first (or $\beta S_E + (1-\beta)S_{CD}$ with user 2 decoded first), where α (or β) ranges from 0 to 1. These rates are achievable, so they are lower bounds. Because the objective is a concave function of the covariance matrices, this lower bound is better than the time-sharing of data rates associated with *B* and *C* (or *D* and *E*). Since the corner points after one iteration (i.e., *B* and *E*) are at most $(K-1)m/2$ nats away from the sum capacity, the lower bound is a close approximation of the capacity region. A typical example is shown in Fig. 5. Extensive numerical simulations show that the lower bound is fairly tight. An upper bound is also plotted by extending the line segments *AB*, *CD*, and *EF*. This is an upper bound because the capacity region is convex.

VI. Conclusion

This correspondence addresses the problem of finding the optimal transmitter covariance matrices that achieve the sum capacity in a Gaussian vector multiple-access channel. The computation of the sum capacity is formulated in a convex optimization framework. A multiuser water-filling condition for achieving the sum capacity is found. It is shown that the sum-rate maximization problem can be solved efficiently using an iterative water-filling algorithm, where each step of the iteration is equivalent to a local maximization of one user's data rate with multiuser interference treated as noise. The iterative water-filling algorithm is shown to converge to the sum capacity from any starting point. The convergence is fast. In particular, it reaches within 1/2 nats per user per output dimension from the sum capacity after just a single iteration. As mentioned before, the vector channel model discussed in this correspondence includes ISI channels and fading channels as special cases. Thus, the iterative water-filling algorithm can also be used to efficiently compute the power allocation across the frequency spectrum for an ISI channel or over time for a fading channel.

Finally, although the iterative water-filling algorithm solves the sum capacity problem efficiently, it does not directly apply to other points in the capacity region. The computations of other capacity points are also convex programming problems. However, how to best exploit the problem structure in these cases is still not clear.

References

[1] R. Ahlswede, "Multi-way communication channels," in *Proc. 2nd Int. Symp. Information Theory*, 1971, pp. 103–105.

[2] H. Liao, "Multiple access channels," Ph.D. dissertation, Univ. Hawaii, 1972.

[3] R. G. Gallager, "Energy limited channels: Coding, multiaccess and spread spectrum," in *Proc. Conf. Information Science and Systems (CISS)*, Mar. 1988, p. 372.

[4] S. Verdú, "On channel capacity per unit cost," *IEEE Trans. Inform. Theory*, vol. 36, pp. 1019–1030, Sept. 1990.

[5] R. G. Gallager, "A perspective on multiaccess channels," *IEEE Trans. Inform. Theory*, vol. IT-31, pp. 124–142, Mar. 1985.

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

[7] R. S. Cheng and S. Verdú, "Gaussian multiaccess channels with ISI: Capacity region and multiuser water-filling," *IEEE Trans. Inform. Theory*, vol. 39, pp. 773–785, May 1993.

[8] R. Knopp and P. A. Humblet, "Information capacity and power control in single-cell multi-user communications," in *Proc. IEEE Int. Conf. Communications (ICC)*, 1995, pp. 331–335.

[9] D. N. C. Tse and S. Hanly, "Multi-access fading channels: Part I: Polymatroid structure, optimal resource allocation and throughput capacities," *IEEE Trans. Inform. Theory*, vol. 44, pp. 2796–2815, Nov. 1998.

[10] W. Yu, "Competition and Cooperation in multi-user communication environments," Ph.D. dissertation, Stanford Univ., Stanford, CA, 2002.

[11] L. Vandenberghe, S. Boyd, and S.-P. Wu, "Determinant maximization with linear matrix inequality constraints," *SIAM J. Matrix Anal. Applic.*, vol. 19, pp. 499–533, 1998.

[12] P. Viswanath, D. N. C. Tse, and V. Anantharam, "Asymptotically optimal water-filling in vector multiple-access channels," *IEEE Trans. Inform. Theory*, vol. 47, pp. 241–67, Jan. 2001.

[13] P. Viswanath and V. Anantharam, "Optimal sequences and sum capacity of synchronous CDMA systems," *IEEE Trans. Inform. Theory*, vol. 45, pp. 1984–1991, Sept. 1999.

[14] S. Ulukus and R. D. Yates, "Iterative signature adaptation for capacity maximization of CDMA systems," in *Proc. Allerton Conf. Commications, Control and Computing*, 1998.

[15] C. Rose, S. Ulukus, and R. D. Yates, "Wireless systems and interference avoidance," *IEEE Trans. Wireless Commun.*, vol. 1, pp. 415–428, July 2002.

[16] M. Médard, "The effect upon channel capacity in wireless communications of perfect and imperfect knowledge of the channel," *IEEE Trans. Inform. Theory*, vol. 46, pp. 935–946, May 2000.

[17] R. A. Horn and C. R. Johnson, *Matrix Analysis*. Cambridge, U.K.: Cambridge Univ. Press, 1990.

[18] S. Boyd and L. Vandenberghe, "Convex Optimization," Cambridge Univ. Press, Cambridge, U.K., available at http://www.stanford.edu/~boyd/cvxbook.html, 2003.

[19] L. Vandenberghe and S. Boyd, "Semidefinite programming," *SIAM Rev.*, vol. 38, no. 1, pp. 49–95, Mar. 1996.

Maximizing the Spectral Efficiency of Coded CDMA Under Successive Decoding

Giuseppe Caire, *Senior Member, IEEE*,
Souad Guemghar, *Student Member, IEEE*,
Aline Roumy, *Member, IEEE*, and Sergio Verdú, *Fellow, IEEE*

***Abstract*—We investigate the spectral efficiency achievable by random synchronous code-division multiple access (CDMA) with quaternary phase-shift keying (QPSK) modulation and binary error-control codes, in the large system limit where the number of users, the spreading factor, and the code block length go to infinity. For given codes, we maximize spectral efficiency assuming a minimum mean-square error (MMSE) successive stripping decoder for the cases of equal rate and equal power users. In both cases, the maximization of spectral efficiency can be formulated as a linear program and admits a simple closed-form solution that can be readily interpreted in terms of power and rate control. We provide examples of the proposed optimization methods based on off-the-shelf low-density parity-check (LDPC) codes and we investigate by simulation the performance of practical systems with finite code block length.**

***Index Terms*—Channel capacity, code-division multiple access (CDMA), low-density parity-check (LDPC) codes, multiuser detection, quaternary phase-shift keying (QPSK) modulation, successive decoding.**

I. Introduction

All points in the capacity region of the scalar Gaussian multiple-access channel are achievable by successive single-user encoding,

Manuscript received September 11, 2002; revised August 22, 2003. The material in this correspondence was presented at the 39th Allerton Conference, Monticello, IL, October 2001.

G. Caire and S. Guemghar are with the Eurecom Institute, 06904 Sophia-Antipolis, France (e-mail: Giuseppe.Caire@eurecom.fr; Souad.Guemghar@eurecom.fr).

A. Roumy is with IRISA-INRIA, 35042 Rennes, France (e-mail: aline.roumy@irisa.fr).

S. Verdú is with the Department of Electrical Engineering, Princeton University, Princeton, NJ 08544 USA (e-mail: verdu@princeton.edu).

Communicated by V. V. Veeravalli, Associate Editor for Detection and Estimation.

Digital Object Identifier 10.1109/TIT.2003.821970

decoding, and interference cancellation (stripping) [1], [2]. The Cover–Wyner capacity region was generalized in [3] to encompass the case of code-division multiple access (CDMA) with arbitrary signature waveforms. As in the scalar case, all points in the boundary of the capacity region of the CDMA channel can also be achieved by stripping and single-user encoding/decoding, as long as the stripping decoders incorporate minimum mean-square error (MMSE) filters against yet undecoded users at each successive cancellation stage [4]. Key to the optimality of stripping is the use of Gaussian codes of rate arbitrarily close to (but not larger) than the capacity of the channel obtained by removing the already decoded users. In this way, optimal spectral efficiency is achieved by simple single-user coding and decoding, with linear complexity in the number of users.

For given signature waveforms, successive stripping generally requires that every user must transmit at a different rate, or must be received at a different signal-to-noise ratio (SNR) level. This can be avoided by designing the signature waveforms such that the equal-rate point coincides with a vertex of the equal-power capacity region. However, optimizing the signature waveforms (e.g., [5], [6]) is highly impractical in real-life applications, where transmission is usually affected by frequency selective fading channels.

On the other hand, existing nonorthogonal CDMA systems [7], [8] are largely based on pseudorandom waveforms. The maximum spectral efficiency of randomly spread (synchronous) CDMA, in the large system limit, where the number of users and the spreading factor grow without bound while their ratio tends to a constant β, was found in the case of power-constrained inputs in [9], [10], and in the case of binary antipodal inputs in [11].

While in the power-constrained case capacity is achieved by Gaussian inputs, practical systems make use of discrete small-size modulation alphabets. Given its widespread application and the fact that quaternary phase-shift keying (QPSK) is optimal in the wide-band low-SNR regime [12] we shall restrict our analysis to QPSK-modulated CDMA.

We consider a pragmatic approach to QPSK-modulated CDMA based on applying single-user binary coding and the same stripping decoding approach which would be optimal for Gaussian codes. Moreover, for evident practical reasons, we constrain our system to have only a finite number of coding rates and/or of received SNR levels. For this setting, we compute the achievable spectral efficiency in the large system regime with optimal (i.e., single-user capacity achieving) binary codes and with the best known low-density parity-check (LDPC) code ensembles [13], in the limit for large code block length, in the cases of equal received SNRs or equal rate users. The proposed equal-power and equal-rate design approaches can be effectively applied to nonasymptotic code block length, and provide a simple tool to dimension CDMA systems for given target bit-error rate (BER), user codes, and desired spectral efficiency.

The rest of the correspondence is organized as follows. Section II presents the basic synchronous CDMA additive white Gaussian noise (AWGN) model where users are grouped into a finite number of classes such that users in a given class have the same rate and received SNR. The existing results on the optimum spectral efficiency of the power-constrained CDMA channel are summarized in Section III with particular emphasis on the finite class model. Our choice for the input constellation is QPSK, justified on the basis of complexity and asymptotic optimality, as shown in Section IV-A. Then, we consider the optimization of the received power profile of the different classes in Section IV-B, assuming that the users employ equal-rate codes. Conversely, in Section IV-C, we optimize the code rate profile assuming equal received SNR for all users. Both problems are formulated as linear programs whose solution can be found explicitly. Section V presents numerical examples of both methods when the user codes are irregular LDPC codes found in [13]. Simulation results for finite block length and finite number of users validate the large-system infinite block length assumption made in the proposed optimization methods.

