for random CDMA in the limit for large K and L.

C. Large-System Limit and Gaussian Approximation

In this subsection, we consider the large-system limit performance of the iterative IC decoders for random CDMA [29]–[31]. We assume that the spreading sequences are randomly generated with elements $s_{m,k} = \frac{1}{L} v_{m,k}$, where $v_{m,k}$ for $m = 1, \ldots, L$ and $k = 1, \ldots, K$ are i.i.d. complex circularly symmetric random variables with $E[v_{m,k}] = 0$, $E[|v_{m,k}|^2] = 1$, and $E[|v_{m,k}|^4] < \infty$, and we let $K \to \infty$ with $L = K/\alpha$, for fixed channel load α. The DE limiting pdfs (provided that they exist), depend on the basic code $\mathcal{C}$, on the user amplitudes $\{\sqrt{\gamma_k}\}$ and on the spreading matrix $\boldsymbol{S}$. Hence, for random $\boldsymbol{S}$ the limiting users' BER is a random variable. We shall show that, in the large-system regime, the BER performance of each user converges almost surely to a deterministic quantity. Moreover, the resulting DE is much simpler than the general DE for given $\boldsymbol{S}$ and finite K, L.

An important *caveat* should be made about the order of the limits with respect to N and K. In fact, Proposition 1 holds for finite K.[7] Then, our results are valid if we let *first* $N \to \infty$ and *then* $K \to \infty$, and not *vice versa*. In practice, this means that a random CDMA system with finite N and K performs with high probability very close to the limit provided in this section if $N \gg K$. Fortunately, this is a realistic assumption for practical CDMA systems [36].

In this subsection and in the rest of the paper we focus on parallel IC algorithms only, since in the limit for infinite K it is not clear how to compute the serial DE. The following result and its corollary characterize the large-system limit of the iterative multiuser IC decoders with parallel scheduling.

Proposition 2: Assume that, for all k, the user squared signal amplitudes satisfy $\gamma_k \in [0, \gamma_{\max}]$ for some finite $\gamma_{\max}$ independent of K, and that the empirical cdf of the γ_k's, defined by

$$G^{(0,K)}(\gamma) = \frac{1}{K}\sum_{k=1}^{K} 1\{\gamma_k \leq \gamma\} \tag{42}$$

converges weakly [60, Ch. 5] to a given (nonrandom) cdf $G^{(0)}(\gamma)$ as $K \to \infty$. Then, as $K, L \to \infty$ with $K/L = \alpha$, $\overline{F}_k^{(\ell)}(z)$ converges almost surely to the Gaussian distribution $\mathcal{N}_{\mathbb{C}}(0, \frac{1}{\gamma_k \eta^{(\ell)}})$ where, for hard IC and SUMF-based soft-IC, $\eta^{(\ell)}$ is given by

$$\eta^{(\ell)} = \frac{1}{1 + \alpha E\left[U^{(\ell-1)}\right]} \tag{43}$$

where the cdf of $U^{(\ell)}$ is defined by the limit

$$G^{(\ell)}(u) = \lim_{K\to\infty} \frac{1}{K}\sum_{k=1}^{K} 1\left\{\gamma_k \left|x_{k,n} - \tilde{x}_{k,n}^{(\ell)}\right|^2 \leq u\right\} \tag{44}$$

(that exists and is independent of n) and, for LMMSE-based soft-IC, $\eta^{(\ell)}$ is given by the unique nonnegative solution of

$$\eta = \frac{1}{1 + \alpha E\left[\frac{U^{(\ell-1)}}{1+\eta U^{(\ell-1)}}\right]} \tag{45}$$

where the cdf of $U^{(\ell)}$ is defined by the limit

$$G^{(\ell)}(u) = \lim_{K\to\infty} \frac{1}{K}\sum_{k=1}^{K} 1\left\{\gamma_k \left(1 - \left|\tilde{x}_{k,n}^{(\ell)}\right|^2\right) \leq u\right\} \tag{46}$$

(that exists and is independent of n).

Proof: See Appendix A. □

[6]The function $e_b(\mathsf{SNR})$ can be either obtained by simulation or by standard union bound [37], or by a combination of both. For example, for low SNR, we can use simulation and for high SNR, where simulation would be too time-consuming and union bound is tight, we can use the union bound.

[7]In general, concentration [27] requires that the maximum degree in the random graph representing the message-passing decoder is finite, while the size of the graph grows to infinity.

Corollary 3: Under the assumptions of Proposition 2, DE with parallel scheduling converges almost surely to a stable fixed point (which exists almost surely), uniquely identified by the limit $\eta^* \stackrel{\Delta}{=} \lim_{\ell \to \infty} \eta^{(\ell)}$, where $\eta^{(\ell)}$ is given in Proposition 2.

Proof: See Appendix A. □

From Proposition 2 it follows that the Gaussian approximation (41) of the SISO decoder input holds exactly in the large-system limit. Next, we propose a Gaussian approximation of the SISO decoder *output* [32]–[34], which makes the iterative computation of $\eta^{(\ell)}$ very simple. The resulting Gaussian approximation of the DE algorithm will be referred to as the GA-DE.

Let $\epsilon = e_s(\mathsf{SNR})$ denote the symbol error rate (SER) versus. SNR function for the basic code $\mathcal{C}$ over AWGN,[8] when decision are made according to the EXT-based hard detection rule

$$\tilde{x}_{j,n} = \arg\max_{a \in \mathcal{A}_j} P_{j,n}(a). \tag{47}$$

The SER of user k at iteration $\ell - 1$ is given by

$$\epsilon = e_s\left(\gamma_k \eta^{(\ell-1)}\right). \tag{48}$$

For hard IC we can write

$$\begin{aligned} &E\left[|x_{k,n} - \tilde{x}_{k,n}|^2\right] \\ &\quad = 2 - 2\mathrm{Re}\left\{E[x_{k,n}^* \tilde{x}_{k,n}]\right\} \\ &\quad \approx 2\left(1 - (1-\epsilon) - \frac{\epsilon}{M-1}\mathrm{Re}\left\{\sum_{m=1}^{M-1} e^{j2\pi m/M}\right\}\right) \\ &\quad = 2\epsilon\,\frac{M}{M-1} \end{aligned} \tag{49}$$

where we assumed that, if a symbol is in error, any other symbol appears with uniform probability $\epsilon/(M-1)$, and we used the fact that $\sum_{m=1}^{M-1} e^{j2\pi m/M} = -1$ for $M \geq 2$. By using (49) and (48) in (43) we obtain the recursion

$$\eta^{(\ell)} = \frac{1}{1 + 2\alpha\,\frac{M}{M-1}\int_0^\infty \gamma e_s\left(\gamma\eta^{(\ell-1)}\right)\,dG^{(0)}(\gamma)} \tag{50}$$

initialized by $\eta^{(0)} = 0$.

For soft IC, we assume that the SISO decoder output messages $P_{k,n}(a)$ are produced by a virtual uncoded AWGN channel $y = x_{k,n} + \nu$, whose SNR is such that the SER with symbol-by-symbol detection is equal to ϵ given in (48).[9] Denote by $\mu(\epsilon)$ the SNR of such virtual channel and, without loss of generality, assume that $x_{k,n} = 1$ is the transmitted symbol. Then, the GA of the SISO output yields the estimated symbol

$$\tilde{x} \stackrel{\Delta}{=} \sum_{a \in \mathcal{A}_M} a\,\frac{e^{-\mu(\epsilon)|\nu+1-a|^2}}{\sum\limits_{a' \in \mathcal{A}_M} e^{-\mu(\epsilon)|\nu+1-a'|^2}}. \tag{51}$$

From the above approximation, we can write

$$E\left[|x_{k,n} - \tilde{x}_{k,n}|^2\right] = E_\nu\left[|1-\tilde{x}|^2\right] \stackrel{\Delta}{=} V(\epsilon) \tag{52}$$

[8]As for the BER function $e_b(\mathsf{SNR})$, also the function $e_s(\mathsf{SNR})$ can be obtained by the combination of simulation (for low SNR) and union bound (for high SNR).

[9]The GA proposed here is based on matching the SER at the decoder output, and differs from the GA proposed in [32], [33], that matches the conditional *mean value* of the EXT log-likelihood ratio $\mathcal{L}_{k,n} = \log\frac{P_{k,n}(+1)}{P_{k,n}(-1)}$ given $x_{k,n} = +1$ and from the GA proposed in [34], that matches the mutual information $I(\mathcal{L}_{k,n}; x_{k,n})$. These two GA's methods can be applied to binary antipodal signaling, and are widely used in the analysis and design of random-like codes under iterative message-passing decoding.

where $E_\nu[\cdot]$ denotes expectation with respect to $\nu \sim \mathcal{N}_{\mathbb{C}}(0, 1/\mu(\epsilon))$, and where we used the fact that, for a given modulation alphabet $\mathcal{A}_M$, $E_\nu[|1-\tilde{x}|^2]$ in (52) depends only on the SNR $\mu(\epsilon)$, which is a one-to-one function of the SER ϵ; therefore, the function $V(\epsilon) \stackrel{\Delta}{=} E_\nu[|1-\tilde{x}|^2]$ is well defined. By using this into (43) we obtain the GA-DE recursion for the SUMF-based iterative soft IC decoder as

$$\eta^{(\ell)} = \frac{1}{1 + \alpha\int_0^\infty \gamma V\left(e_s\left(\gamma\eta^{(\ell-1)}\right)\right)\,dG^{(0)}(\gamma)} \tag{53}$$

with initial condition $\eta^{(0)} = 0$.

For the LMMSE-based algorithm, notice that the statistics of $\tilde{x}$ is completely defined by the parameter ϵ. Define the cdf $H(u|\epsilon) = \Pr(1 - |\tilde{x}|^2 < u)$. Then, from (45), the GA-DE recursion for the LMMSE-based iterative soft IC decoder is given by $\eta^{(\ell)}$ solution of

$$\eta = \frac{1}{1 + \alpha\int_0^\infty\int_0^\infty \frac{\gamma u}{1+\gamma u\eta}\,dH\left(u\,\middle|\,e_s\left(\gamma\eta^{(\ell-1)}\right)\right)\,dG^{(0)}(\gamma)} \tag{54}$$

initialized again by $\eta^{(0)} = 0$. Despite its apparent complication, the recursion (54) is easily computable in several cases of interest (see the example in Appendix B).

Recursions (50), (53), and (54), completely characterize, up to the approximations in (49) and in (51), the behavior of the hard IC, the SUMF-based soft IC, and the LMMSE-based soft IC iterative multiuser decoders in terms of one-dimensional dynamical systems with state variable $\eta^{(\ell)}$. In brief, these dynamical systems can be written as

$$\eta^{(\ell)} = \Psi\left(\eta^{(\ell-1)}, \alpha, G^{(0)}(\gamma)\right), \qquad \ell = 1, 2, \ldots \tag{55}$$

with $\eta^{(0)} = 0$, and where the *mapping function* Ψ depends on the particular algorithm considered and on the system parameters: the channel load α and the cdf of the users' SNRs $G^{(0)}(\gamma)$. By replicating the proof of Corollary 3, it is immediate to see that the iterated map (55) with initial condition $\eta^{(0)} = 0$ converges always to a limit $\tilde{\eta}^\star \in [0, 1]$. As shown in Section V, for most choices of the system parameters $\tilde{\eta}^\star$ is a good approximation of the limit $\eta^\star$ of the exact DE.

V. Results

In this section, we present some numerical examples in order to illustrate the main theoretical results of this work. All results are based on binary convolutional codes (CCs), with bit-interleaving, Gray mapping [37], and QPSK modulation. We focus on this case since it is very important in CDMA applications [36] and the computation of the GA-DE is particularly simple. The details of the DE-GA calculation are given in Appendix B.

