Optimal Power/Rate Allocation and Code Selection for Iterative Joint Detection of Coded Random CDMA

Christian Schlegel, Zhenning Shi, and Marat Burnashev

Abstract—Iterative interference cancellation of coded CDMA using random spreading with linear cancellation is analyzed. If users are grouped into power classes and Shannon-bound approaching codes are used, a geometric power distribution achieves the AWGN channel Shannon bound as the numbers of classes becomes large. The optimal distribution of the size of these classes is shown to be uniform. If users are grouped into different rate classes with equal powers among equal rate users, the Shannon bound for AWGN channels can be achieved with an arbitrary distribution of the classes sizes, provided that the size of the largest rate class obeys the mild condition that its ratio of size to processing gain is much smaller than the inverse of the signal-to-noise ratio.

The case of equal powers and equal rates among all users is addressed as a “worst case” scenario. It is argued that simple repetition codes provide for a larger achievable capacity than stronger codes. It is shown that this capacity monotonically increases as the rate of the code decreases. A density evolution analysis is used to show that the achievable rates exceed those of a minimum-mean square error filter applied to the uncoded signals. This lower bound is tight for small ratios of bit energy to noise power, and otherwise the iterative cancellation receiver provides an appreciably larger capacity. Relating to recent result from the application of statistical mechanics it is shown that the repetition-coded system with iterative cancellation achieves the performance of an equivalent optimal joint detector for uncoded transmission.

Index Terms—Joint detection, iterative decoding, random CDMA, optimal power, optimal rate

I. INTRODUCTION

Iterative joint decoding of code-division multiple access (CDMA) systems using forward error control coding is based on the success of turbo coding [3], and has the potential of realizing a significant portion of the multiple access channel capacity [1], [21], [16], [2], [24], [37]. Iterative decoding breaks the complex task of a multiuser decoder into two operations, viz. (i) a posteriori probability (APP) estimation of the coded symbols, or an approximation thereof, and (ii), parallel soft decoding of single-user forward error control (FEC) codes.

The asymptotic analysis of large-scale CDMA systems has been studied for optimal and some suboptimal multiuser detectors [33], [32], and [28], where it is shown that the output SINRs, approach certain constant values if random spreading is used. The convergence behavior of an iterative decoder can be studied using density evolution (DE) analysis [22], [4]. In [7] and [6] different power levels and respectively a linear programming solution to optimizing these power levels are presented.

In [25] the impact of error control coding on iterative systems using cancellation-type joint detection front-ends is analyzed. It was found that for equal power levels among the users, weak error control codes, such as simple parity check codes or low-memory convolutional codes support substantially higher system loads than stronger turbo or low-density parity-check codes. A complex a posteriori probability (APP) decoder for the multiple-access channel, which is operated jointly with an inner error control decoder, has been used to achieve near capacity performance on small MIMO systems in [30], [29], and on MIMO multiple-access systems in [27] by optimizing the degree distributions of the irregular LDPC codes via linear programming.

In this paper the iterative decoder is viewed as a separation or layering device. Consequently we are interested in the capacities of the layers. We study the role of the error control codes via density evolution, discuss optimal power and system load distributions, as well as equal-rate, equal-power systems, which represent the most difficult situation for such a joint decoder. Even though the iterative cancellation receiver using repetition codes under equal received power levels has the same error performance as an optimal CDMA receiver, analyzed via a statistical mechanics approach in [28], it falls significantly short of what is achievable with optimal power and/or rate assignments. Furthermore, in the low signal-to-noise ratio regime, the decoder’s performance is identical to that of a minimum-mean square error (MMSE) matrix filter receiver [38].

II. SYSTEM MODEL

In an asynchronous CDMA system $K$ transmitters generate independent binary symbols $u_k \in \{-1,1\}$, $k = 1, \ldots, K$ which are encoded by $K$ parallel forward error (FEC) control encoders. Random interleavers separate the error control encoders from the spreading operation. BPSK is chosen as the modulation format.


Christian Schlegel is with the Department of Electrical and Computer Engineering, University of Alberta, Edmonton, Canada.