II. Synchronous CDMA Canonical Model

We consider the complex baseband discrete-time channel model

$$\boldsymbol{y}_i = \boldsymbol{S}\boldsymbol{x}_i + \boldsymbol{n}_i, \qquad i = 1, \dots, n \tag{1}$$

originated by sampling at the chip-rate a synchronous CDMA system [14], where

1) $\boldsymbol{y}_i, \boldsymbol{n}_i \in \mathbb{C}^n$ are the vector of received chip-rate samples and the corresponding AWGN samples $\sim \mathcal{N}_{\mathbb{C}}(0,1)$ received at time i;
2) $\boldsymbol{S} \in \mathbb{C}^{N\times K}$ contains the user spreading sequences by columns. Spreading sequences are proper complex, known to the receiver, with independent and identically distributed (i.i.d.) chips with zero mean, variance $1/N$, and finite fourth-order moment;
3) $\boldsymbol{x}_i \in \mathbb{C}^K$ is the vector of user modulation symbols transmitted at time i, where its kth component $x_{k,i}$ takes on values in some signal constellation, with given average energy per symbol $E[|x_{i,k}|^2] = \alpha_k \mathsf{SNR}$. The scaling factors α_k represent *power control* and, without loss of generality, are normalized such that
$$\frac{1}{K}\sum_{k=1}^{K}\alpha_k = 1;$$
4) N, K, and n denote the spreading factor, the number of users, and the code block length, respectively.

For the purpose of system design, it is convenient to consider a system formed by J user classes. The size of class j is K_j, and we denote by $\beta_j = K_j/N$ the "class load" of class j. Thus, the total channel load is

$$\beta = \sum_{j=1}^{J}\beta_j \text{ users per chip.}$$

Users in class j have the same SNR, denoted by γ_j (i.e., $\alpha_k\mathsf{SNR} = \gamma_j$ for all users k in class j). Without loss of generality, we assume $\gamma_1 \le \cdots \le \gamma_J$. The total system spectral efficiency is given by

$$\rho = \sum_{j=1}^{J}\beta_j R_j$$

where R_j denotes the average rate of users in class j. The users individual E_b/N_0's are in general different. Nevertheless, for the sake of comparison with a reference equal-rate equal-power system, it is convenient to define a "system" E_b/N_0 by

$$\left(\frac{E_b}{N_0}\right)_{\text{sys}} \triangleq \frac{\sum_{j=1}^{J}\beta_j\gamma_j}{\sum_{j=1}^{J}\beta_j R_j} \tag{2}$$

which coincides with the individual E_b/N_0's in the case where users are dynamically assigned to the classes so that each user belongs to class j for a fraction β_j/β of the time.

III. Existing Results on Fundamental Limits

In [10], the spectral efficiency (in bit/s/Hz) of random CDMA in the large system limit ($K, N \to \infty$ with $K/N = \beta$) subject to fading with an input power constraint is found to be

$$\mathsf{C}(\boldsymbol{\beta},\boldsymbol{\gamma}) = \mathsf{C}^{\text{mmse}}(\boldsymbol{\beta},\boldsymbol{\gamma}) + \log_2\frac{1}{\eta} + (\eta - 1)\log_2 e \tag{3}$$

where $\boldsymbol{\beta} \triangleq (\beta_1, \dots, \beta_J)$ and $\boldsymbol{\gamma} \triangleq (\gamma_1, \dots, \gamma_J)$, η is the large-system multiuser efficiency [14] of the linear MMSE receiver, given by the solution of the Tse–Hanly equation [15], which for later use we write as

$$\eta = f_J(\eta, \beta_J) \tag{4}$$

where we define

$$f_j(\eta, z) \triangleq \left(1 + z\frac{\gamma_j}{1+\gamma_j \eta} + \sum_{i=1}^{j-1} \beta_i \frac{\gamma_i}{1+\gamma_i\eta}\right)^{-1} \tag{5}$$

and where $\mathsf{C}^{\mathrm{mmse}}(\boldsymbol{\beta}, \boldsymbol{\gamma})$ is the achievable spectral efficiency of a system based on linear MMSE filtering followed by single-user decoding, given by

$$\mathsf{C}^{\mathrm{mmse}}(\boldsymbol{\beta}, \boldsymbol{\gamma}) = \sum_{j=1}^{J} \beta_j \log_2(1+\gamma_j \eta). \tag{6}$$

The spectral efficiencies in (3) and in (6) are achieved with codes whose empirical distributions are Gaussian.

As shown in [10], the supremum of (3) over all possible J, $\boldsymbol{\beta}$, $\boldsymbol{\gamma}$ (for a fixed E_b/N_0, and β) is achieved by $J = 1$ (one class only). This can be readily seen by writing the total spectral efficiency for finite K, N and given $\boldsymbol{S}$ as $\frac{1}{N}\log_2 \det(\boldsymbol{I} + \mathsf{SNR}\,\boldsymbol{S}\mathsf{A}^2\boldsymbol{S}^H)$, where $\mathsf{A} \triangleq \mathrm{diag}(\sqrt{\alpha_1}, \dots, \sqrt{\alpha_K})$. Then, we notice that, for $K/N = \beta$ and assuming that the empirical distribution of the scaling factors α_k converges to some fixed (nonrandom) distribution as $K \to \infty$, the limit

$$\lim_{K\to\infty} \frac{1}{N}\log_2 \det(\boldsymbol{I} + \mathsf{SNR}\,\boldsymbol{S}\mathsf{A}^2\boldsymbol{S}^H) = \lim_{K\to\infty} \frac{1}{N} E\left[\log_2 \det(\boldsymbol{I} + \mathsf{SNR}\,\boldsymbol{S}\mathsf{A}^2\boldsymbol{S}^H)\right]$$

holds with probability 1 [10]. Finally, by averaging over all $K \times K$ permutation matrices $\boldsymbol{\Pi}$, by noticing that $\boldsymbol{S}$ and $\boldsymbol{S\Pi}$ are identically distributed, and by using Jensen's inequality, it follows from the concavity of $\log\det(\cdot)$ on the cone of nonnegative definite Hermitian symmetric matrices that

$$\begin{aligned} &E[\log_2 \det(\boldsymbol{I} + \mathsf{SNR}\,\boldsymbol{S}\mathsf{A}^2\boldsymbol{S}^H)] \\ &= \frac{1}{K!}\sum_{\boldsymbol{\Pi}} E\left[\log_2 \det(\boldsymbol{I} + \mathsf{SNR}\,\boldsymbol{S\Pi}\mathsf{A}^2\boldsymbol{\Pi}^H\boldsymbol{S}^H)\right] \\ &\le E\left[\log_2 \det\left(\boldsymbol{I} + \frac{\mathsf{SNR}}{K!}\sum_{\boldsymbol{\Pi}} \boldsymbol{S\Pi}\mathsf{A}^2\boldsymbol{\Pi}^H\boldsymbol{S}^H\right)\right] \\ &= E[\log_2 \det(\boldsymbol{I} + \mathsf{SNR}\,\boldsymbol{S}\boldsymbol{S}^H)] \end{aligned}$$

where the last line is achieved when all users are received with the same SNR.

The supremum over β is achieved for $\beta \to \infty$, and coincides with the AWGN single-user capacity $\mathsf{C}^\star$, implicitly given by

$$\frac{2^{\mathsf{C}^\star} - 1}{\mathsf{C}^\star} = \frac{E_b}{N_0}. \tag{7}$$

The spectral efficiency $\mathsf{C}(\boldsymbol{\beta}, \gamma)$ can be achieved by single-user decoding with successive stripping and MMSE filtering against undecoded users. Suppose that users are decoded one by one, starting from users in class J, then class $J-1$, and so on. Then, $\mathsf{C}(\boldsymbol{\beta}, \boldsymbol{\gamma})$ can be written as

$$\mathsf{C}(\boldsymbol{\beta}, \boldsymbol{\gamma}) = \sum_{j=1}^{J} \int_0^{\beta_j} \log_2(1+\gamma_j \eta_j(z))dz \tag{8}$$

where $\eta_j(z)$ is the solution to $\eta = f_j(\eta, z)$. Notice that stripping of the users one by one implies that users in the same class have different rates. Namely, the user decoded in position $\lfloor K_j z/\beta_j \rfloor$ of class j (where $z \in [0, \beta_j]$), transmits at rate $\log_2(1+\gamma_j\eta_j(z))$. Hence, in general, a different rate is required for each user. This makes such system highly impractical from the implementation point of view.

IV. Approaching the Optimal Spectral Efficiency

A. QPSK Input Constellations

Information theory teaches us that one way to approach (3) is to use single-user capacity approaching codes for the AWGN channel, successive interference cancellation, and MMSE filtering at each cancellation stage [4]. Furthermore, substantial progress has been made in the last few years in designing binary codes and low-complexity decoders whose rate comes fairly close to single-user capacity at vanishing BER. Among those modern codes are turbo codes, repeat–accumulate (RA) codes, and LDPC codes, all of which are decoded by efficient iterative techniques (see the special issue [16] and references therein). These code ensembles are characterized by their rate–threshold pair (R, g), such that for SNR $\ge g$ the BER can be made arbitrarily small in the limit of $n \to \infty$. For carefully optimized code ensembles [17]–[20], on the standard single-user AWGN channel, the rate–threshold pairs achieved so far come remarkably close to the curve $R = C_{\mathrm{qpsk}}(\mathsf{SNR})$, where

$$C_{\mathrm{qpsk}}(\mathsf{SNR}) = 2\left(1 - \int_{-\infty}^{\infty} \log_2\left(1 + e^{-2\mathsf{SNR} - 2\sqrt{\mathsf{SNR}}v}\right)\frac{e^{-v^2/2}}{\sqrt{2\pi}}dv\right) \tag{9}$$

is the QPSK-input AWGN channel capacity, as a function of SNR. For example, Fig. 1 shows the QPSK capacity (9) (solid curve) and rate–threshold pairs (marks) corresponding to some LDPC code ensembles from [13].

In the following, we shall optimize a CDMA system under successive decoding assuming error-free decoding at each decoding level when the decoder operates above its threshold SNR. This corresponds to assuming very large code block length (i.e., $n \to \infty$). Our goal is to find the vectors $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$ so that, at each stripping decoder stage, the threshold requirement of each single-user decoder is satisfied. We shall refer to this condition as the successive decodability condition.

In the large system limit, under our system assumptions, it is well known that the residual interference at the output of the MMSE filter at any cancellation stage is complex Gaussian with circular symmetry [21]. Assuming optimal QPSK codes characterized by the rate–threshold pairs $(R, C_{\mathrm{qpsk}}^{-1}(R))$, for $R \in [0, 2]$ (see Fig. 1), the spectral efficiency achieved by a stripping decoder is given by

$$\mathsf{C}_{\mathrm{qpsk}}(\boldsymbol{\beta}, \boldsymbol{\gamma}) = \sum_{j=1}^{J} \int_0^{\beta_j} C_{\mathrm{qpsk}}(\gamma_j \eta_j(z))dz. \tag{10}$$

Fig. 2 shows $\mathsf{C}_{\mathrm{qpsk}}(\beta, \gamma)$ and $\mathsf{C}(\beta, \gamma)$ (for a single-class system, i.e., $J = 1$) versus β, for $E_b/N_0 = 3$ and 10 dB. The corresponding AWGN capacity $\mathsf{C}^\star$ is shown for comparison. We notice that the loss incurred by QPSK codes with respect to Gaussian codes gets more pronounced as E_b/N_0 increases. Although for any fixed E_b/N_0 and sufficiently large β the loss vanishes, for high E_b/N_0 exceedingly large values of β are required to make the loss negligible.