The Ψ Mapping and Its Fixed Points: Fig. 4 shows the function $\Psi(\eta, \alpha, G^{(0)}(\gamma))$ defined in (55) for the LMMSE-based soft IC decoder, different values of the channel load α and equal-power users, i.e., $G^{(0)}(\gamma) = 1\{\gamma \geq \gamma_0\}$. The user basic code is the binary CC of rate $1/2$ with octal generators $(5, 7)$ (see [37]), briefly denoted in the following by "CC $(5, 7)$." Users have all the same E_b/N_0 equal to 6 dB, corresponding to $\gamma_0 = 3.9811$. For $\alpha = 1.0, 1.4, 1.8$, and 2.2,

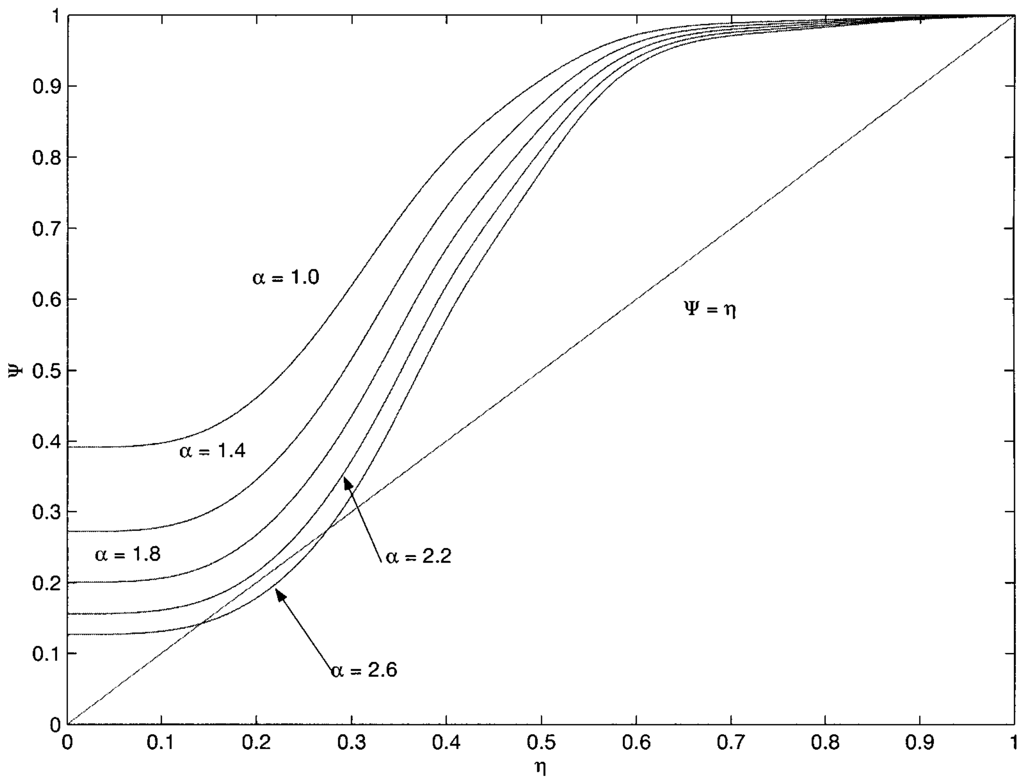

Fig. 4. Ψ mapping function for the GA-DE of LMMSE-based soft IC decoding, with CC $(5, 7)$, QPSK modulation, equal power users with $E_b/N_0 = 6$ dB, and $\alpha = 1.0, 1.4, 1.8, 2.2$, and 2.6.

the equation $\eta = \Psi(\eta, \alpha, G^{(0)}(\gamma))$ has only one solution $\tilde{\eta}^\star \approx 1$, corresponding to the unique (stable) fixed point of the GA-DE. Since $\tilde{\eta}^\star$ is very close to 1, we expect that all users achieve near-single-user BER after a sufficiently large number of decoder iterations. On the contrary, for $\alpha = 2.6$, the equation $\eta = \Psi(\eta, \alpha, G^{(0)}(\gamma))$ has three solutions (given by the intersections of the curve $\Psi(\eta, \alpha, G^{(0)}(\gamma))$ with the first quadrant diagonal $\Psi = \eta$). In this case, the limit $\tilde{\eta}^\star$ is given by the smallest solution, which is ≈ 0.14. Hence, we expect that for $\alpha = 2.6$ all users perform very far from the single-user BER (the SINR penalty in decibels with respect to the single-user SNR is given by $10\log_{10}(\tilde{\eta}^\star) \approx -8.53$ dB).

We notice that the iterative multiuser decoder has a threshold behavior with respect to the channel load α, i.e., there exist a value $\overline{\alpha}$ such that, for all $\alpha < \overline{\alpha}$ near-single-user performance is achieved by all users, while for all $\alpha \geq \overline{\alpha}$, the degradation with respect to the single-user performance is very large for all users. The value $\overline{\alpha}$ corresponds to a *fold bifurcation* [61] of the dynamical system defined by (55), i.e., it is the value below which the second stable fixed point disappears.

The fixed-point analysis is confirmed by Fig. 5, showing a snapshot simulation of a random finite-dimensional system with $L = 60$, $N = 2000$, and the same user basic code, modulation, and E_b/N_0 of the infinite-dimensional system of Fig. 4, but where the user spreading sequences and interleavers are randomly generated at each frame. A total of 10 frames (each corresponding to a block of length N coded QPSK symbols) were generated. The scattered points in Fig. 5 show the minimum and maximum empirical SINR/γ_0 ratio over the user population, for each simulated frame. The empirical SINR for user k at iteration ℓ is evaluated as

$$\mathsf{SINR}_k^{(\ell)} = \frac{1}{\frac{1}{N}\sum_{n=1}^{N}\left|z_{k,n}^{(\ell)} - x_{k,n}\right|^2}.$$

The exact DE and the GA-DE trajectories of $\eta^{(\ell)}$ versus the number of iterations ℓ for the infinite-dimensional system are shown for comparison.

Two observations should be pointed out here. 1) The GA-DE and exact DE trajectories agree very well for $\alpha = 1.8$ and 2.6, i.e., for α far from the threshold $\overline{\alpha}$. On the contrary, for $\alpha = 2.2$ that is close to the threshold $\overline{\alpha}$ the two trajectories converge to nearly the same fixed point but in a different number of iterations. In particular, the GA-DE is slightly optimistic in terms of the number of iterations needed to converge. 2) The behavior of the random finite-dimensional system is very well predicted by the infinite-dimensional DE, and quite accurately predicted (at least in terms of the limiting fixed point) by the GA-DE.

Fig. 6 shows the GA-DE mapping function Ψ for CC $(5, 7)$, equal-power users with $E_b/N_0 = 6$ dB, $\alpha = 1.8$, and LMMSE, SUMF, and hard IC iterative multiuser decoders. The threshold load $\overline{\alpha}$ characterizing the decoder fold bifurcation depends on the decoder algorithm. In the example of Fig. 6, we observe that the LMMSE-based decoder is below its threshold load while the hard-IC and SUMF-based soft IC decoders are above their threshold load.

Finally, Fig. 7 shows the GA-DE mapping function Ψ for the LMMSE-based decoder for a system with CC $(5, 7)$ basic code,

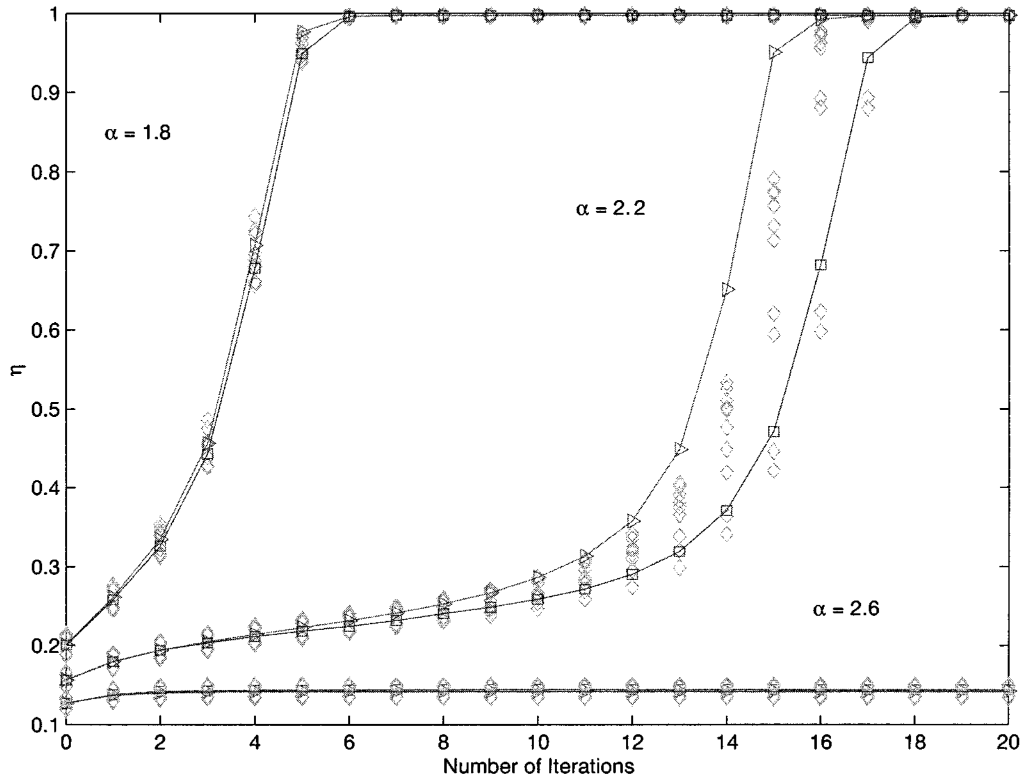

Fig. 5. Evolution of $\eta^{(\ell)}$ versus the number of iterations for LMMSE-based soft IC decoding, with CC $(5, 7)$, QPSK modulation, equal power users with $E_b/N_0 = 6$ dB, and $\alpha = 1.8, 2.2,$ and 2.6. Solid lines with triangles and squares denote the GA-DE and the exact DE trajectories, respectively, and lozenges denote simulation snapshots of a finite-dimensional system with spreading factor $L = 60$ and block length $N = 2000$.

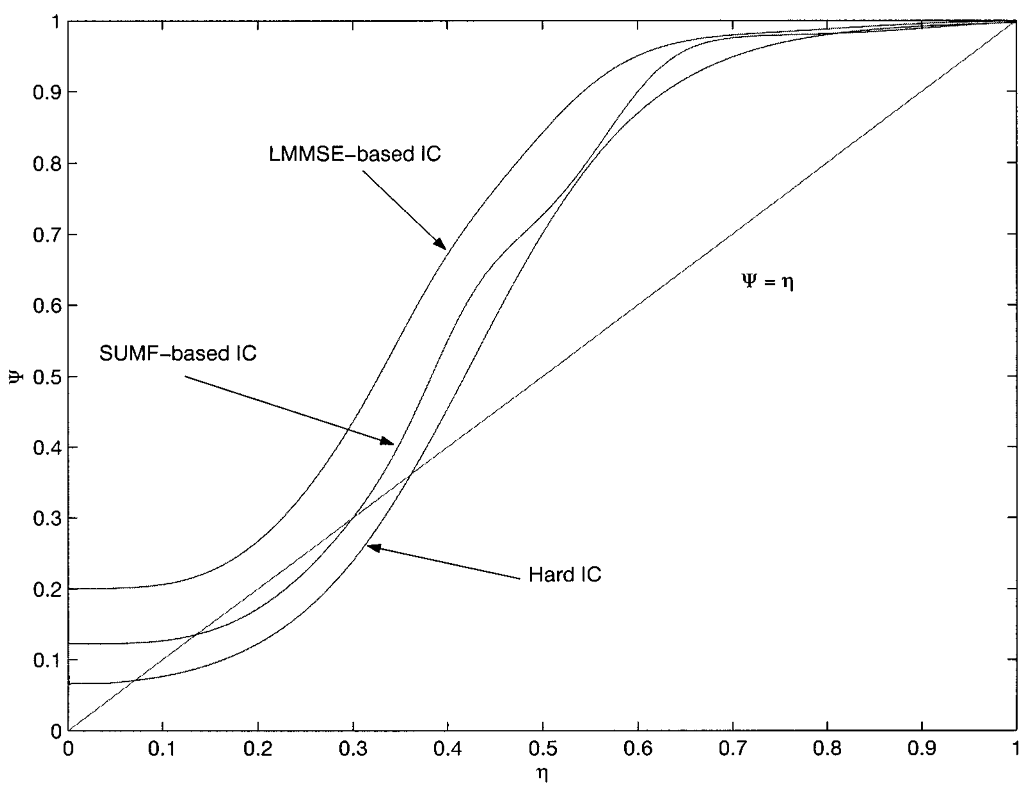

Fig. 6. Ψ mapping function for the GA-DE of different IC decoders, with CC $(5, 7)$, QPSK modulation, equal power users with $E_b/N_0 = 6$ dB, and $\alpha = 1.8$.