Zhenning Shi is with National ICT Australia, and affiliated with the Australian National University, Canberra, ACT, Australia.

Marat Burnashev is with Institute for Problems of Information Transmission, Moscow, Russia.

Supported in part by NSF under Grant CCR 9732962 and iCORE Alberta, and by Grant 03-01-00098 from the Russian Fund for Fundamental Research.
The signal from the spreader of transmitter $k$ is given by

$$x_k(t) = \sum_{l=0}^{L-1} \sqrt{P_k}d_{k,l}a_{k,l}(t - lT - \tau_k)$$

where $L$ is the number of code symbols per user per frame, $d_{k,l}$ is the $l$-th symbol in the data stream of user $k$, $\tau_k < T$ is the time delay, and $P_k$ is the power of user $k$. The spreading waveform of user $k$ during symbol time $l$, $a_{k,l}(t)$, is supported on the interval $[0, T]$, and given by

$$a_{k,l}(t) = \sum_{n=0}^{N-1} a_{k,l,n}p_w(t - nT_c)$$

where $N$ is the spreading gain, $T_c$ is the chip interval, $a_{k,l,n} \in \{-1/\sqrt{N}, 1/\sqrt{N}\}$ is the $n$-th spreading chip for user $k$ at symbol time $l$, and $p_w(t)$ is the chip waveform with normalized energy $E_c = 1/N$, such that $|a_{k,l}(t)|^2 = 1$. The chips of each symbol are chosen independently (see [36]).

We consider an AWGN multiple access channel and the received signal is

$$y(t) = \sum_{k=1}^{K} x_k(t) + n(t)$$

where $n(t)$ is zero mean white Gaussian noise with double-sided power spectral density $\sigma^2 = N_0/2$. Assuming that timing and phase references have been established, the received signal sampled by chip matched filters can be written in the convenient discrete chip-based matrix model [12]

$$y = Ad + n$$

where $n$ is an $(L + 1)N$ vector of sampled white noise with variance $\sigma^2$. $A$ is an $(L+1)N \times LK$ matrix whose $j$-th column is $a_{k,l} = [0_{1N + \tau_k/T}. a_{k,l,0}, \ldots, a_{k,l,N-1}, 0_{1N-\tau_k/T}]^T$; where $l = j/K$ is the symbol time index, $0_i$ is a length-$i$ all-zero column vector, and $d = [d_{1,0}, d_{K,0}, d_{1,1}, \ldots, d_{K,L-1}]^T$ is the vector of encoded symbols.

### III. Iterative (Turbo) Joint Decoding

Figure 1 shows the block diagram of an iterative joint receiver. The CDMA interference front-end generates soft outputs of the encoded symbols $d_{k,l}$ given a received frame $y$, and separates these into $K$ parallel streams suitable for the $K$ outer FEC decoders. Ideally, we might want to use a CDMA channel APP decoder which computes log-likelihood ratios (LLR) of $d_{k,l}$, given by $\lambda(d_{k,l}) = \log(P(d_{k,l} = 1|y)/P(d_{k,l} = -1|y))$, but complexity typically prohibits this.

Using soft cancellation at the front-end [1], [4], user $k$ sees a cancelled signal given by

$$y_k = \sum_{l=1}^{L} \sqrt{P_k}d_{k,l}a_{k,l} + \sum_{l=1}^{L} \sum_{m=1 \atop (m \neq k)}^{K} \sqrt{P_m}(d_{m,l} - \tilde{d}_{m,l})a_{m,l} + n$$

where $\tilde{d}_{m,l}$ is a soft estimate of the coded symbol of user $k'$ at time $l$, which is available from a previous iteration. The signal $y_k$ is filtered by $w_{k,l}$ at time $l$ to obtain a statistic $z_{k,l}$ for symbol $d_{k,l}$ at time $l$, given by

$$z_{k,l} = w_{k,l}^T y_k = \sqrt{P_k}d_{k,l} + \eta_{k,l}$$

Under some quite general conditions, $\eta_{k,l}$ is well approximated by — and approaches in the limit — an independent Gaussian random variable [4]. One such condition is that the decoder follows the extrinsic information exchange principle [26] and that only extrinsic information is used to generate the cancelled signals.