The following result shows that as the system load grows without bound, QPSK suffers no loss of optimality. The result applies to the general case where the received user SNRs are given by $\mathsf{SNR}|\mathsf{A}_k|^2$, under the mild requirement that, as $K \to \infty$, the empirical distribution of the user amplitudes $|\mathsf{A}_k|$ converges to a given nonrandom distribution $F_{|\mathsf{A}|}$. For $|\mathsf{A}_k|^2 = \alpha_k$ (deterministic) we obtain the system model defined in Section III, but the result applies also to the case where $|\mathsf{A}_k|^2$ represents the random fading coefficient of user k, as in [10].[1]

[1]As in all capacity results of CDMA in the large system limit (see, for example, [9], [10], [15]), achievability implies that the user code block length goes to infinity. The natural and meaningful order of the limits with respect to n and K is: *first* let $n \to \infty$ and *then* let $K \to \infty$. This means that the large-system spectral efficiency is the limit of the spectral efficiencies of systems with increasing K.

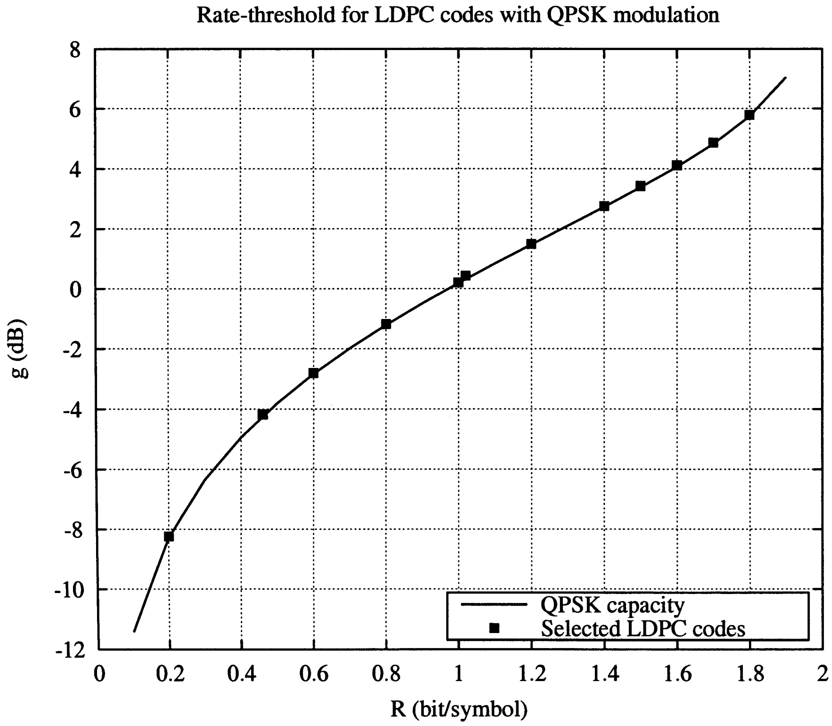

Fig. 1. Rate–threshold pairs corresponding to QPSK capacity and for some LDPC codes from [13].

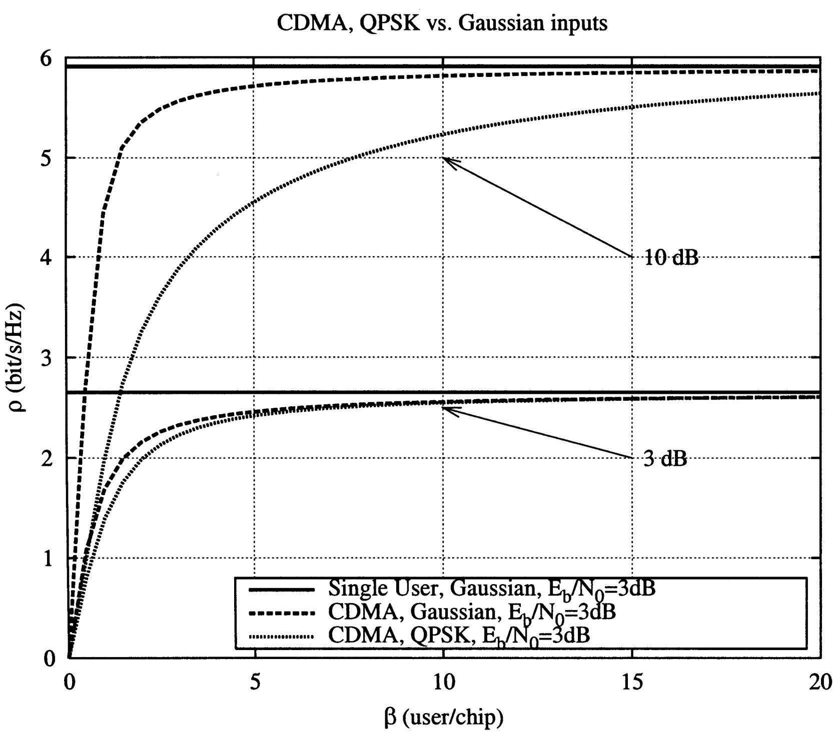

Fig. 2. Spectral efficiency versus β for random CDMA with Gaussian and QPSK inputs (with stripping decoder).

Theorem 1: Let

$$C_Q(\beta, \mathsf{SNR}) = \int_0^\beta E[C_{\rm qpsk}(|\mathsf{A}|^2 \mathsf{SNR}\eta(z, \mathsf{SNR}))]dz \tag{11}$$

where $\eta(z, \mathsf{SNR})$ is the solution to

$$\eta + ZE\left[\frac{\mathsf{SNR}|\mathsf{A}|^2\eta}{1+\mathsf{SNR}|\mathsf{A}|^2\eta}\right] = 1 \tag{12}$$

where $|\mathsf{A}| \sim F_{|\mathsf{A}|}$.

Fix β and $\frac{E_b}{N_0}$ and define

$$\mathsf{C}_Q\left(\beta, \frac{E_b}{N_0}\right) = C_Q(\beta, \mathsf{SNR}) \tag{13}$$

for the SNR satisfying

$$\frac{E_b}{N_0} C_Q(\beta, \mathsf{SNR}) = \beta \mathsf{SNR}. \tag{14}$$

Then, for all $\frac{E_b}{N_0} \geq \log_e 2$

$$\lim_{\beta\to\infty} \mathsf{C}_Q\left(\beta, \frac{E_b}{N_0}\right) = \mathsf{C}^\star. \tag{15}$$

Proof: In view of the result shown in [10, eq. (163)] and since the use of QPSK cannot improve upon the result obtained with Gaussian inputs with the same power, it is enough to show that

$$\lim_{\beta\to\infty} \mathsf{C}_Q\left(\beta, \frac{E_b}{N_0}\right) \geq \mathsf{C}^\star \tag{16}$$

where $\mathsf{C}^\star$ is given by (7) for any $E_b/N_0 > \log_e 2$ (notice that $\mathsf{C}^\star(\log_e 2) = 0$).

To show (16), we will show that for every β and SNR

$$C_Q(\beta, \mathsf{SNR}) \geq \log_2(1+\beta\mathsf{SNR}) - \frac{\kappa(|\mathsf{A}|)\beta\mathsf{SNR}^2}{\beta\mathsf{SNR}+1}\log_2 e \tag{17}$$

where

$$\kappa(|\mathsf{A}|) = \frac{E[|\mathsf{A}|^4]}{(E[|\mathsf{A}|^2])^2}$$

denotes the kurtosis of the distribution of $|\mathsf{A}|$.

The bound in (17) will be sufficient for our purposes because if we choose the following SNR:

$$\mathsf{SNR}_\beta = \frac{\mathsf{C}^\star}{\beta}\frac{E_b}{N_0} \tag{18}$$

then

$$\begin{aligned} C_Q(\beta, \mathsf{SNR}_\beta) &\geq \log_2(1+\beta\mathsf{SNR}_\beta) \\ &\quad - \kappa(|\mathsf{A}|)\mathsf{SNR}_\beta \frac{\frac{E_b}{N_0}\mathsf{C}^\star}{\frac{E_b}{N_0}\mathsf{C}^\star + 1}\log_2 e. &(19)\\ &\to \log_2\left(1+\frac{E_b}{N_0}\mathsf{C}^\star\right) &(20)\\ &= \mathsf{C}^\star. &(21)\end{aligned}$$

Furthermore, the $\frac{E_b}{N_0}$ required by SNR_β is upper-bounded by

$$\begin{aligned} \frac{\beta\mathsf{SNR}_\beta}{C_Q(\beta,\mathsf{SNR}_\beta)} &\leq \frac{\frac{E_b}{N_0}\mathsf{C}^\star}{\log_2\left(1+\frac{E_b}{N_0}\mathsf{C}^\star\right) - \epsilon\mathsf{C}^\star} &(22)\\ &= \frac{E_b}{N_0}\frac{1}{1-\epsilon} &(23)\end{aligned}$$

for an arbitrarily small ϵ, provided β is large enough.

To show (17) we need the following two inequalities:

$$C_{\text{qpsk}}(x) \geq \left(x - x^2\right)\log_2 e \tag{24}$$

and

$$\eta \geq \frac{1}{1+\beta\mathsf{SNR}} \tag{25}$$

where η is the solution to

$$\eta + \beta E\left[\frac{\mathsf{SNR}|\mathsf{A}|^2\eta}{1+\mathsf{SNR}|\mathsf{A}|^2\eta}\right] = 1. \tag{26}$$

To show (25) rewrite the Tse–Hanly equation as

$$\begin{aligned} \mathsf{SNR} &= \eta\mathsf{SNR} + \beta\mathsf{SNR}E\left[\frac{\mathsf{SNR}|\mathsf{A}|^2\eta}{1+\mathsf{SNR}|\mathsf{A}|^2\eta}\right] &(27)\\ &\leq \eta\mathsf{SNR} + \beta\mathsf{SNR}^2\eta &(28)\end{aligned}$$

where the inequality comes from $E[|\mathsf{A}|^2] = 1$. Inequality (24) follows from the definition of C_{qpsk}.

Now, (17) readily follows from

$$\begin{aligned} C_Q(\beta, \mathsf{SNR}) &= E[C_{\text{qpsk}}(|\mathsf{A}|^2\mathsf{SNR}\eta(z,\mathsf{SNR}))]dz &(29)\\ &\geq \int_0^\beta E\left[C_{\text{qpsk}}\left(\frac{|\mathsf{A}|^2\,\mathsf{SNR}}{1+z\mathsf{SNR}}\right)\right]dz &(30)\\ &\geq \int_0^\beta E\left[\frac{|\mathsf{A}|^2\,\mathsf{SNR}}{1+z\mathsf{SNR}}\right]\log_2 e \\ &\quad - E\left[\frac{|\mathsf{A}|^4\,\mathsf{SNR}^2}{(1+z\mathsf{SNR})^2}\right]\log_2 e\,dz &(31)\\ &\geq \int_0^\beta \frac{\mathsf{SNR}}{1+z\mathsf{SNR}}\log_2 e - \frac{\kappa(|\mathsf{A}|)\mathsf{SNR}^2}{(1+z\mathsf{SNR})^2}\log_2 e\,dz &(32)\\ &= \log_2(1+\beta\mathsf{SNR}) - \frac{\kappa(|\mathsf{A}|)\beta\mathsf{SNR}^2}{1+\beta\mathsf{SNR}}\log_2 e &(33)\end{aligned}$$

thus concluding the proof. □

Interestingly, the optimality of QPSK in the large β limit proved in Theorem 1 is different in nature from its wide-band optimality proved in [12]. In fact, as a consequence of the rotational invariance of the spreading sequences, Theorem 1 also holds if the modulation is binary phase-shift keying (BPSK). A related result on the optimality of binary inputs in the absence of spreading admits a very different (central-limit theorem based) proof [3].