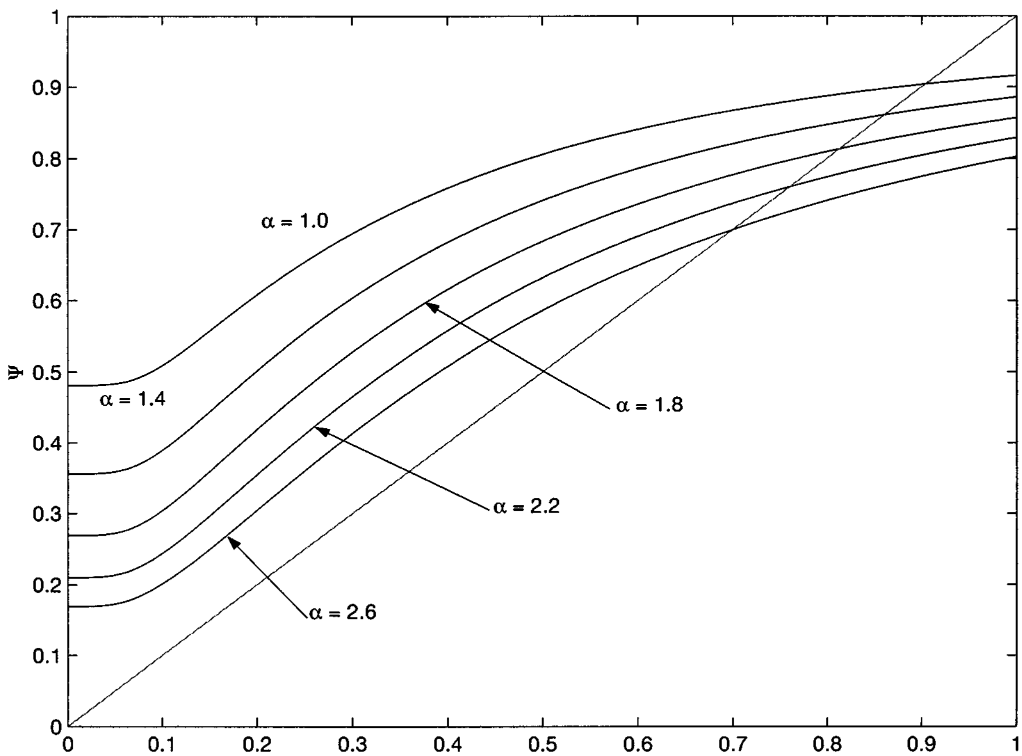

Fig. 7. Ψ mapping function for the GA-DE of LMMSE-based soft IC decoding, with CC $(5, 7)$, QPSK modulation, block-fading with average $E_b/N_0 = 6$ dB, and $\alpha = 1.0, 1.4, 1.8, 2.2$, and 2.6.

for an unequal-power system where the users SNR is distributed according to a truncated exponential pdf, with corresponding cdf given by

$$G^{(0)}(\gamma) = \begin{cases} 0, & \gamma < 0 \\ \frac{1-e^{-\kappa\gamma}}{1-e^{-\kappa\gamma_{\max}}}, & 0 \le \gamma \le \gamma_{\max} \\ 1, & \gamma > \gamma_{\max} \end{cases}$$

(in this example we chose $\gamma_{\max} = 40$ and κ such that E_b/N_0 averaged over all users is equal to 6 dB). This case is representative of a block-fading channel where the channel gain for each user is constant over the whole codeword duration (see [31]) and where the fading gain cdf is given by $G^{(0}(\gamma)$. Suppose that the fading of each user is governed by an ergodic process [62]. Then, the average BER of each user can be evaluated as BER $= E[e_b(\gamma\tilde{\eta}^\star)]$, where expectation is with respect to $\gamma \sim G^{(0)}(\gamma)$ given above, and where $e_b(\cdot)$ is the BER versus SNR function characteristic of the basic user code.

We notice that, with such unequal power distribution, the iterative decoder behavior is quite different from the corresponding equal-power case of Fig. 4. In fact, here the unique decoder fixed point decreases smoothly as the load increases, and no fold bifurcation appears for the considered range of channel load α.

EXT Versus APP Feedback, and the Bias Effect: In Remark 3 of Section III-C we anticipated that feeding back MAI estimates obtained from the APPs rather than from the EXT of SISO decoders makes residual interference conditionally biased and degrades the decoder performance. Figs. 8 and 9 illustrate this fact for the SUMF-based decoder. A random finite-dimensional system with $L = 60$, $N = 2000$, $\alpha = 1.4$, CC $(5, 7)$, equal-power users with $E_b/N_0 = 6$ dB was simulated with EXT-based and APP-based feedback. Fig. 8 shows the simulated SINR/γ_0 ratio (defined as in Fig. 5) for the two decoders, versus the number of iterations. We notice that the EXT-based decoder converges to $\tilde{\eta}^\star \approx 1$, and its behavior agrees very well with the GA-DE trajectory, while the APP-based decoder does not improve its performance with iterations. Fig. 9 shows the corresponding evolution of the conditional residual interference bias for the signal in-phase component, defined by $E[\mathrm{Re}\{\zeta_{k,n}^{(\ell)}\}|\mathrm{Re}\{x_{k,n}\} \geq 0]$. We notice that the EXT-based system converges rapidly to zero bias for all users, while the APP-based system gets trapped into an oscillatory behavior. Notice also that, for the APP-based decoder, the bias averaged over all users is negative, i.e., on the average it reduces the useful signal component.

System Spectral Efficiency: We can use the GA-DE as a *rule of thumb* for a quick evaluation of the spectral efficiency achievable with given codes and iterative multiuser IC decoders. Spectral efficiency is calculated as follows. We fix the target BER to be achieved by all users and we compute E_b/N_0 necessary to achieve the target BER in the single-user case. If the target BER is sufficiently small, because of the threshold behavior illustrated above, it can be achieved only if the decoder works below threshold. Then, the resulting spectral efficiency is given by $\rho = \overline{\alpha}R$, where $\overline{\alpha}$ is the threshold load for the value of E_b/N_0 obtained and R is the coding rate.

For example, Fig. 10 shows the spectral efficiency for target BER equal to 10^{-5} achieved by the optimal CCs of rate $1/2$ with 4, 8, 16, 32, and 64 states [37], with Gray-mapped QPSK modulation and different iterative IC decoding algorithms, for

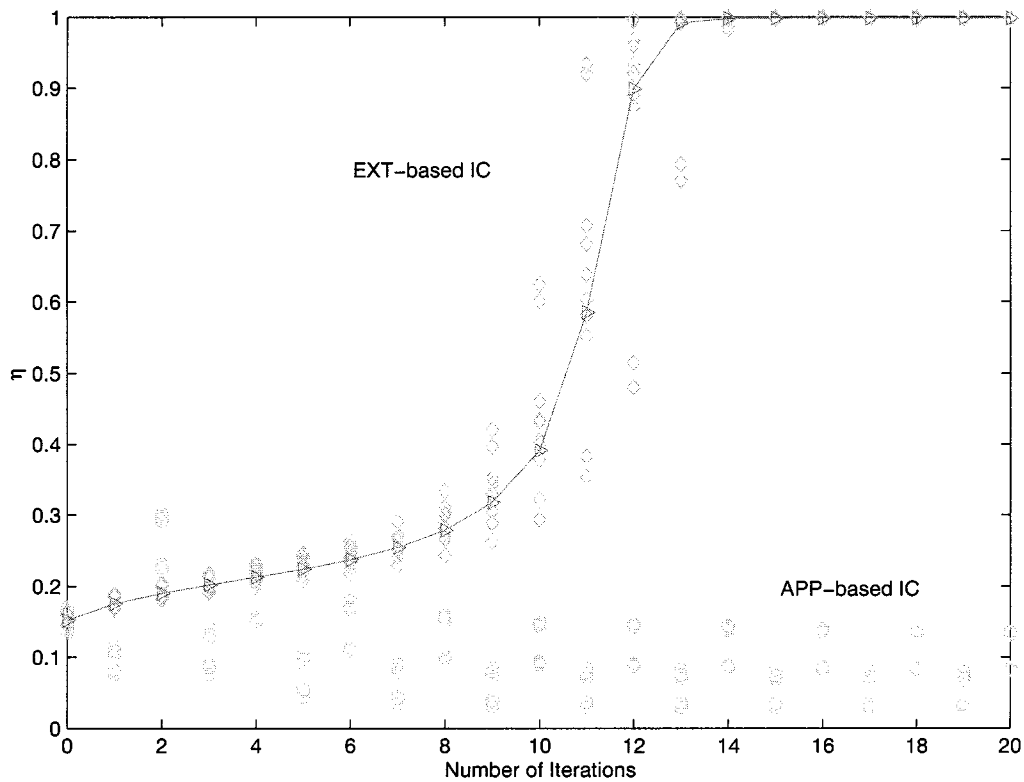

Fig. 8. Evolution of $\eta^{(\ell)}$ versus the number of iterations for SUMF-based soft IC decoding, with CC $(5, 7)$, QPSK modulation, equal power users with $E_b/N_0 =$ 6 dB, and $\alpha = 1.4$. The solid line with triangles denotes the GA-DE trajectory. Lozenges and circles denote simulation snapshots with EXT-based and APP-based IC, respectively, of a finite-dimensional system with spreading factor $L = 60$ and block length $N = 2000$.

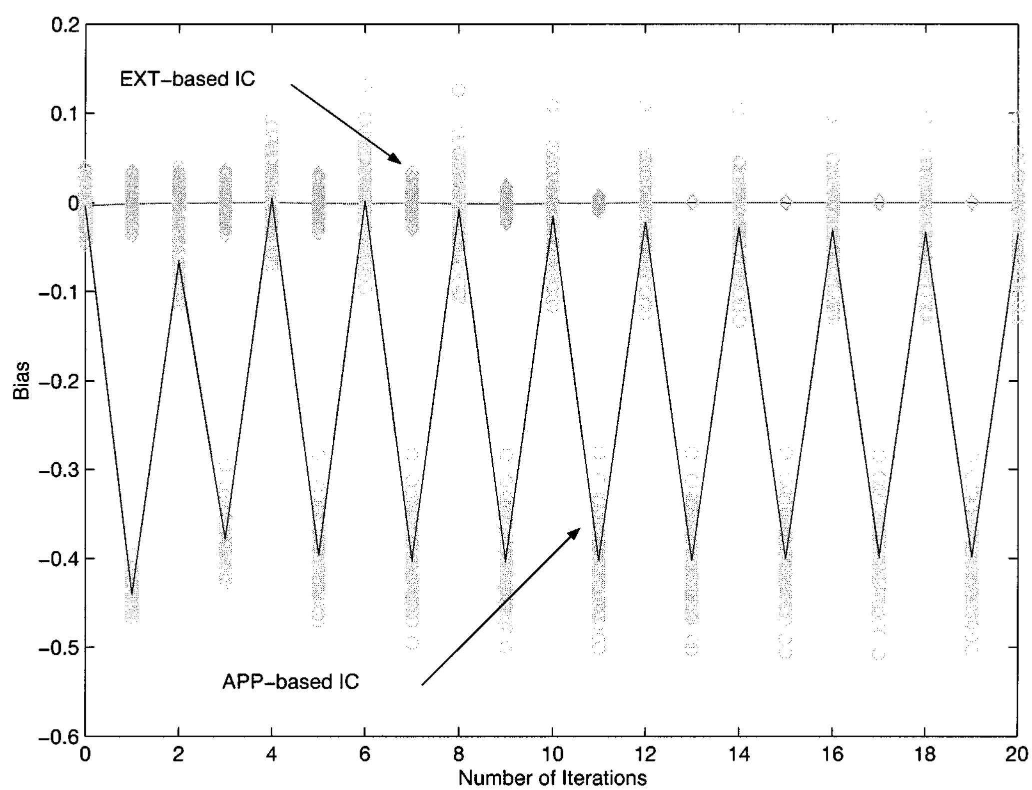

Fig. 9. Evolution of the residual interference conditional bias versus the number of iterations for SUMF-based soft IC decoding, with CC $(5, 7)$, QPSK modulation, equal power users with $E_b/N_0 = 6$ dB, and $\alpha = 1.4$. Lozenges and circles denote simulation snapshots with EXT-based and APP-based IC, respectively, of a finite-dimensional system with spreading factor $L = 60$ and block length $N = 2000$. Solid lines represent the trajectories of the bias mean values (averaged over the users).