The filter $w_{k,l}$ can be a matched filter [1], or a more complex conditional or unconditional MMSE filter [37], [4]. Tradeoffs for different filters are discussed in [4] and [25], who show that the per-user MMSE filters can increase the load of an equal-power system by a factor of at most $1 + N/K$. If $w_{k,l} = a_{k,l}$

$$\eta_{k,l} = \sum_{j=1}^{L} \sum_{m=1 \atop (m \neq k)}^{K} \sqrt{P_m}(d_{m,j} - \tilde{d}_{m,j}) (a_{k,l}^T a_{m,j}) + n_{k,l}$$

Assuming unbiased estimates $\tilde{d}_{k,l}$ allows us to calculate the mean and variance of (7). For i.i.d. random spreading with $\tau_m > \tau_k$, $E[\lambda(a_{k,l}^T a_{m,j})^2] = (N - \tau_m + \tau_k)/N$ if $j = l$, and $(\tau_m - \tau_k)/N$ if $j = l - 1$. It follows that

$$\sigma_k^2 = E[\eta_{k,l}] = \frac{1}{N} \sum_{m=1 \atop (m \neq k)}^{K} P_m E[(d_{m,j} - \tilde{d}_{m,j})^2] + \sigma^2$$

which is independent of the time index $l$.

In general, the squared estimation error of $(d_k - \tilde{d}_k)^2$ will be a function of the input signal-to-noise ratio of each decoder, as well as the particular error control code used. We write it formally as

$$E[(d_k - \tilde{d}_k)^2] = g\left(\frac{\sigma_k^2}{P_k}\right)$$

where $g(x)$ is the variance transfer (VT) characteristic of the code, which is a monotonically increasing function with range $[0, 1]$ for any reasonable error control code. We can write (8) as an iterative equation in $v$

$$\sigma_{k,v}^2 = \frac{1}{N} \sum_{m=1 \atop (m \neq k)}^{K} P_m g\left(\frac{\sigma_{m,v-1}^2}{P_m}\right) + \sigma^2$$

and show that $\sigma_{k,v}^2, v = 0, 1, 2, \ldots$ is monotonically decreasing. To see this, formally set $\sigma_{k,0}^2 = \infty$, and assume that $\sigma_{k,v}^2 \geq \sigma_{k,v+1}^2 \geq \sigma_{k,v-1}^2$. Using the monotonicity property of the function $g(x)$ we show that

$$\sigma_{k,v}^2 \leq \frac{1}{N} \sum_{m=1 \atop (m \neq k)}^{K} P_m g\left(\frac{\sigma_{m,v-2}^2}{P_m}\right) + \sigma^2 = \sigma_{k,v-1}^2$$

Under some quite general conditions, $\eta_{k,i}$ is well approximated by — and approaches in the limit — an independent Gaussian random variable [4]. One such condition is that the decoder follows the extrinsic information exchange principle [26] and that only extrinsic information is used to generate the cancelled signals.
By induction the sequence \( \sigma_{k,v}^2 \) is monotonically decreasing \( \forall k \), and, since \( \sigma_{k,v}^2 \geq 0 \), it converges to the limit

\[
\sigma_{k,\infty}^2 = \frac{1}{N} \sum_{m=1}^{K} P_m g \left( \frac{\sigma_{m,\infty}^2}{P_k} \right) + \sigma^2 \quad (12)
\]

Imposing the Lindeberg condition on asymptotic negligibility on the powers of the users, given formally as

\[
\sum_{k} P_k \frac{1}{\sum_{m} P_m} \frac{K}{\infty} \to 0; \quad \forall k \quad (13)
\]

the \( v \)-th iteration variance (10) as well as the limit variance (12) of the iterative process become independent of the user index and approaches for all users the value

\[
\sigma_v^2 = \frac{1}{N} \sum_{m=1}^{K} P_m g \left( \frac{\sigma_{m,v}^2}{P_m} \right) + \sigma^2; \quad (14)
\]

The decisive observation is that \( \sigma_{v,\infty}^2 \) is independent of \( l \) and \( k \) in the large system case with random spreading. This allows the system behavior to be described by a single-parameter dynamical system, as was done in [1, 24, 4].