In the following, we consider two alternative *pragmatic* CDMA optimization problems: 1) equal-rate, nonuniform power, and 2) equal-power, nonuniform-rate systems. We assume that, in both cases, users in each class i are decoded in parallel by a bank of single-user decoders, while classes are stripped off from J to 1, i.e., in decreasing SNR order (for the equal-rate case) or in increasing rate order (for the equal-power case). Notice that our approach is pragmatic in two ways: it makes use of QPSK rather than Gaussian codes and it performs class-by-class stripping, rather than user-by-user, as implied by (8) and (10).

B. Optimization for Equal-Rate Systems

We assume that users in all classes make use of codes drawn randomly and independently from the same ensemble with rate–threshold pair (R, g). The signal to interference plus noise ratio (SINR) at the output of the MMSE filter for class i users, assuming that all users in classes $i+1, \ldots, J$ have been perfectly canceled, is given by $\gamma_i\eta_i(\beta_i)$. Hence, the condition for successive decodability of all users is

$$\eta_i(\beta_i) \geq \frac{g}{\gamma_i}, \qquad \text{for all } i = 1, \ldots, J.$$

We fix the received power levels $\boldsymbol{\gamma}$, and consider the optimization of the class loads $\boldsymbol{\beta}$. Without loss of generality, we assume $\gamma_1 \geq g$, since for all j such that $\gamma_j < g$, we would have trivially $\beta_j = 0$. This problem can be formulated as a linear program as follows. Because of the monotonicity in the first argument of the function in (5), and the fact that the solution to $\eta_j(z) = f_j(\eta_j(z), z)$ is unique [15], we can conclude that

$$\forall\, x \in [0,\infty) \quad x \leq \eta_j(z) \Leftrightarrow x \leq f_j(x,z) \tag{34}$$

Accordingly, the successive decodability condition is equivalent to

$$\left(1 + \sum_{j=1}^{i}\beta_j\frac{\gamma_j}{1+\gamma_j\frac{g}{\gamma_i}}\right)^{-1} \geq \frac{g}{\gamma_i}, \qquad \forall\, i = 1, \ldots, J \tag{35}$$

which can be written in compact form as

$$\boldsymbol{A}\boldsymbol{\beta} \leq \boldsymbol{b}$$

where $\boldsymbol{A}$ is a $J \times J$ lower triangular matrix with nonzero elements

$$a_{i,j} = \frac{(1+g)\gamma_j}{\gamma_i + \gamma_j g} \in (0,1] \tag{36}$$

and $\boldsymbol{b}$ is a positive vector with elements

$$b_i = \frac{(1+g)(\gamma_i - g)}{\gamma_i g}. \tag{37}$$

Notice that $a_{i,i} = 1$, $a_{i,j}$ (for $1 \leq j \leq i$) is increasing with j and decreasing with i and b_i is increasing with i.

For a desired spectral efficiency $\rho = \beta R$, the optimal vector $\boldsymbol{\beta}$ which achieves (if possible) arbitrarily small BER with minimal $(E_b/N_0)_{\rm sys}$ is the solution to the linear program

$$\begin{cases} \text{minimize} & \sum_{i=1}^{J} \beta_i \gamma_i \\ \text{subject to} & \boldsymbol{A}\boldsymbol{\beta} \leq \boldsymbol{b} \\ & \sum_{i=1}^{J} \beta_i \geq \beta \\ & \boldsymbol{\beta} \geq \boldsymbol{0} \end{cases} \tag{38}$$

given by the following result.

Proposition 1: The equation $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ has a unique solution with nonnegative components $\boldsymbol{\tau}$. Furthermore, the feasible set in (38) is nonempty if and only if $\beta \leq \sum_{j=1}^{J} \tau_j$. The solution of (38) is given explicitly by

$$\beta_i^\star = \begin{cases} \tau_i, & i = 1, \dots, \hat{J} - 1 \\ \beta - \sum_{j=1}^{\hat{J}-1} \tau_j, & I = \hat{J} \\ 0, & i = \hat{J} + 1, \dots, J \end{cases} \tag{39}$$

where $\hat{J}$ denotes the minimum i for which $\beta \leq \sum_{j=1}^{i} \tau_j$.

Proof: See Appendix A □

C. Optimization for Equal-Power Systems

In this case, we assume that the users in all classes have fixed SNR γ, but users in each class j make use of a different code ensemble, characterized by the rate–threshold pair (R_j, g_j). Without loss of generality, we assume $R_1 \geq \cdots \geq R_J$ and $g_1 \geq \cdots \geq g_J$.[2] For example, assuming optimal binary codes corresponds to obtaining the pairs (R_j, g_j) by sampling the curve of Fig. 1 in given J desired rate values. Without loss of generality, we assume $\gamma \geq g_1$, since for all j such that $\gamma < g_j$, we would have trivially $\beta_j = 0$.

The successive decodability condition is given by

$$\eta_i(\beta_i) \geq \frac{g_i}{\gamma}, \qquad \text{for all } i = 1, \dots, J$$

which translates into

$$\sum_{j=1}^{i} \beta_j \leq b_i, \qquad i = 1, \dots, J \tag{40}$$

with

$$b_i = \frac{(1 + g_i)(\gamma - g_i)}{\gamma g_i} \tag{41}$$

using again property (34). Hence, for given rate–threshold pairs (R_j, g_j), the spectral efficiency $\rho = \sum_{i=1}^{J} \beta_i R_i$ maximized over the class loads is obtained as the solution to the following linear program:

$$\begin{cases} \text{maximize} & \sum_{i=1}^{J} \beta_i R_i \\ \text{subject to} & \boldsymbol{L}\boldsymbol{\beta} \leq \boldsymbol{b} \\ & \sum_{i=1}^{J} \beta_i \leq \beta \\ & \boldsymbol{\beta} \geq \boldsymbol{0} \end{cases} \tag{42}$$

where $\boldsymbol{L}$ is a lower triangular $J \times J$ matrix with $l_{ij} = 1$ for all $i \geq j$ and where $\boldsymbol{b} = (b_1, \dots, b_J)^T$ with b_i given in (41). We have the following result.

[2]Any good family of codes satisfies the condition that codes with larger rate have larger SNR thresholds.

Proposition 2: The problem (42) is always feasible and its solution is given explicitly by

$$\beta_i^\star = \begin{cases} b_i - b_{i-1}, & i = 1, \dots, \hat{J} - 1 \\ \beta - b_{\hat{J}-1}, & i = \hat{J} \\ 0, & i = \hat{J} + 1, \dots, J \end{cases} \tag{43}$$

where $b_0 \triangleq 0$, and $\hat{J}$ denotes the minimum i for which $\beta \leq b_i$.

Proof: See Appendix A. □

V. Numerical Examples

In this section, we give examples of the equal-rate and the equal-power system designs using the off-the-shelf LDPC code ensembles optimized for the binary-input AWGN channel, found in [13].

Equal-rate design. In Fig. 3, the curves labeled by "LDPC, $R = 0.2, 1.0, 1.8$" are the spectral efficiencies achieved by the equal-rate design with the LDPC codes of rate 0.2, 1.0, and 1.8 bits per QPSK symbol (corresponding to binary rate $0.1, 0.5,$ and 0.9), in the family whose rate–threshold pairs are represented by the marks in Fig. 1. The equal-rate spectral efficiency curves were obtained considering increasing values of β, and, for each β, a vector $\boldsymbol{\gamma}$ obtained by discretizing the interval $[g, \bar{\gamma}(\beta)]$ with step of 0.1 dB, where $\bar{\gamma}(\beta)$ is the minimum γ_J for which the feasible set of (38) is nonempty. The single-user capacity $\mathsf{C}^\star$ is shown for comparison. We notice that the equal-rate design is able to approach quite closely $\mathsf{C}^\star$ for low R at the price of a very large load β. For example, for $R = 0.2$ at $\rho = 2$ bits/s/Hz, the gap from $\mathsf{C}^\star$ is ≈ 0.2 dB, with load $\beta = 10$ users/chip. Figs. 4 and 5 show the distributions (mass functions) of β_j versus γ_j corresponding to $\rho = 2$ and LDPC codes of rate $R = 0.2$ and $R = 1.0$ (achieving 1.4-dB gap from $\mathsf{C}^\star$ at $\beta = 2$ users/chip).

Equal-power design. In Fig. 6, the curves labeled by "LDPC" and "discr.QPSK" are the spectral efficiencies achieved by the equal-power design with the LDPC code family found in with rate–threshold pairs corresponding to the marks in Fig. 1, and rate–threshold pairs obtained by sampling the QPSK capacity curve from $R = 0.05$ to 1.95 with step 0.1, respectively. The fact that the "LDPC" curve does not approach the "discr.QPSK" curve at high $\frac{E_b}{N_0}$ is due to the fact that the largest rate available in the LDPC code family of [13] is limited to 1.8-bits/channel use. It turns out that in order to approach $\mathsf{C}^\star$, it is necessary to have many different classes. Even a relatively finely discretized distribution of optimal (i.e., single-user capacity-achieving) rates, such as curve "discr.QPSK, "suffers some loss away from $\mathsf{C}^\star$. In Fig. 6, the low $(E_b/N_0)_{\rm Sys}$ behavior of spectral efficiency is dominated by the class with lowest coding rate (and SNR threshold). In fact, the value at which spectral efficiency becomes zero is given by g_J/r_J, which is the minimum E_b/N_0 to have a vanishing fraction of users at nonzero rate. We conclude that in order to approach optimal spectral efficiency with the equal power scheme, a wide range of coding rates is needed and, in particular, very-high- and very-low-rate codes must be designed. This might pose some practical problems for code design.

It is worthwhile to mention that the equal-power spectral efficiency curves are obtained as the upper envelope of the solution of (42), over all $\gamma \geq g_J$ and $\beta \in [0, b_J]$, i.e., for all pairs (γ, β) for which the solution (43) exists.