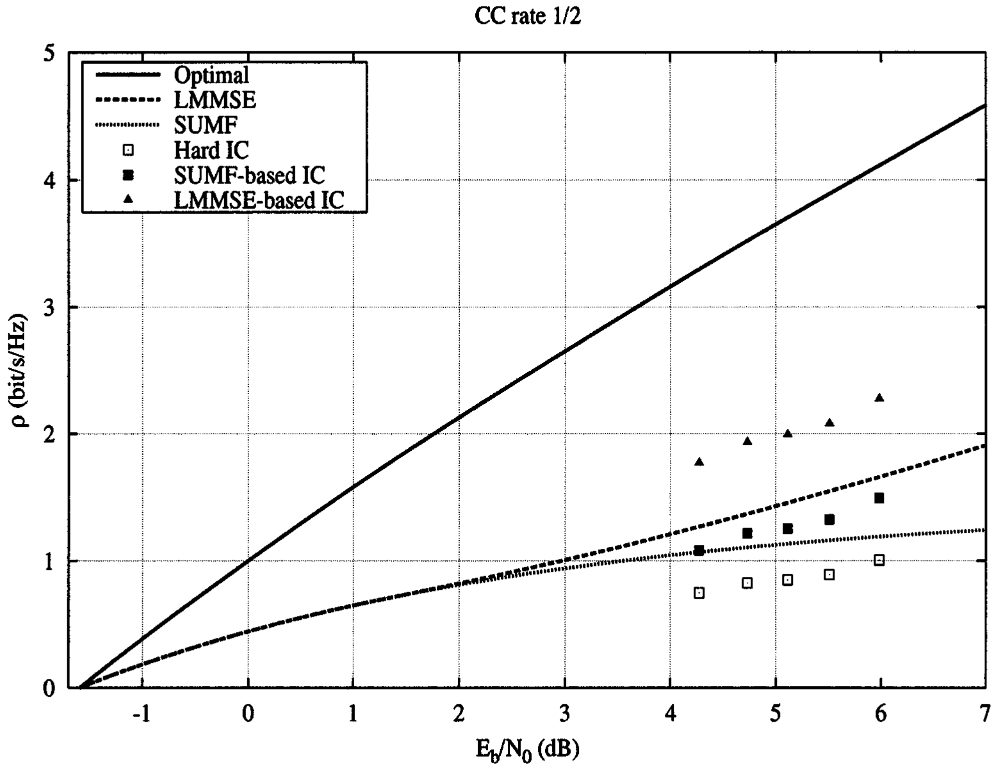

Fig. 10. Spectral efficiency of iterative IC decoders with equal power users and CCs of rate $1/2$ with 4, 8, 16, 32, and 64 states and QPSK modulation. The different code/decoder pairs are represented by points in the $(E_b/N_0,\, \rho)$ plane. The number of trellis states increases from right to left (i.e., the rightmost point corresponds to the four-state code and the leftmost point to the 64-state code). Solid, dashed, and dotted lines represent the achievable spectral efficiency of CDMA with optimal Gaussian random codes and joint decoding, linear MMSE front-end followed by single-user decoding, and SUMF followed by single-user decoding, respectively.

equal-power users. The different code/decoder pairs are represented by points in the $(E_b/N_0,\, \rho)$ plane. The number of trellis states increases from right to left (i.e., the rightmost point corresponds to the four-state code and the leftmost point to the 64-state code). Larger spectral efficiency is achieved by simpler codes, at the price of a larger required E_b/N_0 to achieve the target BER.

For the sake of comparison, the spectral efficiency with random spreading in a large-system regime achieved by linear detectors (SUMF and LMMSE) with single-user decoding and by the optimal joint decoder with ideal Gaussian random codes is also shown (adapted from [30]).

VI. Concluding Remarks

We would like to conclude by pointing out some recent results originated by this work and some suggestions for future research.

The GA-DE analysis proposed in this paper was recently used to study the performance of iterative IC decoding with LDPC codes in [63]. We expect that, by following in the footsteps of the concentration result provided here, a rigorous concentration result and the corresponding DE can be proved for several classes of user codes, and for systems where users make use of different codes with possibly different coding rates.

By using the GA-DE, the optimization of achievable spectral efficiency with respect to the SNR distribution $G^{(0)}(\gamma)$ is solved for binary CCs and QPSK modulation in [56]. Interestingly, it turns out that this problem is a linear program that can be solved very efficiently. In [56] and [63], it is also shown that a nonuniform user SNR distribution is the key to achieve large spectral efficiencies with iterative multiuser decoding. These results are in agreement with the conjectures made in [9], [10], based on information-theoretic arguments. Intuitively, the nonuniform SNR distribution shapes the multiaccess capacity region such that the transmitted rate K-tuple is close to a vertex (i.e., to a successively decodable point [2], [3]).

In [64], an improved iterative IC decoder based on widely linear filtering is proposed. Namely, it is recognized that the residual interference $\zeta_{k,\,n}^{(\ell)}$ conditioned with respect to the phases $\{\theta_{k,\,n}\}$ and EXT messages $\{P_{j,\,n}^{(\ell-1)}\colon j \neq k\}$ is not circularly symmetric. Therefore, a widely linear filter (see [65] and references therein) instead of the complex LMMSE filter (25) can be used in the IC-MUD nodes, providing better residual interference mitigation.

The framework developed in this paper can be extended to more general CDMA models, involving asynchronous transmission and multipath propagation. From the analysis point of view, DE may be extended by using the results of [57] to handle asynchronism and frequency-selective channels in the finite-dimensional case, and the results of [66] and [67] of to handle asynchronism and frequency-selective channels in the large-system limit.

Finally, iterative multiuser decoding can be naturally coupled with iterative channel estimation in the case of unknown user channels. First steps in this direction are reported in [46], [68],

[69]. A rigorous concentration result and the corresponding DE analysis of iterative joint data detection and channel estimation is still missing, and it represents an interesting area of research.

APPENDIX A
PROOFS

A.1 Proof of Lemma 1

The neighborhood $\mathcal{N}_{k,n}^{(\ell)}$ is cycle-free if and only if no IC-MUD node appears more than once in its node collection. The total number of IC-MUD nodes is $m = \frac{\mathcal{K}^{\ell+1}-1}{\mathcal{K}-1}$, where $\mathcal{K} = (K-1)((2W+1)s-1)$ was already introduced in (32). We enumerate the IC-MUD nodes in $\mathcal{N}_{k,n}^{(\ell)}$ from top to bottom and, inside each layer, from left to right (see Fig. 3), and define the sequence of nested neighborhoods $\mathcal{N}_{k,n}^{(\ell)}[i]$, for $i = 1, \ldots, m$ such that $\mathcal{N}_{k,n}^{(\ell)}[i]$ contains all IC-MUD nodes from 1 to i, and the edges connected to them. Then, $\mathcal{N}_{k,n}^{(\ell)}$ [1] is obviously cycle-free, and $\mathcal{N}_{k,n}^{(\ell)}[m] = \mathcal{N}_{k,n}^{(\ell)}$.

The events $\mathcal{E}_i = \{\mathcal{N}_{k,n}^{(\ell)}[i] \text{ is cycle-free}\}$ are nested, i.e., $\mathcal{E}_m \subseteq \mathcal{E}_{m-1} \subseteq \cdots \subseteq \mathcal{E}_1$. Hence,[10]

$$\overline{\mathcal{E}_m} = \bigcup_{i=2}^{m} \{\overline{\mathcal{E}_i} \cap \mathcal{E}_{i-1}\}$$

and, since the $\mathcal{E}_i$'s are nested, the events in the above union are disjoint. Therefore, we have

$$\begin{aligned}\Pr(\overline{\mathcal{E}_m}) &= \sum_{i=2}^{m} \Pr\left(\overline{\mathcal{E}_i} \cap \mathcal{E}_{i-1}\right)\\ &\leq \sum_{i=2}^{m} \Pr\left(\overline{\mathcal{E}_i} | \mathcal{E}_{i-1}\right)\\ &= \sum_{i=2}^{m} \frac{i-1}{N}\\ &= \frac{m(m-1)}{2N}\end{aligned} \tag{56}$$

where we have used the fact that, for a given set of distinct IC-MUD nodes of size $i-1$, the probability that a new randomly selected node out of the N possible nodes belongs to the set is $(i-1)/N$. This concludes the proof.

A.2 Proof of Lemma 2

Without loss of generality, we assume that the all-one sequence $\mathbf{1}$ is a codeword. Any geometrically uniform code $\mathcal{C}$ over $\mathcal{A}_M$ as defined in Section IV-B can be written as $\mathcal{C} = \{\boldsymbol{x} \in \mathbb{C}^N : \boldsymbol{x} = \boldsymbol{U}\mathbf{1}, \boldsymbol{U} \in G\}$, where G is a group of diagonal $N \times N$ matrices with diagonal elements $u_{j,j} \in \mathcal{A}_M$.

Let $\hat{\boldsymbol{x}} = \hat{\boldsymbol{U}}\mathbf{1}$ be the transmitted codeword, such that $\hat{u}_{n,n} = b$ for some $b \in \mathcal{A}_M$, and let $\boldsymbol{y} = \hat{\boldsymbol{x}} + \boldsymbol{\nu}$ be the channel output. Let $\mathcal{C}_n^{(a)}$ denote the subcode of all codewords $\boldsymbol{x} \in \mathcal{C}$ having symbol a in position n. The nth output message produced by the SISO decoder is given by

$$\begin{aligned}P_n(a) &\propto \sum_{\boldsymbol{x}\in\mathcal{C}_n^{(a)}} \prod_{j\neq n} p_\nu(y_j - x_j)\\ &\propto \sum_{\boldsymbol{U}\in H_n} \prod_{j\neq n} p_\nu\left(\nu_j + \hat{u}_{j,j} - u_{j,j}^{(a)} u_{j,j}\right)\end{aligned} \tag{57}$$

where H_n is the subgroup of all matrices $\boldsymbol{U} \in G$ having $u_{n,n} = 1$, and $\boldsymbol{U}^{(a)}$ is any matrix in G having $u_{n,n}^{(a)} = a$. We have the coset decomposition $G = \bigcup_{a\in\mathcal{A}_M} \boldsymbol{U}^{(a)} H_n$, and for all $a \in \mathcal{A}_M$ we can write

$$\mathcal{C}_n^{(a)} = \{\boldsymbol{x} = \boldsymbol{U}^{(a)}\boldsymbol{U}\mathbf{1} : \boldsymbol{U} \in H_n\}$$

for some coset leader $\boldsymbol{U}^{(a)} \in G$.

Let $P_n'(a)$ be the nth SISO output message when $\hat{\boldsymbol{x}}' = \hat{\boldsymbol{U}}'\mathbf{1}$ is the transmitted codeword, with $\hat{u}_{n,n}' = b'$ for some $b' \in \mathcal{A}_M$. We have

$$\begin{aligned}P_n'(a) &\propto \sum_{\boldsymbol{U}\in H_n} \prod_{j\neq n} p_\nu\left(\nu_j + \hat{u}_{j,j}' - u_{j,j}^{(a)} u_{j,j}\right)\\ &\overset{(a)}{\propto} \sum_{\boldsymbol{U}\in H_n} \prod_{j\neq n} p_{\hat{u}_{j,j}^* \hat{u}_{j,j}' \nu}\left(\nu_j + \hat{u}_{j,j}' - u_{j,j}^{(a)} u_{j,j}\right)\\ &\overset{(b)}{\propto} \sum_{\boldsymbol{U}\in H_n} \prod_{j\neq n} p_\nu\left(\nu_j' + \hat{u}_{j,j} - \hat{u}_{j,j}(\hat{u}_{j,j}')^* u_{j,j}^{(a)} u_{j,j}\right)\\ &\overset{(c)}{\propto} \sum_{\boldsymbol{U}\in H_n} \prod_{j\neq n} p_\nu\left(\nu_j' + \hat{u}_{j,j} - u_{j,j}^{(a')} u_{j,j}\right)\\ &\overset{(d)}{\sim} P_n(a')\end{aligned} \tag{58}$$

where (a) follows from the definition of $\mathcal{A}_M$-invariance of $p_\nu(z)$, (b) follows from the obvious relation

$$p_{c\nu}(z) = \frac{1}{|c|^2}\, p_\nu(c^{-1}z)$$

valid for any nonzero $c \in \mathbb{C}$, (c) follows by letting $\hat{\boldsymbol{U}}(\hat{\boldsymbol{U}}')^*\boldsymbol{U}^{(a)} = \boldsymbol{U}^{(a')}$ and by noticing that, since $(\hat{\boldsymbol{U}}')^* \in G$, then also $\boldsymbol{U}^{(a')} \in G$ and the coset $\boldsymbol{U}^{(a')}H_n$ generates the subcode $\mathcal{C}_n^{(a')}$, and (d) follows from (57), after noticing that $\nu_j' \triangleq \hat{u}_{j,j}(\hat{u}_{j,j}')^*\nu_j$ and ν_j are identically distributed.

We have shown that $\{P_n'(a) : a \in \mathcal{A}_M\}$ is distributed as a permuted version of $\{P_n(a) : a \in \mathcal{A}_M\}$, where the permutation $\pi^{-1} : a \mapsto a'$ is the cyclic shift defined by $a' = b(b')^* a$. This concludes the proof.

A.3 Proofs of Corollaries 1 and 2

Corollary 1 follows from Lemma 2 and from the fact that the residual interference at the output of any IC-MUD node in a cycle-free neighborhood is $\mathcal{A}_M$-invariant, since the Gaussian noise is circularly-symmetric and the complex amplitudes $w_{i,j}$ have i.i.d. phases $\theta_{i,j}$ uniformly distributed on $\{2\pi m/M : m = 0, \ldots, M-1\}$. Therefore, multiplying the residual interference term by any $a \in \mathcal{A}_M$ does not change its pdf.