The dynamics of an equal-power system with \( \sigma_v^2 = f(g(\sigma_{v-1}^2)) \) are illustrated in Figure 2, showing the variance transfer between cancellation, \( \sigma_{v,\text{cif}}^2 \), and decoder, \( \sigma_{v,\text{d}}^2 \), for an \( R = 1/3 \) convolutional code. The trajectory is a simulation for \( K = 45 \) users. Decoding starts at the right top portion of the diagram, and alternates between the the code and canceller curves.

The systems have either one or two stable fixed points. A first one, the interference limitation, occurs at high system loads. If no interference fixed point exists, the iterations proceed through a “bottleneck” region and converge to the noise fixed point. During this process, the BER drops rapidly (with SNR).

At high SNR the BERs follow the single user performance curve, which acts as lower asymptote. This is the noise limitation where performance is limited by the channel noise. These two effects lead to the turbo cliff/error floor behavior which is well known in the context of turbo codes [19].

IV. UNEQUAL RATES AND POWER

Assume that the \( K \) users are separated into \( J \) groups of powers \( P_{1}, \ldots, P_{J} \), with user numbers \( K_{1}, \ldots, K_{J} \), ordered as \( P_{1} \leq P_{2} \leq \cdots \leq P_{J} \). Assuming an “ideal” error control code with SNR threshold \( \tau \), group \( k \) can be decoded via serial cancellation [35] if

\[
\tau P_k \geq \sum_{j=1}^{k} \alpha_j P_j + \sigma^2 \quad (15)
\]

where \( \alpha = \sum_{j=1}^{J} K_j / N = \sum \alpha_j \) is the system load, and \( \alpha_j \) is the partial load of the \( j \)-th group, and we assuming \( K_j \ll N \).

This leads to

\[
P_j = \sigma^2 \tau^{j-1} \prod_{l=1}^{j} \frac{1}{\tau - \alpha_l}; \quad 1 \leq j \leq J \quad (16)
\]

and average power

\[
\bar{P} = \min_{\{\alpha_j\}} \frac{\sum_{j=1}^{J} P_j \alpha_j}{\sum_{j=1}^{J} \alpha_j} \quad (17)
\]

The following lemma is proven in the appendix:

**Lemma 1:** The average power (17) given (16) is minimized by the uniform load distribution \( \alpha_j = \alpha / J; \forall j \).
The weak codes support 1.3 bits/dim at 8.68 dB, 50% of the little difference between the supportable loads of both systems. Figure 3 presents (19) using both strong concatenated codes and weak 4-state convolutional codes for three power groups, distributed as $P_j = P_1, P_2 = 2P, P_3 = 4P$. There is little difference between the supportable loads of both systems. The weak codes support 1.3 bits/dim at 8.68 dB, 50% of the AWGN capacity.

The following lemma shows that the attainable spectral efficiency increases with the number of groups.

**Lemma 2**: The sequence $(1 - \alpha/(J \tau))^J$ is monotonically decreasing with $J$ and converges to $e^{\alpha/\tau}$ in the range $[J^*, \infty)$, where $J^*$ denotes the smallest integer such that $\alpha/(J^* \tau) < 1$.

**Proof**: See appendix.

In the limit as $J \to \infty$ the required threshold $E_b/N_0$ has to obey

$$
\frac{E_b}{N_0}_{\text{lim}} = \lim_{J \to \infty} \frac{1}{2R\alpha} \left[ \left( \frac{\tau}{\tau - \alpha/J} \right)^J - 1 \right] = e^{\alpha/\tau} - 1
$$

where we have used $(1 - 1/x)^J = e$ for $x \to 0$. Assuming ideal codes with $1/\tau = 2^{2R} - 1$ in (20) we get for $R \to 0$

$$
\frac{E_b}{N_0}_{\text{lim}} = \lim_{R \to 0} \frac{e^{\alpha/(2^{2R} - 1)} - 1}{2R\alpha} = \frac{2^{2C} - 1}{2C}
$$

where $C = R\alpha$ is the total spectral efficiency and we applied the rule of de l’Hôpital to the 0/0 limit in (21).