Effect of finite n and K. We discuss the effect of finite block length n and finite K and N on the system achievable performance throughan example. For finite code length, the decoding SNR threshold g, above which the post-decoding BER is vanishing, is not defined (strictly speaking, it is infinite). However, in practical system design, for each class j we set a target SNR threshold g_j such that if

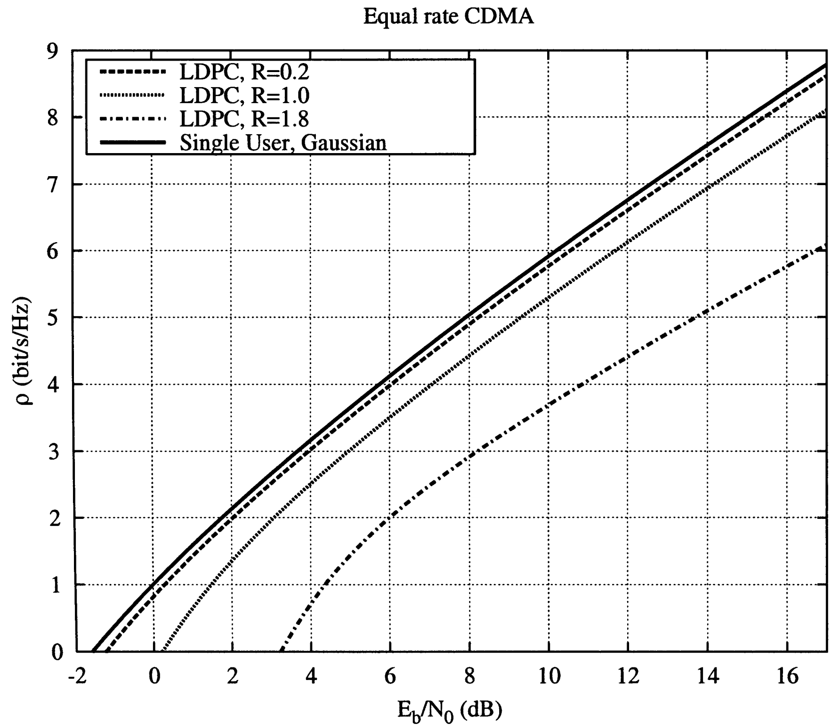

Fig. 3. Spectral efficiency of some LDPC codes with equal-rate design.

the SINR at the input of the decoders of users in class j is above g_j, then the post-decoding BER is so small that it has negligible impact on subsequent decoding stages of classes $j-1, j-2, \ldots, 1$. A natural question is whether such small but nonvanishing BER has a catastrophic effect, preventing the successive stripping decoder from decoding some class of users. We argue that for n sufficiently larger than K, and post-decoding BER sufficiently small, the effect of residual errors is indeed negligible. In fact, assuming random errors (if errors are correlated after decoding, we can use independent interleaving for each user), the expected number of incorrectly decoded symbols interfering with a user of class j is given by $\epsilon \sum_{i=j+1}^{j} K_i$, where ϵ is the residual BER. Typically, LDPC codes have the property that, at the transition between the *waterfall* and the *error flattening* regions of the BER curve, the BER is $\epsilon = O(1/n) = \kappa/n$ for some constant $\kappa \ll n$. Turbo codes have similar behavior, where $\epsilon = O(1/I)$ and I is the size of the interleaver of the turbo encoder (this effect is called "interleaving gain" in [22]). Hence, by letting the thresholds g_j correspond to the SNR at the transmission between waterfall and flattening, and by letting $n \gg K$, we have that the expected number of incorrectly decoded symbols interfering with a user of class j is given by $\kappa(\sum_{i=j+1}^{j} K_i)/n \ll 1$.

Eventually, we conclude that a *sensible* approach for the design of practical CDMA systems based on successive decoding is to use the large-system optimization methods developed before, while replacing the infinite block-length thresholds g_j by some target SNR values chosen according to the BER versus SNR performance of actual finite-length codes. While this argument provides only a heuristic design approach, extensive simulations show that the resulting systems are very good.

Fig. 8 shows the spectral efficiency achievable by the equal-rate design with $R = 1.0$. The curve denoted by "optimal" corresponds to the threshold $g = 0.186$ dB, of ideal infinite-length QPSK random coding. The curve denoted by "practical" corresponds to the threshold $g =$ 0.933 dB, of the finite-length irregular LDPC code ensemble generated from the degree distribution found in [13].[3] From Fig. 7 we observe that this threshold corresponds to BER $\approx 5 \cdot 10^{-3}$ for $n = 5000$ and BER $\approx 5 \cdot 10^{-4}$ for $n = 10\,000$. We simulated a system with $\rho = 2.0$ bits/s/Hz, spreading factor $N = 64$, and $K = 128$ users, corresponding to the mark in Fig. 8. Moreover, in order to improve the robustness of the stripping decoder to residual errors, *soft stripping* can be used, and successive decoding of the users from class J to class 1 can be iterated more than once (we refer to this approach as the multipass soft-stripping decoder, where one decoding pass consists of decoding once all the users). Soft stripping (see, for example, [23]) consists of subtracting from the signal the MMSE estimates of the signals of the already-decoded users instead of their hard decisions. When the user decoders are symbol-by-symbol maximum *a posteriori* (MAP) (or approximations thereof via belief-propagation decoder, as in LDPC decoding [17], [18]), the MMSE estimate of the ith symbol of user k, $x_{i,k}$ is obtained as the conditional mean

$$\widetilde{x}_{i,k} = E[x|\mathrm{EXT}_{i,k}]$$

where $\mathrm{EXT}_{i,k}$ denotes the kth decoder *extrinsic information* for the ith symbol. For example, with QPSK symbols we have

$$\widetilde{x}_{i,k} = \frac{1}{\sqrt{2}} \tanh\left(\frac{\mathcal{L}_{i,k}^{(I)}}{2}\right) + \frac{j}{\sqrt{2}} \tanh\left(\frac{\mathcal{L}_{i,k}^{(Q)}}{2}\right) \tag{44}$$

where $\mathcal{L}_{i,k}^{(I)}$ and $\mathcal{L}_{i,k}^{(Q)}$ denote the extrinsic belief-propagation decoder messages for the variable nodes corresponding to the bits modulated in the in-phase and quadrature components of the ith QPSK symbol (see [23] for more details). Soft stripping has the advantage that if decisions are not reliable, i.e., $|\mathcal{L}_{i,k}^{(I)}|$ and/or $|\mathcal{L}_{i,k}^{(Q)}|$ are small, then the effect of

[3]The threshold SNR $g = 0.933$ dB was obtained by trial-and-error as the minimum value for which the multipass successive decoder removes all interference in two decoding passes for block length $n = 5000$. For different codes, block length, and channel load this value might be different.

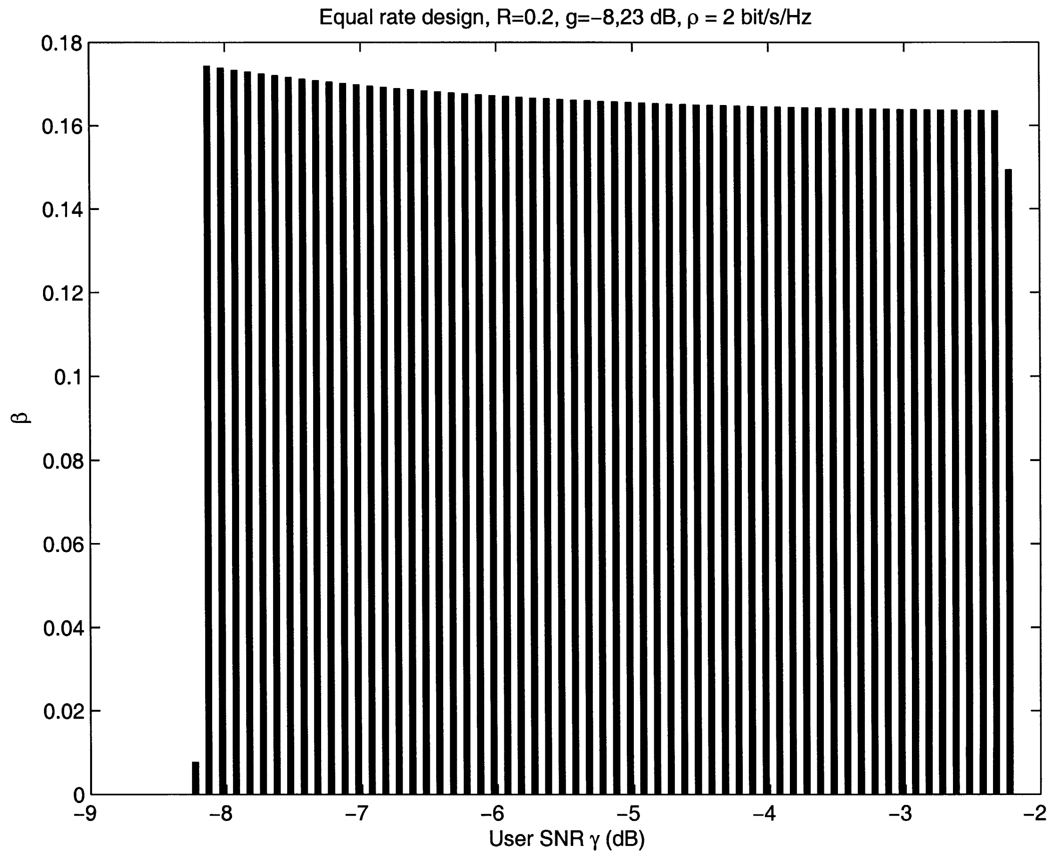

Fig. 4. Load distribution ($\{\beta_j\}$ versus $\{\gamma_j\}$) for the equal-rate design with LDPC-coded QPSK of rate 0.2-bit/channel use and $\rho = 2$ bits/s/Hz.

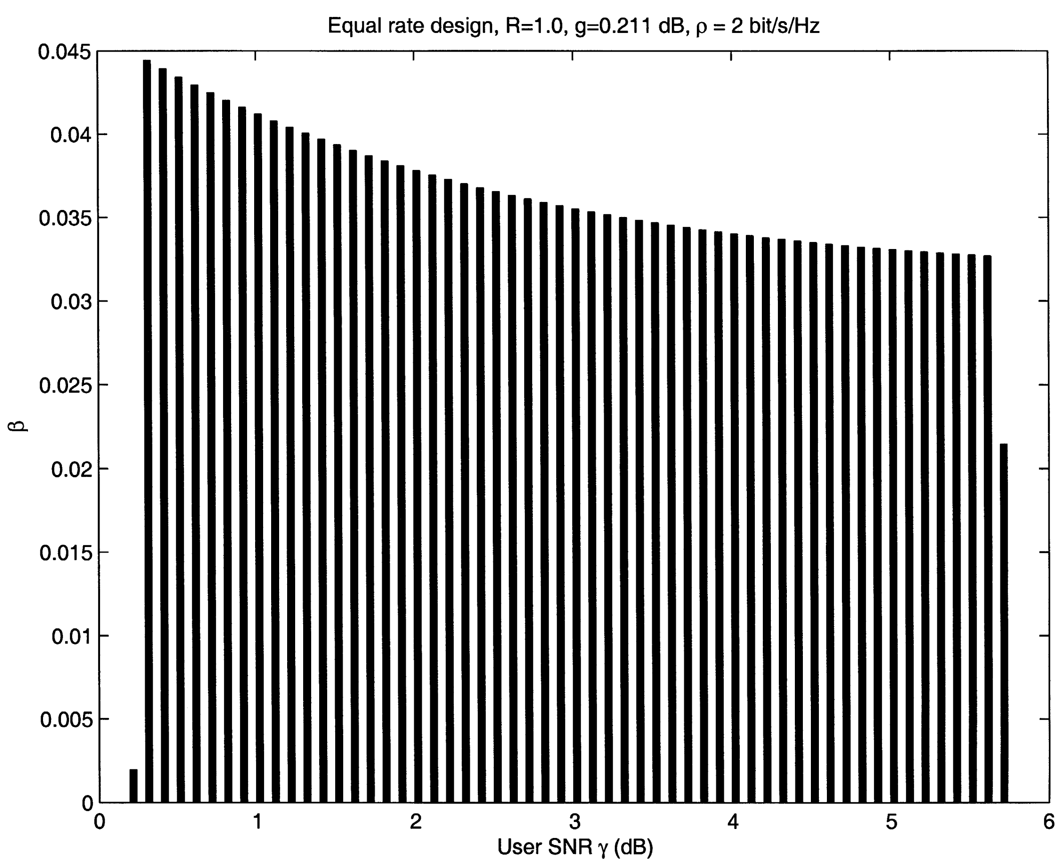

Fig. 5. Load distribution ($\{\beta_j\}$ versus $\{\gamma_j\}$) for the equal-rate design with LDPC-coded QPSK of rate 1.0-bit/channel use and $\rho = 2$ bits/s/Hz.