Consider a trellis window T_i of the ith SISO decoder in the neighborhood $\mathcal{N}_{k,n}^{(\ell)}$, connected to the jth IC-MUD node, and let $\boldsymbol{x}^i$ be the ith user transmitted codeword. From Lemma 2 we have that the distribution of the EXT message produced by the trellis window T_i for symbol $x_{i,j}$ is the same for all codewords $\boldsymbol{x}'$ that coincide with $\boldsymbol{x}^i$ in the jth position. As a consequence, the statistics of residual interference at the output of the jth IC-MUD node depends only on the symbols labeling its input edges.

By definition, the collection of such symbols is the neighborhood type $\mathcal{T}(\mathcal{N}_{k,n}^{(\ell)})$. Since $\zeta_{k,n}^{(\ell)}$ is a function of the messages propagated through the neighborhood, we conclude that $\zeta_{k,n}^{(\ell)}$ is independent of $\boldsymbol{X}$ given $\mathcal{T}(\mathcal{N}_{k,n}^{(\ell)}) = \boldsymbol{t}$.

[10] The complement of an event $\mathcal{E}$ is denoted by $\overline{\mathcal{E}}$.

Corollary 2 is proved by iterating expectation. In the following, all expectations and probabilities are conditioned with respect to the event that $\mathcal{N}_{k,n}^{(\ell)}$ is cycle-free. We do not write it explicitly for the sake of notational simplicity, as it is obvious from the context. We have

$$\begin{aligned}
&\overline{F}_k^{(\ell)}(z|\boldsymbol{X})\\
&\quad = E\left[E\left[E\left[1\left\{\mathrm{Re}\left\{\zeta_{k,n}^{(\ell)}\right\}\le \mathrm{Re}\{z\},\right.\right.\right.\right.\\
&\qquad\qquad \left.\left.\left.\left.\mathrm{Im}\left\{\zeta_{k,n}^{(\ell)}\right\}\le \mathrm{Im}\{z\}\right\}\right|\boldsymbol{t},\boldsymbol{X},\boldsymbol{\Theta}\right]\right|\boldsymbol{t},\boldsymbol{X}\right]\right|\boldsymbol{X}\right]\\
&\quad \overset{(a)}{=} E\left[E\left[E\left[1\left\{\mathrm{Re}\left\{\zeta_{k,n}^{(\ell)}\right\}\le \mathrm{Re}\{z\},\right.\right.\right.\right.\\
&\qquad\qquad \left.\left.\left.\left.\mathrm{Im}\left\{\zeta_{k,n}^{(\ell)}\right\}\le \mathrm{Im}\{z\}\right\}\right|\boldsymbol{t},\boldsymbol{\Theta}\right]\right|\boldsymbol{t}\right]\right|\boldsymbol{X}\right]\\
&\quad \overset{(b)}{=} E\left[E\left[\overline{F}_k^{(\ell)}(z|\boldsymbol{t},\boldsymbol{\Theta})\right|\boldsymbol{t}\right]\right|\boldsymbol{X}\right]\\
&\quad \overset{(c)}{=} E\left[\frac{1}{M^{KN}}\sum_{\boldsymbol{\Theta}}\overline{F}_k^{(\ell)}(z|\boldsymbol{t},\boldsymbol{\Theta})\right|\boldsymbol{X}\right]\\
&\quad \overset{(d)}{=} E\left[\frac{1}{M^{KN}}\sum_{\boldsymbol{\Theta}}\overline{F}_k^{(\ell)}(z|\mathbf{1},\boldsymbol{\Theta})\right|\boldsymbol{X}\right]\\
&\quad \overset{(e)}{=} \frac{1}{M^{KN}}\sum_{\boldsymbol{\Theta}}\overline{F}_k^{(\ell)}(z|\mathbf{1},\boldsymbol{\Theta}) \triangleq \overline{F}_k^{(\ell)}(z)
\end{aligned} \tag{59}$$

where (a) follows from the fact that, by Corollary 1 and by the general fact that $E[E[Z|X,Y]|X] = E[E[Z|Y]|X]$ for the Markov chain $X \to Y \to Z$, (b) follows from defining

$$\overline{F}_k^{(\ell)}(z|\boldsymbol{t},\boldsymbol{\Theta}) \triangleq E\left[1\left\{\mathrm{Re}\left\{\zeta_{k,n}^{(\ell)}\right\}\le \mathrm{Re}\{z\},\ \mathrm{Im}\left\{\zeta_{k,n}^{(\ell)}\right\}\le \mathrm{Im}\{z\}\right\}\right|\boldsymbol{t},\boldsymbol{\Theta}\right]$$

(c) follows from the fact that $\boldsymbol{\Theta}$ is uniformly distributed over all possible phase sequences, (d) follows from the fact that for every $\boldsymbol{t} = \{x_{i,j}\}$ there exists $\boldsymbol{\Theta} = \{\theta_{i,j}\}$ such that $x_{i,j}e^{j\theta_{i,j}} = 1$, therefore, after summing over all possible $\boldsymbol{\Theta}$, the result is independent of $\boldsymbol{t}$. Finally, (e) follows from the fact that the argument in the expectation is deterministic and independent of $\boldsymbol{X}$.

A.4 *Proof of Proposition 1*

The proof of the concentration result stated in Proposition 1 follows closely the proofs given in [27], [57]. We include the main steps here for the sake of completeness.

Fix $\epsilon > 0$. For random variables $A, B, C \in \mathbb{R}$ we have

$$\begin{aligned}
&\Pr(|A-B|\ge\epsilon)\\
&\quad \overset{(a)}{\le} \Pr(|A-C|+|C-B|\ge\epsilon)\\
&\quad \le \Pr(\{|A-C|\ge\epsilon/2\}\cup\{|C-B|\ge\epsilon/2\})\\
&\quad \overset{(b)}{\le} \Pr(\{|A-C|\ge\epsilon/2\})+\Pr(\{|C-B|\ge\epsilon/2\})
\end{aligned} \tag{60}$$

where (a) follows from the triangular inequality and (b) from the union bound. We define

$$\overline{F}_k^{(\ell,N)}(z|\boldsymbol{X}) = E[F_k^{(\ell,N)}(z|\boldsymbol{X})|\boldsymbol{X}]$$

where expectation is with respect the random computational flowgraph ensemble $\{(\mathcal{G},\boldsymbol{\Theta})\}$ of block length N and with respect to the channel noise $\boldsymbol{N}$, conditioned on the transmitted codewords $\boldsymbol{X}$. From (60) we have

$$\begin{aligned}
&\Pr\left(\left|F_k^{(\ell,N)}(z|\boldsymbol{X})-\overline{F}_k^{(\ell)}(z)\right|\ge\epsilon\right)\\
&\quad \le \Pr\left(\left|F_k^{(\ell,N)}(z|\boldsymbol{X})-\overline{F}_k^{(\ell,N)}(z|\boldsymbol{X})\right|\ge\epsilon/2\right)\\
&\qquad + \Pr\left(\left|\overline{F}_k^{(\ell,N)}(z|\boldsymbol{X})-\overline{F}^{(\ell)}(z)\right|\ge\epsilon/2\right).
\end{aligned} \tag{61}$$

We show that for $N > 2c_1/\epsilon$, where c_1 is the constant appearing in Lemma 1, the second term in the RHS of (61) is equal to zero. In fact, let $\lambda = \Pr(\mathcal{N}_{k,1}^{(\ell)}$ is cycle-free), then

$$\begin{aligned}
&\overline{F}_k^{(\ell,N)}(z|\boldsymbol{X})\\
&\quad = \frac{1}{N}\sum_{n=1}^{N} E\left[1\left\{\mathrm{Re}\left\{\zeta_{k,n}^{(\ell)}\right\}\le\mathrm{Re}\{z\},\ \mathrm{Im}\left\{\zeta_{k,n}^{(\ell)}\right\}\le\mathrm{Im}\{z\}\right\}\right|\boldsymbol{X}\right]\\
&\quad \overset{(a)}{=} E\left[1\left\{\mathrm{Re}\left\{\zeta_{k,1}^{(\ell)}\right\}\le\mathrm{Re}\{z\},\ \mathrm{Im}\left\{\zeta_{k,1}^{(\ell)}\right\}\le\mathrm{Im}\{z\}\right\}\right|\boldsymbol{X}\right]\\
&\quad = \lambda E\left[1\left\{\mathrm{Re}\left\{\zeta_{k,1}^{(\ell)}\right\}\le\mathrm{Re}\{z\},\ \mathrm{Im}\left\{\zeta_{k,1}^{(\ell)}\right\}\le\mathrm{Im}\{z\}\right\}\right|\boldsymbol{X},\mathcal{N}_{k,1}^{(\ell)}\text{ is cycle-free}\right]\\
&\qquad + (1-\lambda)E\left[1\left\{\mathrm{Re}\left\{\zeta_{k,1}^{(\ell)}\right\}\le\mathrm{Re}\{z\},\ \mathrm{Im}\left\{\zeta_{k,1}^{(\ell)}\right\}\le\mathrm{Im}\{z\}\right\}\right|\boldsymbol{X},\mathcal{N}_{k,1}^{(\ell)}\text{ is not cycle-free}\right]
\end{aligned} \tag{62}$$

where (a) follows from symmetry. Since the indicator function is ≤ 1, by using Lemma 1 and recalling that, by definition

$$\overline{F}_k^{(\ell)}(z) = E\left[1\left\{\mathrm{Re}\left\{\zeta_{k,1}^{(\ell)}\right\}\le\mathrm{Re}\{z\},\ \mathrm{Im}\{\zeta_{k,1}^{(\ell)}\}\le\mathrm{Im}\{z\}\right\}\right|\boldsymbol{X},\mathcal{N}_{k,1}^{(\ell)}\text{ is cycle-free}\right]$$

which is independent of $\boldsymbol{X}$ by Corollary 2, we have that, for all $\boldsymbol{X}$

$$\begin{aligned}
\overline{F}_k^{(\ell,N)}(z|\boldsymbol{X}) &\le \overline{F}_k^{(\ell)}(z)+\frac{c_1}{N}\\
\overline{F}_k^{(\ell,N)}(z|\boldsymbol{X}) &\ge \overline{F}_k^{(\ell)}(z)\left(1-\frac{c_1}{N}\right)
\end{aligned} \tag{63}$$

which implies

$$\left|\overline{F}_k^{(\ell,N)}(z|\boldsymbol{X})-\overline{F}_k^{(\ell)}(z)\right|\le c_1/N.$$

Then, by choosing $N > 2c_1/\epsilon$ we have

$$\Pr\left(\left|\overline{F}_k^{(\ell,N)}(z|\boldsymbol{X})-\overline{F}_k^{(\ell)}(z)\right|\ge\epsilon/2\right)=0.$$

For the first term in the RHS of (61), for $\boldsymbol{X}$ fixed, we form an edges-phases-and-noise revealing Doob's martingale on the joint probability space $\{(\mathcal{G},\boldsymbol{\Theta},\boldsymbol{N})\}$ and apply Azuma's inequality (see [27] and references therein). For a given $\mathcal{G}$, we order its edges such that the ith edge in the list, for $i = 0,\ldots,KN-1$, corresponds to $\vec{e}_{j,n}$ with $j = [i]_{\mathrm{mod}\,K}+1$ and $n = \lfloor i/K\rfloor + 1$. Then, we expose the edges of $\mathcal{G}$ one by one, in the above order, and in the subsequent N steps we expose the phase randomization vector $\boldsymbol{\theta}_n$ and the noise vectors $\boldsymbol{\nu}_n$ corresponding to the nth IC-MUD node, for $n = 1,\ldots,N$.

We define a sequence of equivalence relations $\{\equiv_i\}$ inducing a sequence of nested partitions of $\{(\mathcal{G}, \boldsymbol{\Theta}, \boldsymbol{N})\}$ into equivalence classes as follows: $(\mathcal{G}', \boldsymbol{\Theta}', \boldsymbol{N}') \equiv_i (\mathcal{G}, \boldsymbol{\Theta}, \boldsymbol{N})$ if they coincide for the first i steps of the above edges-phases-and-noise revealing procedure. Then, for every fixed $z \in \mathbb{C}$ and $k \in \{1, \ldots, K\}$, the Doob's martingale is defined by

$$D_i = \sum_{n=1}^{N} E\left[1\left\{\mathrm{Re}\{\zeta'_{k,n}\} \le \mathrm{Re}\{z\}, \mathrm{Im}\{\zeta'_{k,n}\} \le \mathrm{Im}\{z\}\right\} \Big| \boldsymbol{X}, (\mathcal{G}', \boldsymbol{\Theta}', \boldsymbol{N}') \equiv_i (\mathcal{G}, \boldsymbol{\Theta}, \boldsymbol{N})\right] \quad (64)$$

for $i = 0, 1, \ldots, NK + N$, where $\zeta'_{k,n}$ is the residual interference at edge $\vec{e}_{k,n}$ of the graph $(\mathcal{G}', \boldsymbol{\Theta}', \boldsymbol{N}')$, after ℓ steps of the iterative IC decoder. Clearly

$$\frac{1}{N} D_0 = \overline{F}_k^{(\ell, N)}(z|\boldsymbol{X})$$

and

$$\frac{1}{N} D_{NK+N} = F_k^{(\ell, N)}(z|\boldsymbol{X}).$$

Since every edge participates in at most $n_t(\ell)$ neighborhoods of depth 2ℓ, the difference $|D_{i+1} - D_i|$ can be uniformly bounded by a constant κ that depends on K, W, s and ℓ, but it is independent of N.[11] Hence, Azuma's inequality yields

$$\Pr\left(|D_{NK+N} - D_0| \ge N\epsilon/2\right) \le 2e^{-\frac{(N\epsilon)^2}{8\sum_{i=1}^{NK+N}\kappa^2}}$$

which yields

$$\Pr\left(\left|F_k^{(\ell, N)}(z|\boldsymbol{X}) - \overline{F}_k^{(\ell, N)}(z|\boldsymbol{X})\right| \ge \epsilon/2\right) \le 2e^{-c_2\epsilon^2 N}$$

with $c_2 = \frac{1}{8(K+1)\kappa^2}$. This concludes the proof.