Instead of optimizing the power profile, we may assign different code rates to different groups to achieve a similar effect. All users transmit with power $P$, and we assign rates $r_j : r_1 > r_2, \ldots, > r_J$ to $J$ groups. Since the SNR $P/\sigma_j^2 = 1/(4^{r_j} - 1)$ for Shannon-type ideal codes is monotonically decreasing with $r_j$, lower rate users converge before higher rate users. The power constraint equation for group $j$ is given by $r_j P \geq \sum_{m=1}^{j} \alpha_m P + \sigma^2$, which implies the rate constraints

$$
r_j \leq \frac{1}{2} \log \left( 1 + \frac{1}{\sum_{m=1}^{j} \alpha_m + b} \right); \quad b = \sigma^2/P
$$

The variational problem at hand is to maximize the sum of the rates $r_j$ for a given system load, i.e., maximize

$$
g(\{\alpha_j\}, J, b) = \frac{1}{2} \sum_{j=1}^{J} \alpha_j \log \left( 1 + \frac{1}{\sum_{m=1}^{j} \alpha_m + b} \right)
$$

given that $\sum_{j=1}^{J} \alpha_j = \alpha$

over all choices $\{\alpha_j\}$.

The following lemma gives implicit upper and lower bounds on the achievable sum rate.

For finite groups (23) can be formulated as a linear programming optimization problem similar to [7] and solved numerically.
Fig. 3. Illustration of different power levels and effective VT characteristics shown for a serial turbo code and a convolutional code, both using three equal-sized power groups with ratios $P_1/P_2/P_3 = 1/2/4$. The unequal power alleviates the disadvantage of the strong code system. The trajectory is a simulation.

**Lemma 3:** Let $\alpha_0 = \max_j \alpha_j$, then the following inequalities hold:

$$
\frac{1}{2} \int_0^\alpha \log \left(1 + \frac{1}{u + \alpha_0 + b}\right) \, du \leq g(\{\alpha_j\}, J, b)
$$

$$
\leq \frac{1}{2} \int_0^\alpha \log \left(1 + \frac{1}{u + b}\right) \, du
$$

The latter integral can be evaluated as $(\alpha + b + 1) \log(\alpha + b + 1) - (\alpha + b) \log(\alpha + b) - (b + 1) \log(b + 1) + b \log b$.

**Proof:** see appendix.

We can draw some interesting conclusions from Lemma 3. As long as $\alpha_0 \ll b$, i.e., the maximum partial load is small w.r.t. the inverse $b$ of the SNR, any load distribution gives the same maximum sum rate. A corollary to this is that the case of a single group, the lower limit in (24) is minimized, i.e., the equal rate case is the worst case for cancelation using strong codes.

The achievable capacity for $a_0/b \to 0$ is given by

$$
C(\alpha, b) = \frac{\alpha + b + 1}{2} \log(\alpha + b + 1) - \frac{\alpha + b}{2} \log(\alpha + b)
$$

$$
= \frac{b + 1}{2} \log(b + 1) + \frac{b}{2} \log b
$$

$$
= \frac{1}{2} \log \left(1 + \frac{\alpha}{b + 1}\right) + \frac{\alpha + b}{2} \log \left(1 + \frac{1}{\alpha + b}\right)
$$

Expressing $E_b/N_0$ in terms of capacity, signal power, and noise variance as

$$
\frac{E_b}{N_0} = \frac{\alpha}{2bC(\alpha, b)}; \quad b = \frac{\sigma^2}{P}
$$

and considering large loads, i.e., $\alpha \to \infty$, we necessarily have low total SNRs, that is, $b \to \infty$. In this case

$$
C(\alpha, b) = \frac{1}{2} \log \left(1 + \frac{\alpha}{b}\right) - \frac{\alpha \log e}{4b(\alpha + b)} + O \left(\frac{1}{b^2}\right), \quad b \to \infty
$$

Furthermore using the limiting approximation $\alpha/b \approx 2^{2C(\alpha, b)/C(\alpha + b)} - 1$ in (25) we obtain

$$
\frac{E_b}{N_0} \approx \frac{(2^{2C} - 1)}{2C} \left[1 + \frac{(2^{2C} - 1)}{2\alpha}\right], \quad \alpha \to \infty
$$

where we have used the abbreviation $C = C(\alpha, b)$. For large system loads $\alpha$ (27) again assumes the familiar form of the Shannon bound, similar to (21), which we express in the following.