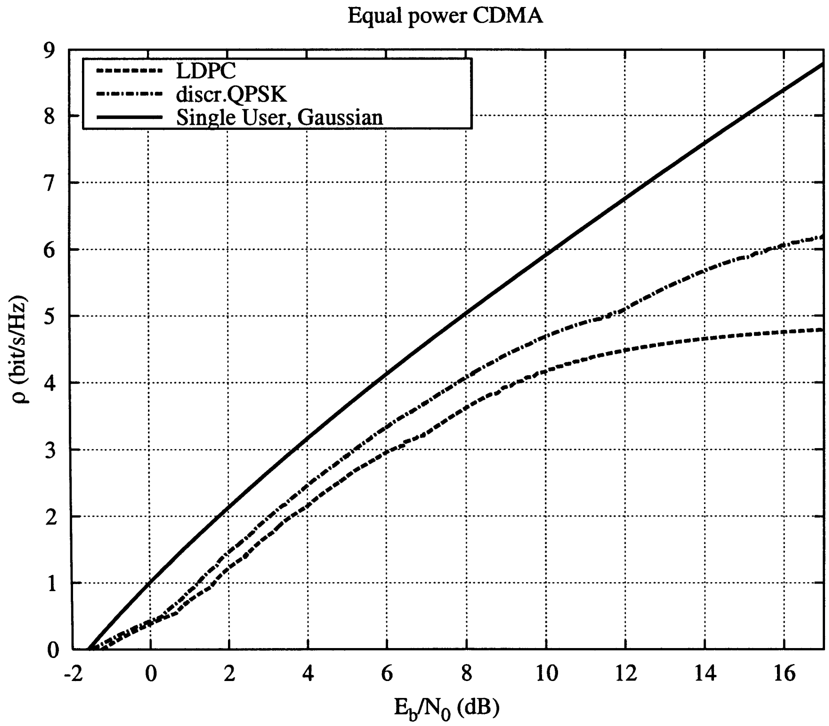

Fig. 6. Spectral efficiency of LDPC and optimal QPSK codes with equal-power design.

this symbol on the residual interference signal is attenuated.The evolution of the users' SINR versus the successive stripping decoder iterations is shown in Fig. 9 for $n = 5000$ and in Fig. 10 for $n = 10\,000$. These curves are snapshots obtained by random generation of the noise, of the spreading sequences, of the information sequences and of the LDPC code graphs. Each user is characterized by the evolution of its SINR with the successive decoding steps. For the sake of clarity, we have shown only the minimum, maximum, and average SINR over all users. The successive decoder makes use of soft stripping. Each LDPC decoder is run for a maximum of 200 iterations, and three interference cancellation passes are performed.

For $n = 5000$, we need to perform two soft-stripping successive decoding passes before all users reach their target SNR threshold. For $n = 10\,000$, since the threshold is much more conservative, a single pass is sufficient. Equivalently, one could have lowered the threshold and achieved the same spectral efficiency at a smaller E_b/N_0. In any case, we notice that the gap between an optimal (infinite block length) system and a practical finite-length system depends almost entirely on the fact that finite-length codes need SNR significantly larger (0.747 dB in this example) than the infinite-length decoding threshold. On the other hand, no catastrophic error propagation effect is observed when the system is designed according to the rules outlined above.

VI. Conclusion

We considered the optimization of a canonical coded synchronous CDMA system characterized by random spreading and QPSK signaling, in the limit for large number of users, large spreading gain, and large user code block length. Such assumptions may be regarded as "pragmatic" in the sense that they are all motivated by actual CDMA systems. The CDMA system considered here has low complexity, as it assumes successive stripping with MMSE filters. Excellent approximations to the MMSE filters can be precomputed using the large random matrix design approach of [24], with complexity $O(K^2)$. Moreover, powerful long user codes such as LDPC codes can be decoded iteratively, with linear complexity in the block length. Hence, the overall complexity per decoded information bit of the multiuser decoder is linear in K and constant in the code block length, i.e., comparable with the complexity of standard CDMA systems based on single-user detection and separated single-user decoding. Nevertheless, the proposed system optimization, driven by recent information-theoretic results, yields spectral efficiencies remarkably close to the optimal (i.e., optimizing also with respect to the user signature waveforms and using Gaussian codebooks).

We quantified the loss in spectral efficiency due to the use of QPSK *in lieu* of Gaussian inputs. The loss for high SNR is not as pronounced as in the single-user case and, in fact, we showed that it vanishes for large channel load β. Then, we considered two special cases of the general rate and power allocation problem: namely, the optimization of the received SNR distribution for an equal-rate system, and the optimization of the user rate distribution for an equal-power system, subject to the successive decodability condition imposed by the stripping decoder. Both problems yield linear programs that admit closed-form explicit solutions.

From a practical viewpoint, the equal-rate system design is probably more attractive than its equal-power counterpart since it can approach optimal spectral efficiency uniformly, for all E_b/N_0's, provided that the individual users coding rate is small. Moreover, controlling the received user SNR is much easier and closer to existing power-control schemes than allocating coding rates (and channel codes) to the users.

Numerical results show that the system optimization carried out in the large-system limit and for infinite code block length can be used effectively to dimension practical systems, provided that the SNR thresholds are chosen according to the actual BER performance of the finite-length user codes. Systems optimized according to the proposed method do not suffer from catastrophic error propagation of the successive stripping decoder even if, in general, finite-length codes have nonvanishing post-decoding BER.

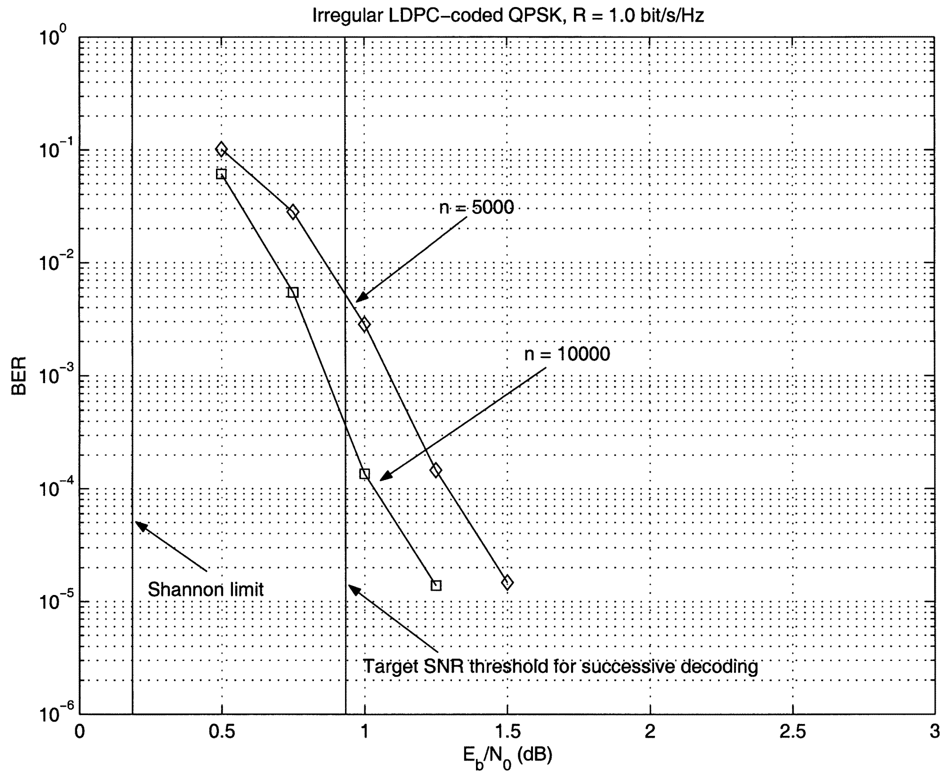

Fig. 7. Average BER performance of irregular LDPCs of rate $1/2$ over (single-user) AWGN with QPSK modulation (spectral efficiency $R = 1.0$ bit/s/Hz). The BER is obtained by averaging over randomly generated parity-check matrices with the given degree distributions (max left degree 100, average right degree 11). The belief-propagation decoder performs up to 200 iterations.

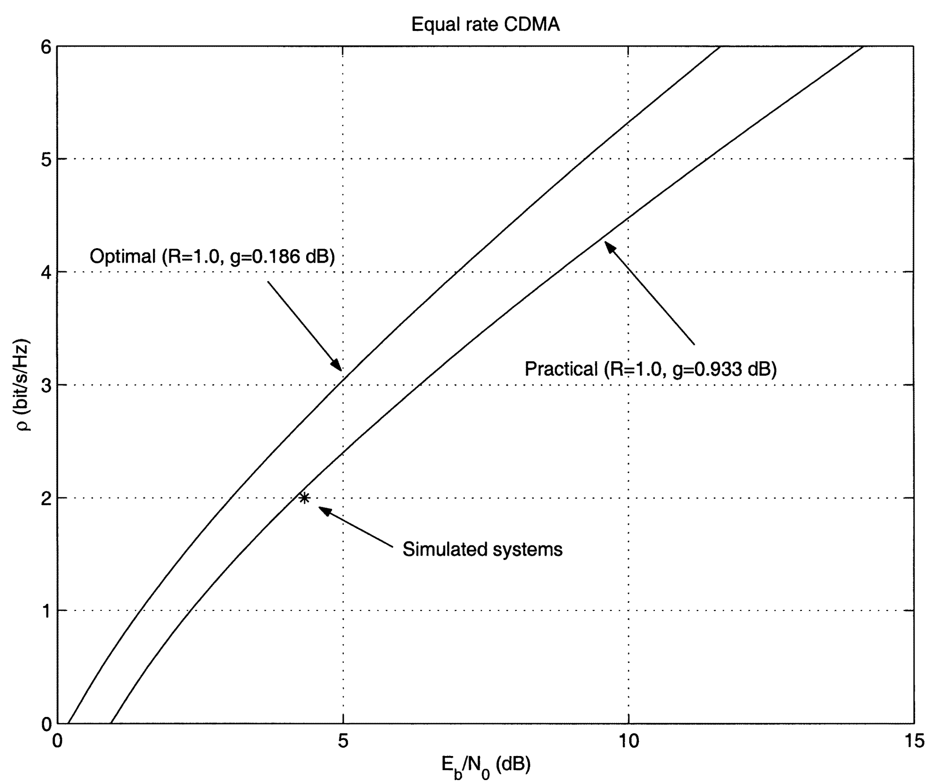

Fig. 8. Spectral efficiency achievable by an ideal random coding QPSK ensemble of rate 1.0, and by a suboptimal code ensemble with the same rate and SNR larger threshold. The mark corresponds to the simulated system.