A.5 Proof of Proposition 2

Recall that we take the limits with respect to the block length N and to the number of users K and spreading factor L by letting first $N \to \infty$ and then $K, L \to \infty$, with fixed ratio $K/L = \alpha$. Since for all finite K, L, the block length N grows without bounds, by Proposition 1 the iterative multiuser IC decoder after ℓ iterations is completely characterized by the DE of Section IV-B, i.e., by the expected cdfs of the messages propagated through a randomly generated cycle-free neighborhood $\mathcal{N}_{k,n}^{(\ell)}$. Hence, we can prove Proposition 2 by induction on the DE iterations.

For the sake of brevity, we show Proposition 2 for the case of LMMSE-based algorithms (the cases of SUMF-based soft IC and hard-IC follow the same lines). The main results from the analysis of random CDMA in the large system limit, which will be used to prove Proposition 2, are summarized in the following result [29]–[31], [70].

Large-System Lemma: Consider a K-user CDMA system with spreading factor K defined by

$$\boldsymbol{y} = \boldsymbol{S}\boldsymbol{W}\boldsymbol{x} + \boldsymbol{\nu}$$

where $\boldsymbol{S}$ is random, with statistics defined in Section IV-C, $\boldsymbol{\nu} \sim \mathcal{N}_{\mathbb{C}}(\boldsymbol{0}, \boldsymbol{I})$ where the symbols x_k are independent, with mean zero, and covariance $E[\boldsymbol{x}\boldsymbol{x}^H] = \boldsymbol{I}$, and where $\boldsymbol{W} = \mathrm{diag}(\sqrt{\gamma_1}, \ldots, \sqrt{\gamma_K})$. Assume that the γ_k's are uniformly upper-bounded by a constant $\gamma_{\max}$ independent of K, and, as $K \to \infty$, the empirical cdf of the γ_k's converges weakly to a given cdf $G(\gamma)$.

Then, as $K, L \to \infty$ with fixed ratio α, the conditional distribution of the residual interference plus noise given $\boldsymbol{S}$ at the output of the kth user SUMF and LMMSE filters converges almost surely to a complex circularly symmetric Gaussian distribution independent of $\boldsymbol{S}$ and the SINR at the filter output converges almost surely to the nonrandom quantity $\gamma_k\eta$, where η is given by

$$\eta = \frac{1}{1 + \alpha\int_0^\infty \gamma\, dG(\gamma)}$$

for the SUMF and by the unique nonnegative solution of the equation

$$\eta = \frac{1}{1 + \alpha\int_0^\infty \frac{\gamma}{1+\eta\gamma}\, dG(\gamma)}$$

for the LMMSE filter. □

At step $\ell = 1$ of the DE recursion, the symbol estimates $\tilde{x}_{k,n}^{(0)}$ are all equal to zero. Then, because of the assumptions of Proposition 2, we find that the Large-System Lemma directly applies by letting $G(\gamma) = G^{(0)}(\gamma)$. Now, suppose that Proposition 2 holds for step $\ell - 1$. If Proposition 2 holds for iteration ℓ assuming that it holds for iteration $\ell - 1$, the proof follows by induction.

In order to show that Proposition 2 holds at step ℓ, it is sufficient to show that the empirical cdf of the interfering residual symbol instantaneous conditional variances $\gamma_k\xi_{k,n}^{(\ell-1)}$ converges weakly to a nonrandom cdf $G^{(\ell-1)}(u)$. For M-PSK, we have (see (20))

$$\xi_{k,n}^{(\ell-1)} = \sum_{a\in\mathcal{A}_M}\left|a - \tilde{x}_{k,n}^{(\ell-1)}\right|^2 P_{k,n}^{(\ell-1)}(a) = 1 - \left|\tilde{x}_{k,n}^{(\ell-1)}\right|^2.$$

The empirical cdf of $\gamma_k\xi_{k,n}^{(\ell-1)}$ is given by

$$G^{(\ell-1, K)}(u) \triangleq \frac{1}{K}\sum_{k=1}^{K} 1\left\{\gamma_k\left(1 - \left|\tilde{x}_{k,n}^{(\ell-1)}\right|^2\right) \le u\right\}. \quad (65)$$

Hence, the proof is complete if we show that the limit in (46) holds for step $\ell - 1$.

Since Proposition 2 is assumed true at step $\ell - 1$, the input of each kth SISO decoder at step $\ell - 1$ is of the form $z = x + \zeta$, where $\zeta \sim \mathcal{N}_{\mathbb{C}}(0, 1/(\gamma_k\eta^{(\ell-1)}))$. Moreover, since the neighborhood is cycle-free, the interference ζ in all inputs contributing to the same trellis window of the windowed BCJR decoder is an i.i.d. sequence. Finally, since ζ is rotationally invariant (and hence also $\mathcal{A}_M$-invariant), by Lemma 2 we incur no loss of generality by assuming $x_{k,n} = 1$.

In order to prove convergence of $G^{(\ell-1, K)}(u)$, we need to treat the SISO decoder as a mapping between (random) messages, and consider the associated mapping between the

[11] After noticing that $\xi'_{k,n}$ is a function of the variables appearing in the neighborhood $\mathcal{N}_{k,n}^{(\ell)}$, and that the maximum number of neighborhoods containing a given edge $\vec{e}_{k,n}$ is finite and independent of N, the explicit evaluation of κ is a straightforward exercise. We skip the details for the sake of brevity, since they are totally analogous to [27].

messages distributions. Let $\mathcal{P}$ be the space of all pmfs over $\mathcal{A}_M$. Then, a SISO decoder can be seen as a sequence of mappings

$$M_n\colon \underbrace{\mathcal{P}\times\cdots\times\mathcal{P}}_{N\text{ times}} \to \mathcal{P}, \qquad n=1,\ldots,N \tag{66}$$

for $n=1,\ldots,N$. For the exact sum-product SISO decoder, $M_n(Q_1,\ldots,Q_N)=P_n$ is given by (11). However, (66) is more general, allowing for approximated implementations of the exact SISO decoder. For example, the windowed BCJR decoder takes the form (66), where the mappings M_n act on a sliding window of the input message sequence $(Q_1,\ldots,Q_N)$ centered on symbol n and comprising $2W+1$ trellis sections.

A pmf in $\mathcal{P}$ can be seen as a real M-dimensional vector $\boldsymbol{p}$ lying in the simplex defined by $\sum_{i=1}^M p_i=1$, $\boldsymbol{p}\in\mathbb{R}_+^M$. For a given transmitted codeword, the SISO decoder input and output messages are random vectors in $\mathcal{P}$. We denote by $\mathcal{F}$ the space of all joint cdfs over $\mathcal{P}$. When the SISO decoder input is produced by a memoryless time-invariant channel, the input messages $(Q_1,\ldots,Q_N)$ form an independent sequence conditionally on the transmitted codeword. We define the message cdf mapping associated to the SISO decoder as

$$\mathcal{M}_n\colon \underbrace{\mathcal{F}\times\cdots\times\mathcal{F}}_{N\text{ times}} \to \mathcal{F}, \qquad n=1,\ldots,N \tag{67}$$

such that if

$$(Q_1,\ldots,Q_N)\sim\prod_{i=1}^N F_i(\boldsymbol{q}_i)$$

then $P_n\sim F_n'(\boldsymbol{p})=\mathcal{M}_n(F_1,\ldots,F_N)$.

We have the following definition [71].

Definition (Unisotropy Degree): Consider a geometrically uniform code $\mathcal{C}$ over $\mathcal{A}_M$ and (without loss of generality) assume that the all-ones sequence is a codeword. Consider the transmission of the all-ones codeword over a complex circularly symmetric AWGN channel. The unisotropy degree d of $\mathcal{C}$ with respect to the SISO decoder given by $\{M_n\colon n=1,\ldots,N\}$ is defined as the number of distinct output message cdfs generated by the associated mappings $\{\mathcal{M}_n\colon n=1,\ldots,N\}$. $\Diamond$

When $\mathcal{C}$ is a trellis code with trellis termination and the SISO decoder is the windowed BCJR algorithm, the unisotropy degree is finite for any N. In fact, trellis termination affects a finite number (say, $2m$) of trellis sections (m at the beginning and m at the end of the block), therefore, only a finite number $2s(m+W)$ of output messages are affected by the asymmetry due to trellis termination. For all the remaining $\max\{N-2s(m+W),0\}$ messages, the SISO decoder behaves as if the trellis were infinite, and due to the fact that the trellis is time-invariant with s symbols per trellis section, the SISO decoder can produce at most s distinct cdfs. By letting $N\to\infty$, the fraction $2s(m+W)/N$ of symbols affected by trellis termination vanishes. Hence, in the limit for large N we get that the unisotropy degree of $\mathcal{C}$ with respect to the windowed BCJR decoder is upper-bounded by s.

For simplicity, we assume that the user amplitudes can take on values in a discrete and finite set of values $0\le\overline{\gamma}_1\le\cdots\le\overline{\gamma}_m=\gamma_{\max}$, for some integer m, and let

$$\delta_h=G^{(0)}(\overline{\gamma}_h)-\lim_{\gamma\uparrow\overline{\gamma}_h}G^{(0)}(\gamma)$$

be the limiting fraction of users transmitted with squared complex amplitude $\overline{\gamma}_h$, for $h=1,\ldots,m$.[12] Consider the K symbols transmitted at time n, for an arbitrary choice of $n\in\{1,\ldots,N\}$. Let $\mathcal{K}_{h,r}$ denote the subset of such symbols transmitted with signal squared amplitude $\overline{\gamma}_h$ and appearing in position r of their branch label in the corresponding trellis section of the basic code $\mathcal{C}$. Then, we can write

$$G^{(\ell-1,K)}(u) = \frac{1}{K}\sum_{h=1}^m\sum_{r=1}^s\sum_{k\in\mathcal{K}_{h,r}} 1\left\{\overline{\gamma}_h\left(1-\left|\tilde{x}_{k,n}^{(\ell-1)}\right|^2\right)\le u\right\}. \tag{68}$$

The soft symbol estimates $\tilde{x}_{k,n}^{(\ell-1)}$ for all $k\in\mathcal{K}_{h,r}$ are i.i.d., since they are produced by the same SISO decoder mapping with identically distributed input (recall that the unisotropy degree of $\mathcal{C}$ is at most s).

From the strong law of large numbers [60], the fraction of user symbols $x_{k,n}$ belonging to subset $\mathcal{K}_{h,r}$ converges almost surely to δ_h/s, since for a random selection of the neighborhood $\mathcal{N}_{k,n}^{(\ell)}$ a symbol transmitted at time n occupies position $1\le r\le s$ in its trellis branch label with equal probability. Hence, the following limit holds almost surely:

$$\begin{aligned}
G^{(\ell-1)}(u) &= \lim_{K\to\infty}\frac{1}{K}\sum_{h=1}^m\sum_{r=1}^s\sum_{k\in\mathcal{K}_{h,r}} 1\left\{\overline{\gamma}_h\left(1-\left|\tilde{x}_{k,n}^{(\ell-1)}\right|^2\right)\le u\right\}\\
&= \lim_{K\to\infty}\sum_{h=1}^m\sum_{r=1}^s\frac{|\mathcal{K}_{h,r}|}{K}\frac{1}{|\mathcal{K}_{h,r}|}\sum_{k\in\mathcal{K}_{h,r}}\\
&\quad\cdot 1\left\{\overline{\gamma}_h\left(1-\left|\tilde{x}_{k,n}^{(\ell-1)}\right|^2\right)\le u\right\}\\
&= \sum_{h=1}^m\frac{\delta_h}{s}\sum_{r=1}^s \Pr\left(\overline{\gamma}_h\left(1-\left|\tilde{x}_{k,n}^{(\ell-1)}\right|^2\right)\le u\,\middle|\,k\in\mathcal{K}_{h,r}\right)\\
&= \sum_{h=1}^m\delta_h H_h(u)
\end{aligned} \tag{69}$$

where we define the cdf

$$H_h(z)=\frac{1}{s}\sum_{r=1}^s\Pr\left(\overline{\gamma}_h\left(1-\left|\tilde{x}_{k,n}^{(\ell-1)}\right|^2\right)\le u\,\middle|\,k\in\mathcal{K}_{h,r}\right).$$

This concludes the proof.