**Theorem 1:** Ideal (equal-power) codes can achieve the AWGN Shannon bound as long as the number of user rate groups is sufficiently large, i.e.,

$$
\frac{E_b}{N_0} \to \frac{(2^{2C} - 1)}{2C}; \quad \alpha \to \infty
$$

for $\alpha \to \infty$, and given the maximum partial load obeys $P_{\alpha_0}/\sigma^2 \to 0$.

**V. EQUAL POWER AND EQUAL RATE USERS**

The system with a single power group, $J = 1$, with $P_j = P$, and equal rates, presents the most challenging situation for an iterative cancellation receiver, since neither power nor rate distributions can be used to maximize spectral efficiency.
In this case weaker error control codes, which exhibit a steeper VT curve than stronger codes [24] (Figure 2), achieve significantly higher loads $\alpha$. In a weak code, however, the distinction between interference and noise limitation is blurred. In fact, no non-trivial bound on the code VT curve exists, which can be seen as follows: Consider a hypothetical code where every information sequence is mapped into the same codeword. An APP decoder can now perfectly estimate the transmitted codewords, and the VT curve of such a code would coincide with the vertical axis, allowing $\alpha \rightarrow \infty$. However, the (information) error performance of such a code is catastrophic since it is not possible to decode the information sequences from the coded sequences.

Figure 4 shows the variance transfer curve of a $R = 1/3$ repetition code. The distinction between interference and noise limitation has vanished, and low BER performance is not possible due to the weak nature of the code independently of the load.

![Interference cancellation with a weak rate $R = 1/3$ repetition code, acting as non-linear layering filter.](image)

The intersection point which corresponds to the fixed point of the iterative system equation corresponds to a certain variance in the log-likelihood ratios of the decoded information bits. As shown below, this LLR is Gaussian distributed, and we therefore have a binary-input, Gaussian-output channel between input information bits and their corresponding output information bit LLRs. The spectral efficiency of this channel is

$$ C = \alpha R C_B \left( \frac{P}{\sigma_{ll}^2(\alpha, C)} \right) $$

(29)

where $C_B(P/\sigma_{ll}^2(\alpha, C))$ is the capacity of the binary-input Gaussian noise channel, and $\sigma_{ll}^2$ is the variance of the Gaussian LLR of the information bit at the decoder output at the iteration fixpoint. The output signal-to-noise ratio $P/\sigma_{ll}^2$ does depend on the system load $\alpha$ as well as on the particular code $C$ used.

In general it is difficult to calculate the code VT function, and this makes optimization over the codes $C$ in (29) hard. The VT function of many different error control codes have been studied, a selection of which are shown in Appendix 2, and the weaker codes show steeper VT functions with less pronounced thresholds.

Based on this, we focus on repetition codes, which were observed to provide the largest system spectral efficiencies among a large number of codes studied [25]. The output extrinsic LLR of any of the coded bits in a length-$L$ repetition code at iteration $m$ is simply

$$ \lambda^{(E)}(d_i) = \sum_{j \neq i} \lambda(y_j) = \frac{2}{\sigma_m^2} \sum_{j \neq i} y_j $$

(30)

and the information bit LLR can be calculated as

$$ \lambda(u) = \sum_{i=1}^{L} \lambda(y_i) $$

(31)

From the extrinsic LLRs $\lambda^{(E)}(d_i)$ of the coded bits the decoder computes the soft bits $\hat{d}_k(i) = \tanh(\lambda^{(E)}(d_i)/2)$ used for cancellation. Decoding is simple since most of the complexity lies in the tanh() function.