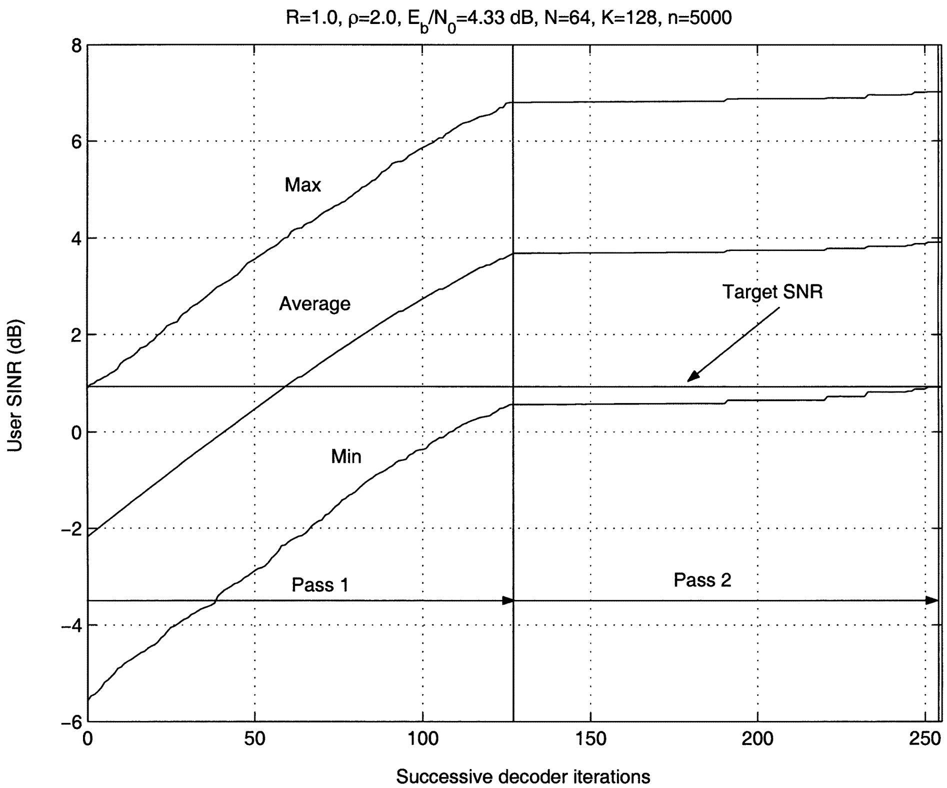

Fig. 9. Evolution of the user SINR at the LDPC decoder input versus the successive decoding steps, for the multipass soft-stripping decoder with LDPC code length $n = 5000$. Each decoder pass corresponds to 128 decoding steps (i.e., to the decoding of all $K = 128$ users).

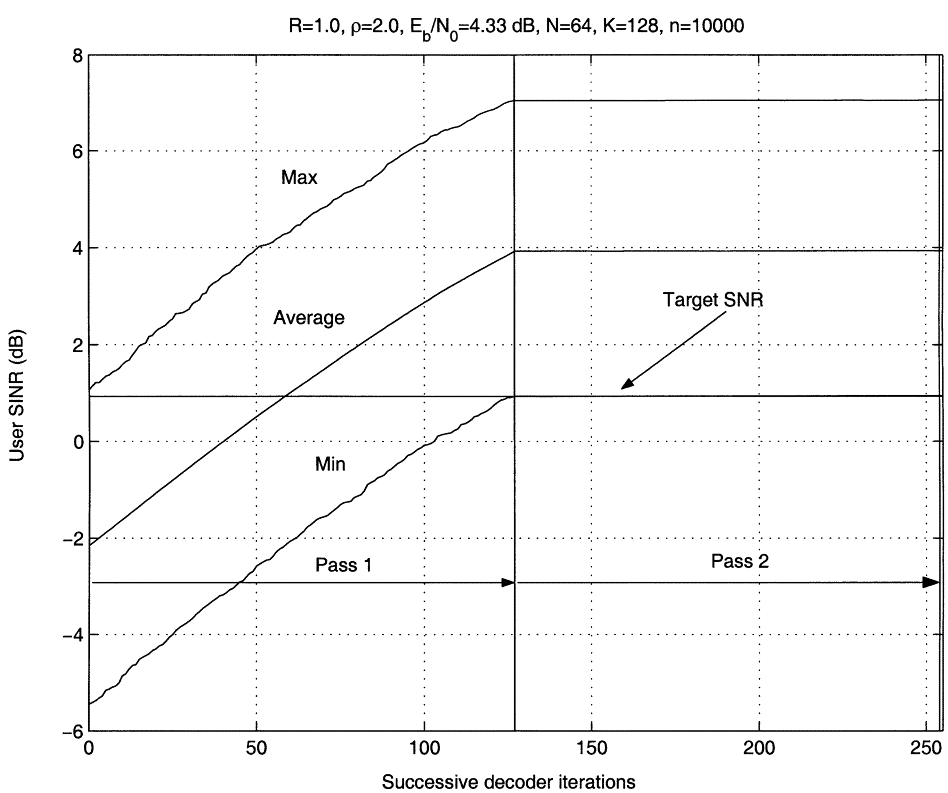

Fig. 10. Evolution of the user SINR at the LDPC decoder input versus the successive decoding steps, for the multipass soft-stripping decoder with LDPC code length $n = 10\,000$. Each decoder pass corresponds to 128 decoding steps (i.e., to the decoding of all $K = 128$ users).

APPENDIX

A. Proofs

Proof of Proposition 1: A necessary condition for $\boldsymbol{\beta}$ minimizing the objective function in (38) is that the constraint $\sum_j \beta_j \geq \beta$ holds with equality. Hence, without loss of generality we rewrite (38) in the canonical form

$$\begin{cases} \text{minimize} & \boldsymbol{\gamma}^T \boldsymbol{\beta} \\ \text{subject to} & -\boldsymbol{A}\boldsymbol{\beta} \geq -\boldsymbol{b} \\ & \mathbf{1}^T \boldsymbol{\beta} = \beta \\ & \boldsymbol{\beta} \geq \mathbf{0}. \end{cases} \tag{45}$$

The dual linear program is given by

$$\begin{cases} \text{maximize} & (-\boldsymbol{b}^T, \beta) \begin{bmatrix} \boldsymbol{y} \\ \alpha \end{bmatrix} \\ \text{subject to} & [-\boldsymbol{A}^T, \mathbf{1}] \begin{bmatrix} \boldsymbol{y} \\ \alpha \end{bmatrix} \leq \boldsymbol{\gamma} \\ & \boldsymbol{y} \geq \mathbf{0} \end{cases} \tag{46}$$

where α can be either positive or negative.

From the properties of the coefficients $a_{i,j}$ and b_j we get immediately that $\boldsymbol{A}$ is invertible and the vector $\boldsymbol{\tau}$ such that $\boldsymbol{\tau} = \boldsymbol{A}^{-1}\boldsymbol{b}$ has nonnegative components. The vector $\boldsymbol{\beta} \in \mathbb{R}_+^J$ maximizing $\mathbf{1}^T\boldsymbol{\beta}$ and satisfying $\boldsymbol{A}\boldsymbol{\beta} \leq \boldsymbol{b}$ is $\boldsymbol{\tau}$ (this is easily shown by contradiction, since $\boldsymbol{\tau}$ is the unique nonnegative vector $\boldsymbol{\beta}$ that makes the inequality $\boldsymbol{A}\boldsymbol{\beta} \leq \boldsymbol{b}$ componentwise tight). Hence, if $\mathbf{1}^T\boldsymbol{\tau} < \beta$ the primal problem is infeasible. On the other hand, if $\mathbf{1}^T\boldsymbol{\tau} \leq \beta$, the primal problem is feasible, and a feasible point is given by (39). In order to show that this is indeed the desired solution, we shall assume that $\mathbf{1}^T\boldsymbol{\tau} \geq \beta$ and find a feasible point for the dual problem. Then, we show that the value of the dual problem at this point is equal to the value of the primal problem at (39).

We rewrite the inequality constraint and the objective function in the dual problem (46) as

$$\boldsymbol{A}^T\boldsymbol{y} \geq \alpha\mathbf{1} - \boldsymbol{\gamma} \tag{47}$$

and

$$-\boldsymbol{b}^T\boldsymbol{y} + \alpha\beta. \tag{48}$$

The vector $\alpha\mathbf{1} - \boldsymbol{\gamma}$ has decreasing components. For fixed α, let K_α denote the number of positive elements of $\alpha\mathbf{1} - \boldsymbol{\gamma}$. It is clear from (47) and (48) that the objective function is maximized by letting the last $J - K_\alpha$ components of $\boldsymbol{y}$ be equal to zero. We introduce the following short-hand notation: for a vector $\boldsymbol{x} \in \mathbb{R}^J$ and a matrix $\boldsymbol{M} \in \mathbb{R}^{J\times J}$, we let $\boldsymbol{x}_\alpha$ and $\boldsymbol{M}_\alpha$ denote the $K_\alpha \times 1$ subvector of $\boldsymbol{x}$ formed by its first K_α components, and the $K_\alpha \times K_\alpha$ submatrix of $\boldsymbol{M}$ formed by its first K_α rows and columns, respectively. Then, a feasible point for the dual problem is the vector $\boldsymbol{\pi}$ such that its first K_α components are given by

$$\boldsymbol{\pi}_\alpha = \alpha(\boldsymbol{A}_\alpha^T)^{-1}\mathbf{1}_\alpha - (\boldsymbol{A}_\alpha^T)^{-1}\boldsymbol{\gamma}_\alpha \tag{49}$$

and the remaining $J - K_\alpha$ components are equal to zero.