Notice that the fact that the basic code $\mathcal{C}$ has *finite* unisotropy degree is crucial. In fact, if the number of different output message distributions generated by the SISO decoder grows linearly with the block length N, the convergence (69) is not ensured any longer.

[12]The extension to a continuous limiting distribution $G^{(0)}(u)$ is a straightforward but tedious exercise, and follows from standard continuity arguments.

A.6 *Proof of Corollary 3*

Define the mapping function $\psi(\eta)$ such that $\psi: \eta^{(\ell-1)} \mapsto \eta^{(\ell)}$ coincides with (43) or with (45), depending on the IC algorithm considered. All DE fixed points must satisfy the equation $\psi(\eta) = \eta$ for some $\eta \in [0, 1]$. By inspection, we have that the mapping $\psi(\cdot)$, defined on the domain $[0, 1]$, is continuous and has range $[0, m]$, with $m \leq 1$. Hence, the equation $\psi(\eta) = \eta$ has at least one solution in $[0, 1]$. Moreover, we have that $\psi(0) > 0$. Therefore, $\eta^\star = \min\{\eta \in [0, 1]: \psi(\eta) = \eta\}$ is a stable fixed point and the iterated map ψ with initial condition $\eta^{(0)} = 0$ converges to the limit $\eta^\star$. From Proposition 2, it follows that the DE recursion with initial condition $\eta^{(0)} = 0$ has limit

$$\lim_{\ell\to\infty} \overline{F}_k^{(\ell)}(z) = \mathcal{N}_{\mathbb{C}}(0, 1/(\gamma_k \eta^\star))$$

for all $k = 1, \ldots, K$, that is uniquely identified by $\eta^\star$.

APPENDIX B
QPSK WITH GRAY MAPPING AND BIT INTERLEAVING

The QPSK signal set $\mathcal{Q} = \{\pm\frac{1}{\sqrt{2}} \pm \frac{j}{\sqrt{2}}\}$ is naturally matched to the binary field $\mathbb{F}_2$ via the Gray mapping $\mu: \mathbb{F}_2 \times \mathbb{F}_2 \to \mathcal{Q}$ such that

$$\mu(c_1, c_2) = ((1-2c_1) + j(1-2c_2))/\sqrt{2}.$$

In CDMA applications, the case of binary convolutional codes Gray-mapped onto QPSK is particularly important [36]. In this appendix, we provide the details of the calculation of the DE-GA used to generate the results of Section V.

The Gray-mapped QPSK does not fit directly the theory developed for M-PSK and linear trellis codes over the ring $\mathbb{Z}_M$, since the additive group of $\mathbb{F}_2 \times \mathbb{F}_2$ is not isomorphic to the cyclic group of order 4. However, the results of Section IV can be extended to this case by considering binary convolutional codes with *bit interleaving* [72]. Namely, we shall assume that each user code $\mathcal{C}_k$ is obtained by concatenating the same "basic" binary convolutional code $\mathcal{C}$ with a randomly and independently selected bit interleaver, and that the QPSK modulation with Gray mapping is applied to the resulting bit-interleaved binary sequence. Finally, phase randomization is applied independently on the real and imaginary parts of the modulated symbols, so that the resulting CDMA channel model is given by

$$\boldsymbol{y}_n = \frac{1}{\sqrt{2}} \boldsymbol{S}\boldsymbol{A}\left(\boldsymbol{\Phi}_n^{(I)}\boldsymbol{a}_n + \boldsymbol{\Phi}_n^{(Q)}\boldsymbol{b}_n\right) + \boldsymbol{\nu}_n \tag{70}$$

where $\boldsymbol{a}_n, \boldsymbol{b}_n \in \{\pm 1\}^K$ contain the in-phase (I) and quadrature (Q) components of the transmitted symbols, $\boldsymbol{A} = \mathrm{diag}(\sqrt{\gamma_1}, \ldots, \sqrt{\gamma_K})$ is a real diagonal matrix of amplitudes, and $\boldsymbol{\Phi}_n^{(I)}, \boldsymbol{\Phi}_n^{(Q)}$ are diagonal matrices with diagonal elements in $\{\pm 1\}$, representing the I and Q phase randomization sequences. We shall refer to the model (70) as the "I&Q" channel.

With bit-interleaving, the I and Q components of the residual interference contributing to the input messages to any trellis window in a cycle-free neighborhood $\mathcal{N}_{k,n}^{(\ell)}$ are independent. The concentration result of Section IV-B extends immediately to the I&Q channel, and it is easy to show that, in the large-system limit, the residual interference at the input of the SISO decoder of user k at iteration ℓ is complex Gaussian, with identically distributed I and Q components, with mean zero and per-component variance $1/(2\gamma_k\eta^{(\ell)})$, where $\eta^{(\ell)}$ is given by Proposition 2 (the proof of these facts is a straightforward exercise by following the proofs of Proposition 1 and 2).

The SISO decoders produce output messages in the form of EXT log-likelihood ratios

$$\mathcal{L}_{k,n} = \log \frac{P_{k,n}(+1)}{P_{k,n}(-1)}$$

for the coded binary symbols (not for the modulated QPSK symbols!). From now on, we drop the user and time indexes as there is no risk of confusion. Let $e_s(\mathsf{SNR})$ denote the SER versus SNR function of $\mathcal{C}$ transmitted with Gray mapping and QPSK modulation over AWGN, were decision about the coded symbol a is made according to the rule

$$\tilde{x} = \mathrm{sign}(\mathcal{L}_x). \tag{71}$$

($\mathcal{L}_x$ is a short-hand notation to indicate the EXT log-likelihood ratio relative to a certain coded symbol x.) As in Section IV-C, we want to approximate $\eta^{(\ell)}$ defined in Proposition 2 directly in terms of the SER at iteration $\ell - 1$, that for a user received at SNR level γ is given by $\epsilon = e_s(\gamma\eta^{(\ell-1)})$.

We denote by a and b the I and Q binary antipodal symbols relative to a Gray-mapped QPSK symbol, and by $\widetilde{a}$ and $\widetilde{b}$ their estimate produced by the SISO decoder. For the hard-IC decoder (49) becomes

$$E\left[\left|\frac{1}{\sqrt{2}}(a-\tilde{a}) + \frac{j}{\sqrt{2}}(b-\tilde{b})\right|^2\right] = 2 - (1-\epsilon) + \epsilon = 4\epsilon. \tag{72}$$

Hence, (50) reduces to

$$\eta^{(\ell)} = \frac{1}{1 + 4\alpha\int_0^\infty \gamma e_s\left(\gamma\eta^{(\ell-1)}\right)\, dG^{(0)}(\gamma)} \tag{73}$$

initialized by $\eta^{(0)} = 0$.

For soft IC, by using (20) with binary antipodal symbols, we obtain the soft estimates of the I and Q symbols as

$$\tilde{a} = \tanh(\mathcal{L}_a/2), \qquad \tilde{b} = \tanh(\mathcal{L}_b/2).$$

Equation (52) reduces to

$$\begin{aligned} &E\left[\left|\frac{1}{\sqrt{2}}(a-\tilde{a}) + \frac{j}{\sqrt{2}}(b-\tilde{b})\right|^2\right] \\ &\quad= E\left[(1-\tanh(\mathcal{L}_a))^2 | a=+1\right] \\ &\quad= E\left[\left.\frac{4}{(1+e^{\mathcal{L}_a})^2}\right| a=+1\right]. \end{aligned} \tag{74}$$

As in (51), we make a Gaussian approximation of the SISO decoder output and let [32], [33]

$$\mathcal{L}_x \sim \mathcal{N}(\mu, 2\mu) \tag{75}$$

for $x = +1$, where μ is chosen to match the SER. By using (75) in (71), we get $\epsilon = Q(\sqrt{\mu/2})$, where

$$Q(x) \triangleq \int_x^\infty \frac{1}{\sqrt{2\pi}} e^{-z^2/2}\, dz$$

which yields

$$\mu = \mu(\epsilon) \triangleq 2\left(Q^{-1}(\epsilon)\right)^2. \tag{76}$$

Hence, (74) gives the function

$$V(\epsilon) \triangleq \int \frac{4}{(1+e^{\lambda})^2}\, h(\lambda|\epsilon)\, d\lambda \tag{77}$$

where, from (75) and (76), the log-likelihood pdf is given by

$$h(\lambda|\epsilon) \triangleq \frac{1}{\sqrt{4\pi\mu(\epsilon)}}\, e^{-\frac{|\lambda-\mu(\epsilon)|^2}{4\mu(\epsilon)}}. \tag{78}$$

The function $V(\epsilon)$ in (77) can be used in the GA-DE recursion (53) for the SUMF-based decoder.

For the LMMSE-based decoder, we have

$$\begin{aligned} 1-\left|\frac{\tilde{a}}{\sqrt{2}}+\frac{j\tilde{b}}{\sqrt{2}}\right|^2 &= \frac{1}{2}\left(1-|\tanh(\mathcal{L}_a/2)|^2\right) \\ &\quad + \frac{1}{2}\left(1-|\tanh(\mathcal{L}_b/2)|^2\right) \\ &= \frac{2e^{\mathcal{L}_a}}{(1+e^{\mathcal{L}_a})^2}+\frac{2e^{\mathcal{L}_b}}{(1+e^{\mathcal{L}_b})^2}. \end{aligned} \tag{79}$$

Then, after a change of variable and some algebra, the inner integral (with respect to u) in (54) can be written as

$$\iint \frac{4\gamma(1+e^{\lambda_2})^2}{(1+e^{\lambda_1})^2(1+e^{\lambda_2})^2+2\eta\gamma\left(e^{\lambda_2}(1+e^{\lambda_1})^2+e^{\lambda_1}(1+e^{\lambda_2})^2\right)} \cdot h(\lambda_1|\epsilon)h(\lambda_2|\epsilon)\, d\lambda_1\, d\lambda_2 \tag{80}$$

where (80) is obtained by using repeatedly the identity $h(\lambda|\epsilon)e^{-\lambda} = h(-\lambda|\epsilon)$, a symmetry condition of log-likelihood ratios that holds for any binary-input symmetric-output memoryless channel [73].

Integrals (77) and (80) can be easily and accurately computed by using Gauss–Hermite quadratures, and make DE-GA computationally much simpler than exact DE.

ACKNOWLEDGMENT

The authors would like to thank Dr. Ralf Müller and Dr. Hesham El Gamal for the helpful discussions about the topics of this paper.

REFERENCES

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

[2] T. Cover and J. Thomas, *Elements of Information Theory*. New York: Wiley, 1991.

[3] B. Rimoldi and R. Urbanke, "A rate splitting approach to the Gaussian multipleaccess channel," *IEEE Trans. Inform. Theory*, vol. 42, pp. 364–375, Mar. 1996.

[4] M. Varanasi and T. Guess, "Optimum decision feedback multiuser equalization with successive decoding achieves the total capacity of the Gaussian multiple access channel," in *Proc. Asilomar Conf.*, Pacific Grove, CA, Nov. 1997.

[5] P. Alexander, A. Grant, and M. Reed, "Iterative detection in code-division multiple access with error control coding," *European Trans. Telecommun.*, vol. 9, no. 5, pp. 419–425, Sept. 1999.

[6] T. Aulin and R. Espineira, "Trellis coded multiple access (TCMA)," in *Proc. Int. Communications Conf. 1999*, Vancouver, BC, Canada, June 1999.

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

[8] L. Brunel, "Algorithmes de decodage de canal pour l'acces multiple à etalement de spectre," Ph.D. dissertation, Ecole Nationale Supérieure des Télécommunications (ENST), Paris, France, 1999.

[9] N. Chayat and S. Shamai (Shitz), "Iterative soft onion peeling for multiaccess and broadcast channels," in *Proc. PIMRC'98*, Boston, MA, Sept. 1998.

[10] ——, "Convergence properties of iterative soft onion peeling," in *Proc. IEEE Information Theory Workshop (ITW 1999)*, Kruger Nat. Park, South Africa, June 1999, p. 9.

[11] T. Giallorenzi and S. Wilson, "Multiuser ML sequence estimation for convolutional coded asynchronous DS-CDMA systems," *IEEE Trans. Commun.*, vol. 44, pp. 997–1008, Aug. 1996.

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

[13] R. Ratasuk and M. L Honig, "Large system error probability of multiuser decision feedback receivers," in *Proc. IEEE Int. Symp. Information Theory (ISIT 2000}*, Sorrento, Italy, June 2000, p. 385.

[14] M. L. Honig and Y. Sun, "Performance of iterative multiuser decision-feedback receivers," in *Proc. IEEE Information Theory Workshop (ITW 2001)*, Cairns, Australia, Sept. 2001, p. 30.

[15] M. Moher, "An iterative multiuser decoder for near-capacity communications," *IEEE Trans. Commun.*, vol. 46, pp. 870–880, July 1998.

[16] M. Reed, C. Schlegel, P. Alexander, and J. Asenstorfer, "Iterative multiuser detection for CDMA with FEC: Near single-user performance," *IEEE Trans. Commun.*, vol. 46, pp. 1693–1699, Dec. 1998.

[17] C. Schlegel, "Joint detection in multiuser systems via iterative processing," in *Proc. IEEE Int. Symp. Information Theory (ISIT 2000)*, Sorrento, Italy, June 2000, p. 274.

[18] Z. Shi and C. Schlegel, "Joint decoding of serially concatenated coded CDMA: Iterative schedule study," in *Proc. IEEE Information Theory Workshop (ITW 2001)*, Cairns, Australia, Sept. 2001, p. 58.

[19] F. Tarkoy, "Iterative multiuser decoding for asynchronous users," in *Proc. IEEE Int. Symp. Information Theory {ISIT '97)*, Ulm, Germany, July 1997, p. 30.

[20] X. Wang and V. Poor, "Iterative (turbo) soft interference cancellation and decoding for coded CDMA," *IEEE Trans. Commun.*, vol. 47, pp. 1047–1061, July 1999.

[21] H. El Gamal and E. Geraniotis, "Iterative multiuser detection for coded CDMA signals in ANVGN and fading channels," *IEEE J. Select. Areas Commun.*, vol. 18, pp. 30–41, Jan. 2000.

[22] F. Kschischang, B. Frey, and H.-A. Loeliger, "Factor graphs and the sum- product algorithm," *IEEE Trans. Inform. Theory*, vol. 47, pp. 498–519, Feb. 2001.

[23] R. M. Tanner, "A recursive approach to low complexity codes," *IEEE Trans. Inform. Theory*, vol. IT-27, Sept. 1981.

[24] J. Pearl, *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*. San Francisco, CA: Morgan Kaufmann, 1988.

[25] N. Wiberg, H. Loeliger, and R. Kotter, "Codes and iterative decoding on general graphs," *European Trans. Telecomm.*, vol. 6, Sept. 1995.

[26] S. Benedetto, D. Divsalar, G. Montorsi, and F. Pollara, "Soft-input soft-output building blocks for the construction of distributed iterative decoding of code networks," *European Trans. Commun.*, vol. 9, no. 2, Apr. 1998.

[27] T. Richardson and R. Urbanke, "The capacity of low-density parity check codes under message passing decoding," *IEEE Trans. Inform. Theory*, vol. 47, pp. 599–618, Feb. 2001.

[28] "Special Issue on Codes on Graphs Iterative Algorithms," *IEEE Trans. Inform. Theory*, vol. 47, pp. 493–853, Feb. 2001.

[29] D. Tse and S. Hanly, "Linear multiuser receivers: Effective interference, effective bandwidth and capacity," *IEEE Trans. Inform. Theory*, vol. 45, pp. 641–675, Mar. 1999.

[30] S. Verdú and S. Shamai (Shitz), "Spectral efficiency of CDMA with random spreading," *IEEE Trans. Inform. Theory*, vol. 45, pp. 622–640, Mar. 1999.

[31] ——, "The impact of frequency-flat fading on the spectral efficiency of CDMA," *IEEE Trans. Inform. Theory*, vol. 47, pp. 1302–1327, May 2001.

[32] H. El Gamal and R. Hammons, "Analyzing the turbo decoder using the Gaussian approximation," *IEEE Trans. Inform. Theory*, vol. 47, pp. 671–686, Feb. 2001.

[33] S. Y. Chung, T. Richardson, and R. Urbanke, "Analysis of sum-product decoding of low-density parity-check codes using a Gaussian approximation," *IEEE Trans. Inform. Theory*, vol. 47, pp. 657–671, Feb. 2001.

[34] S. Ten Brink, "Convergence behavior of iteratively decoded parallel concatenated codes," *IEEE Trans. Commun.*, vol. 49, pp. 1727–1737, Oct. 2001.

[35] F. Neeser and J. Massey, "Proper complex random processes with applications to information theory," *IEEE Trans. Inform. Theory*, vol. 39, pp. 1293–1302, July 1993.

[36] 3GGP, "TS 25.224 V3.1.0, '3GPP-TSG-RAN-WG1; Physical layer procedures (FDD)'," ETSI, Dec. 1999.

[37] J. Proakis, *Digital Communications*, 3rd ed. New York: McGraw-Hill, 1995.

[38] C. Berrou and A. Glavieux, "Near optimum error-correcting coding and decoding: Turbo codes," *IEEE Trans. Commun.*, vol. 44, Oct. 1996.

[39] H. Jin, A. Khandeka, and R. McEliece, "Irregular repeat-accumulate codes," in *Proc. 2nd Int. Symp. Turbo Codes and Related Topics*, Brest, France, Sept. 2000, pp. 1–8.

[40] J. Hagenauer, E. Offer, and L. Papke, "Iterative decoding of binary block and convolutional codes," *IEEE Trans. Inform. Theory*, vol. 42, pp. 429–445, Mar. 1996.

[41] L. Bahl, J. Cocke, F. Jelinek, and J. Raviv, "Optimal decoding of linear codes for minimizing symbol error rate," *IEEE Trans. Inform. Theory*, vol. IT-20, pp. 284–287, Mar. 1974.

[42] G. Woodward, M. L. Honig, and P. Alexander, "Adaptive multiuser parallel decision feedback with iterative decoding," in *Proc. IEEE Int. Symp. Information Theory (ISIT 2000)*, Sorrento, Italy, June 2000, p. 335.

[43] M. Damen, "Joint coding/decoding in a multiple access system: Applications to mobile communications," Ph.D. dissertation, Ecole Nationale Supérieure des Télécommunications (ENST), Paris, France, 1999.

[44] M. Buehrer, S. Nicoloso, and S. Gollamudi, "Linear versus nonlinear interference cancellation," *J. Commun. and Networks*, vol. 1, no. 2, pp. 118–133, June 1999.

[45] S. Marinkovic, B. Vucetic, and J. Evans, "Improved iterative parallel interference cancellation," in *Proc. IEEE Int. Symp. Information Theory (ISIT 2001)*, Washington, DC, June 2001, p. 34.

[46] P. Alexander and A. Grant, "Iterative channel and information sequence estimation in CDMA," in *Proc. IEEE 6th Int. Symp. Spread-Spectrum Technology and Applications (ISSSTA 2000)*. Newark, NJ, Sept. 2000.

[47] ——, "Iterative decoding and channel estimation," in *Proc. IEEE Int. Symp. Information Theory (ISIT 2000)*, Sorrento, Italy, June 2000, p. 171.

[48] M. Varanasi and B. Aazhang, "Multistage detection in asynchronous code-division multiple-access communications," *IEEE Trans. Commun.*, vol. 38, pp. 509–519, Apr. 1990.

[49] D. Divsalar, M. Simon, and D. Raphaeli, "Improved parallel interference cancellation for CIDMA," *IEEE Trans. Commun.*, vol. 46, pp. 258–268, Feb. 1998.

[50] S. Moshavi, E. Kanterakis, and D. Schilling, "Multistage linear receivers for DS-CDMA systems," *Int. J. Wireless Inform. Networks*, vol. 3, no. 1, pp. 1–17, Jan. 1996.

[51] P. Tan and L. Rasmussen, "Linear interference cancellation in CDMA based on iterative techniques for linear equation systems," *IEEE Trans. Commun.*, vol. 48, pp. 2099–2108, Dec. 2000.

[52] P. Tan, L. Rasmussen, and T. Lim, "Constrained maximum-likelihood detection in CDMA," *IEEE Trans. Commun.*, vol. 49, pp. 142–153, Jan. 2001.

[53] R. Müller and S. Verdú, "Design and analysis of low-complexity interference mitigation on vector channels," *IEEE J. Select. Areas Commun.*, vol. 19, pp. 1429–1441, Aug. 2001.

[54] M. L. Honig and W. Xiao, "Performance of reduced-rank linear interference suppression for DS-CDMA," *IEEE Trans. Inform. Theory*, vol. 47, pp. 1928–1946, July 2001.

[55] L. Nelson and V. Poor, "Iterative multiuser receivers for CDMA channels: An EM-based approach," *IEEE Trans. Commun.*, vol. 44, pp. 1700–1710, Dec. 1996.

[56] G. Caire and R. Müller, "The optimal received power distribution of IC-based iterative multiuser joint decoders," in *Proc. 39th Annu. Allerton Conf. on Communications, Control and Computing*, Monticello, IL, Oct. 2001.

[57] A. Kavčić, B. Marcus, M. Mitzenmacher, and B. Wilson, "Deriving performance bounds for ISI channels using Gallager codes," in *Proc. IEEE Int. Symp. Information Theory (ISIT 2001)*, Washington, DC, June 2001, p. 345.

[58] D. Forney, Jr., "Geometrically uniform codes," *IEEE Trans. Inform. Theory*, vol. 37, pp. 1241–1260, Sept. 1991.

[59] S. Benedetto, D. Divsalar, G. Montorsi, and F. Pollara, "Algorithms for continuous decoding of turbo codes," *Electron. Lett.*, vol. 32, pp. 314–315, Feb. 1996.

[60] P. Billingsley, *Probability and Measure*. New York: Wiley, 1986.

[61] Y. Kuznetsov, *Elements of Applied Bifurcation Theory*, ser. Applied Mathematical Science. New York: Springer-Verlag, 1998, vol. 112.

[62] E. Biglieri, J. Proakis, and S. Shamai (Shitz), "Fading channels: Information-theoretic and communications aspects," *IEEE Trans. Inform. Theory*, vol. 44, pp. 2619–2692, Oct. 1998.

[63] H. El Gamal and D. Hamon, "On the user capacity of turbo multiuser decoders," in *Proc. 39th Annu. Allerton Conf. Communication, Control and Computing*, Monticello, IL, Oct. 2001.

[64] A. Lampe, R. Shober, and W. Gerstacker, "Iterative multiuser detection for complex modulation schemes," in *Proc. IEEE Int. Symp. Information Theory (ISIT 2001)*, Washington, DC, June 2001, p. 33.

[65] A. Tulino and S. Verdú, "Asymptotic analysis of improved linear receivers for BPSK-CDMA subject to fading," *IEEE J. Select. Areas Commun.*, vol. 19, pp. 1544–1555, Aug. 2001.

[66] Kiran and D. N. C. Tse, "Effective interference and effective bandwidth of linear multiuser receivers in asynchronous systems," *IEEE Trans. Inform. Theory*, vol. 46, pp. 1426–1447, July 2000.

[67] J. Evans and D.N. C. Tse, "Large system performance of linear multiuser receivers in multipath fading channels," *IEEE Trans. Inform. Theory*, vol. 46, pp. 2059–2078, Sept. 2000.

[68] M. Kobayashi, J. Boutros, and G. Caire, "Successive interference cancellation with SISO decoding and EM channel estimation," *IEEE J. Select. Areas Commun.*, vol. 19, pp. 1450–1460, Aug. 2001.

[69] G. Caire, A. Tulino, and E. Biglieri, "Iterative multiuser joint decoding and parameter estimation: A factor-graph approach," in *Proc. IEEE Information Theory Workshop (ITW 2001)*, Cairns, Australia, Sept. 2001, p. 36.

[70] E. Chong, J. Zhang, and D. Tse, "Output MAI distribution of linear MMSE multiuser receivers in DS-CDMA systems," *IEEE Trans. Inform. Theory*, vol. 47, pp. 1128–1144, Mar. 2001.

[71] S. Vialle and J. Boutros, "Performance limits of concatenated codes with iterative decoding," in *Proc. IEEE Int. Symp. Information Theory (ISIT 2000)*, Sorrento, Italy, June 1997, p. 150.

[72] G. Caire, G. Taricco, and E. Biglieri, "Bit-interleaved coded modulation," *IEEE Trans. Inform. Theory*, vol. 44, pp. 927–946, May 1998.

[73] T. Richardson and R. Urbanke, "Design of capacity-approaching irregular low-density parity-check codes," *IEEE Trans. Inform. Theory*, vol. 47, pp. 619–637, Feb. 2001.