The value of the objective function (48) at this point is given by

$$f(\alpha) = \boldsymbol{b}_\alpha^T(\boldsymbol{A}_\alpha^T)^{-1}\boldsymbol{\gamma}_\alpha + \alpha\left(\beta - \boldsymbol{b}_\alpha^T(\boldsymbol{A}_\alpha^T)^{-1}\mathbf{1}_\alpha\right). \tag{50}$$

It is not difficult to see that $f(\alpha)$ is a continuous and piecewise-linear function of α, for $\alpha \leq [\gamma_1, \gamma_J]$. The assumption $\mathbf{1}^T\boldsymbol{\tau} \geq \beta$ can be rewritten as $\beta - \mathbf{1}^T\boldsymbol{A}^{-1}\boldsymbol{b} \leq 0$. Hence, for $\alpha > \gamma_s$, for some $1 \leq s \leq J$, the slope of $f(\alpha)$ is negative, while for $\alpha < \gamma_s$ the slope is positive. Therefore, the maximum of $f(\alpha)$ with respect to α is achieved for $\alpha = \gamma_s$ and, by definition, s is the minimum index in $1, \ldots, J$ such that $\sum_{j=1}^s \tau_j \geq \beta$, i.e., $s = \hat{J}$ defined in (39). The primal objective function evaluated at the feasible point (39) is given by

$$(\gamma_1, \ldots, \gamma_s, 0, \ldots, 0)\boldsymbol{A}^{-1}\boldsymbol{b} + \gamma_s\left(\beta - (\underbrace{1, \ldots, 1}_{s}, 0, \ldots, 0)\boldsymbol{A}^{-1}\boldsymbol{b}\right).$$

It is immediate to see that this coincides with the dual objective function $f(\alpha)$ evaluated at $\alpha = \gamma_s$. Hence, we conclude that (39) is the sought solution.

Proof of Proposition 2: The proof follows immediately by observing that, for $\beta \leq b_J$, the program (42) can be reformulated as the $\hat{J}$-dimensional polymatroid program [25]

$$\begin{cases} \text{maximize} & \sum_{i=1}^{\hat{J}} \beta_i R_i \\ \text{subject to} & \sum_{i\in S} \beta_i \leq r(S), \quad \forall\, S \subseteq \{1, \ldots, \hat{J}\} \\ & \boldsymbol{\beta} \geq \mathbf{0} \end{cases} \tag{51}$$

where the rank function $r(S)$ is defined by

$$r(S) = \sum_{i=1}^{\max\{S\}} \Delta_i$$

where $\Delta_i = b_i - b_{i-1}$ for $i = 1, \ldots, \hat{J} - 1$ and $\Delta_{\hat{J}} = \beta - b_{\hat{J}-1}$. Since $(b_0, \ldots, b_{\hat{J}-1}, \beta)$ is increasing, $r(S)$ is submodular.

REFERENCES

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

[2] B. Rimoldi and R. Urbanke, “A rate splitting approach to the Gaussian multiple-access channel,” *IEEE Trans. Inform. Theory*, vol. 42, pp. 364–375, Mar. 1996.

[3] S. Verdú, “Capacity region of Gaussian CDMA channels: The symbol-synchronous case,” in *Proc. 24th Allerton Conf. Communication, Control and Computing*, Monticello, IL, Oct. 1986, pp. 1025–1034.

[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 Groove, CA, Nov. 1997.

[5] P. Viswanath and V. Anantharam, “Optimal sequences and sum capacity of synchronous CDMA systems,” *IEEE Trans. Inform. Theory*, vol. 45, pp. 1984–1991, Sept. 1999.

[6] W. Yu, W. Rhee, S. Boyd, and J. Cioffi, “Iterative water-filling for Gaussian vector multiple-access channels,” *IEEE Trans. Inform. Theory*, vol. 50, pp. 145–152, Jan. 2004.

[7] A. J. Viterbi, *CDMA—Principles of Spread Spectrum Communications*. Reading, MA: Addison-Wesley, 1995.

[8] “TS 25.224 V3.1.0,3GPP-TSG-RAN-WG1; Physical Layer Procedures (FDD),” ETSI, 3GGP, 1999.

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

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

[11] T. Tanaka, “A statistical mechanics approach to large-system analysis of CDMA multiuser detectors,” *IEEE Trans. Inform. Theory*, vol. 48, pp. 2888–2910, Nov. 2002.

[12] S. Verdú, “Spectral efficiency in the wideband regime,” *IEEE Trans. Inform. Theory*, vol. 48, pp. 1319–1343, June 2002.

[13] R. Urbanke *et al.*. (2002). [Online]. Available: http://lthcwww.epfl.ch/research/ldpcopt/

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

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

[16] B. J. Frey, R. Koetter, G. D. Forney Jr., F. R. Kschischang, R. J. McEliece, and D. A. Spielman, Eds., “Special Issue on Codes on Graphs and Iterative Decoding,” *IEEE Trans. Inform. Theory*, vol. 47, Feb. 2001.

[17] 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.

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

[19] 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. 657–670, Feb. 2001.
[20] S.-Y. Chung, "On the construction of some capacity-approaching coding schemes," Ph.D. dissertation, MIT, Cambridge, MA, 2000.
[21] E. Chong, J. Zhang, and D.N. C. Tse, "Output MAI distribution of linear MMSE multiuser receivers in DS-CDMA systems," *IEEE Trans. Inform. Theory*, vol. 47, pp. 1128–1144, Mar. 2001.
[22] S. Benedetto and G. Montorsi, "Unveiling turbo codes: Some results on parallel concatenated coding schemes," *IEEE Trans. Inform. Theory*, vol. 42, pp. 409–428, Mar. 1996.
[23] J. Boutros and G. Caire, "Iterative multiuser decoding: Unified framework and asymptotic performance analysis," *IEEE Trans. Inform. Theory*, vol. 48, pp. 1772–1793, July 2002.
[24] L. Li, A. Tulino, and S. Verdú, "Asymptotic eigenvalue moments for linear multiuser detection," *Commun. Inform. Syst.*, vol. 1, pp. 273–304, Fall 2001.
[25] J. Edmonds, "Submodular functions, matroids and certain polyhedra," in *Proc. Calgary Int. Conf. Combinatorial Structures and Applications*, Calgary, AB, Canada, June 1969, pp. 69–87.

Simple Polynomial Detectors for CDMA Downlink Transmissions on Frequency-Selective Channels

Walid Hachem, *Member, IEEE*

***Abstract*—In code-division multiple-access (CDMA) transmissions, computing the multiuser minimum mean-squred error (MMSE) detector coefficients requires the inversion of the covariance matrix associated to the received vector signal, an operation usually difficult to implement when the spreading factor and the number of users are large. It is therefore interesting to approximate the inverse by a matrix polynomial. In this correspondence, means for computing the polynomial coefficients are proposed in the context of CDMA downlink transmissions on frequency-selective channels, the users having possibly different powers. Derivations are made in the asymptotic regime where the spreading factor and the number of users grow toward infinity at the same rate. Results pertaining to the mathematics of large random matrices, and in particular to free probability theory, are used. Spreading matrices are modeled as isometric random matrices (spreading vectors orthonormality is a natural assumption in downlink) and also as random matrices with independent and identically distributed (i.i.d.) elements.**

***Index Terms*—Code-division multiple access (CDMA), free probability, large systems analysis, polynomial detectors, reduced rank Wiener filters.**

I. Introduction

In direct-sequence code-division multiple access (DS-CDMA), spectral efficiency is considerably increased when a multiuser detector such as the linear minimum mean-squared error (MMSE) detector is used in place of the conventional single-user matched filter (SUMF). However, computation of this detector has a high complexity when the spreading factor and the number of users are large: the main difficulty comes from the inversion of the covariance matrix associated to the received vector signal, necessary for this computation. When the spreading sequences are time variant, this operation has to be done at every symbol interval. In [1], the inverse of the covariance matrix is approximated by a matrix polynomial with a degree depending on the allowable complexity. The natural problem here is to compute simply the polynomial coefficients, considered now as a constrained, thus suboptimal, MMSE detector. Note that this polynomial detector can be seen as a reduced-rank MMSE detector, which is a filter operating in a subspace having a dimension smaller than the dimension of the signal subspace. Specifically, it is implicitly shown in [2, Theorem 2] that the polynomial detector considered in this correspondence is equivalent to the so-called reduced rank multistage Wiener filter (MSWF).

Recently, the mathematics that describe the asymptotic behavior of large random matrices have gained a growing interest in the communications research community (see, for instance, [3] and the references therein). In the context of polynomial detectors, [4] and [5] proposed simple algorithms for computing polynomial coefficients when the spreading factor and the number of interferers are large. They relied on the observation that when spreading matrices are modeled as random matrices with independent and identically distributed (i.i.d.) entries, then the covariance matrix has a limit eigenvalue distribution when the spreading factor grows toward infinity and the system load, which is the ratio of the number of data streams to the spreading factor, converges toward a constant. The optimum asymptotic detector coefficients in the sense of the mean-squared error (MSE) criterion are deduced from the moments of this limit distribution, and these depend on the statistical properties of the signal model in a simple fashion. In the asymptotic regime, the optimum coefficients are computed independently of the spreading matrix realization. [6] used the same type of tools and exploited the additional information associated to the noncircular nature of data symbols when these have their values in a binary phase-shift keying (BPSK) constellation. The authors in [7] and [8] considered iterative techniques for inverting the covariance matrix. These techniques show close relationships with the polynomial approaches.

All these contributions considered frequency flat-fading channels. However, in high-rate transmissions, often the signal bandwidth is not negligible compared to the channel coherence bandwidth. In this correspondence, the polynomial detector for downlink CDMA transmissions is studied for frequency-selective channels and unequal powers allocated to data streams. Orthonormal and nonorthonormal spreading vectors are considered. In the first case, the spreading matrix is modeled as a random isometry matrix while in the second case, it is modeled as a matrix having i.i.d. entries. Performance of the polynomial detectors in the asymptotic conditions is assessed in terms of the signal-to-interference-and-noise ratio (SINR) at the detector output. As a particular case, a closed-form expression of the asymptotic performance of the SUMF, which is a zero-degree polynomial detector, is also given. The impact of channel frequency selectivity on the performance of this detector is thus quantified.

Note that when the spreading matrix has i.i.d. entries, it is established that the noise at the output of various kinds of linear detectors can be considered as Gaussian when K and N grow large ([9], [10]). Therefore, relying on the SINR for computing the bit-error rate (BER) in the asymptotic regime is rigorously justified in this case.

The mathematical tool for studying the asymptotic behavior of large random matrices used in this correspondence is free probability theory ([11]). This theory has been used in a certain number of contributions (see, for instance, [3], [12]) to study the performance of detectors used for multiple-access communications.

This correspondence is organized as follows. The problem is formulated in Section II. A procedure for computing the polynomial detector coefficients in the case where the spreading matrix is isometric and in the case where it has i.i.d. elements is presented in Section III.

Manuscript received July 17, 2002; revised May 13, 2003. The material in this correspondence was presented in part at the 36th Asilomar Conference on Signals, Systems and Computing, Pacific Grove, CA, 2002.
The author is with the Ecole Supérieure d'Electricité (Supélec), Service Radioélectricité, Plateau de Moulon, 91192 Gif-Sur-Yvette Cedex, France (e-mail: Walid.Hachem@supelec.fr).
Communicated by D. N. C. Tse, Associate Editor for Communications.
Digital Object Identifier 10.1109/TIT.2003.821985