The PDFs of both the a priori and extrinsic LLRs are Gaussian due to (30) and (31) with variance

$$ \sigma_E^2 = \frac{\sigma_m^2}{P(L-1)} \quad \text{and} \quad \frac{\sigma_h^2}{P} = \frac{\sigma_m^2}{PL} $$

(32)

respectively. However, the PDF of the soft-bits $\hat{d}_i(k)$ is not Gaussian. It is given by

$$ p(d|\hat{d}_i(k) = 1) = \frac{\exp \left( \frac{1}{2\sigma_E^2} - \left( \log \left( \frac{1+d}{1-d} \right) - \frac{\sigma_d^2}{\sigma_E^2} \right)^2 \right)}{\sqrt{2\pi\sigma_E^2}(1-d^2)} $$

(33)

Applying the central limit theorem to (5) as $K \rightarrow \infty$ and the fact that $\text{var}(d_i(k)) \leq 1$, the residual interference term is converging to a Gaussian distribution with variance $\alpha \sigma_d^2$, where

$$ \sigma_d^2 = \int_{-1}^{1} p(d|\hat{d}_i(k) = 1)(1-d)^2 = g_d(\sigma_E^2) $$

(34)

which has to be evaluated by numerical methods. Equations (34) and (32) lead to the following lemma:

**Lemma 4:** The system spectral efficiency in (29) using repetition codes is maximized for $L \rightarrow \infty$.

**Proof:** Let the point $(x, y) = (g(\sigma_{ll}^2), \sigma^2)$ be the intersection between the load line $\sigma^2 + \alpha \sigma_d^2$ and the variance transfer curve $g(\sigma_m^2/P)$ of the repetition code, i.e.,

$$ \alpha x - \sigma^2 = g_d^{-1}(x) $$

(35)

The spectral efficiency of the layered channels is then given from (29) as

$$ C = \alpha \frac{1}{L} C_B( \frac{LP}{\sigma_{ll}^2} ) = \frac{1}{L} C_B \left( \frac{LP}{\sigma_m^2} \right) $$

$$ = \frac{(L-1)\sigma_E^2 - \sigma^2}{g_d(\sigma_E^2)} \frac{1}{L} C_B \left( \frac{L}{L-1}\sigma_E^2 \right) $$

(36)
which can easily be shown to be monotonically increasing with \( L \).

For the repetition code we obtain from (9) and (30)

\[
d_i(k) = \tanh \left( \frac{P}{\sigma_m^2} \sum_{j \neq i} y_j \right)
= \tanh \left( \frac{(L-1)P}{\sigma_m^2} + \sqrt{\frac{(L-1)P}{\sigma_m^2}} \xi \right)
\]  

(37)

where \( \xi \sim \mathcal{N}(0, 1) \). Now

\[
g \left( \frac{\sigma_m^2}{P} \right) = E \left[ 1 - \tanh \left( b^2 + b\xi \right) \right]^2 ;
\]

(38)

\[
b = \sqrt{\frac{(L-1)P}{\sigma_m^2}}
\]

For our further development, we will make use of some fundamental approximation, given in the following

**Lemma 5:** This lemma consists of two parts:

i) The function \( g(b) \), \( b \geq 0 \) is monotonically decreasing, and for any \( b > 0 \) the following relation holds

\[
E \left[ 1 - \tanh \left( b^2 + b\xi \right) \right]^2 = 4e^{-b^2/2} \sum_{n=0}^{\infty} (-1)^n e^{b^2(2n+1)/2} Q \left[ b(2n+1) \right]
\]

(39)

where \( Q(x) = 1/(\sqrt{2\pi}) \int_x^\infty e^{-u^2/2} du \) is the Gaussian error function.

ii) For any \( b \geq 0 \) the following inequality is true

\[
E \left[ 1 - \tanh \left( b^2 + b\xi \right) \right]^2 \leq \min \left\{ \frac{1}{1+b^2 \pi Q(b)}, \frac{1}{1+b^2} \right\}
\]

(40)

**Remarks.** Equation (40) taken as an approximation has an accuracy better than 90\% for all \( b \geq 0 \). Note also that \( \lim_{b \to \infty} g(b)/(\pi Q(b)) = 1 \). For the proof see [5].

Using equal powers \( P_j = 1 \), we find \( \sigma^2_{m+1} = \alpha g(\sigma_m^2) + \sigma^2 \), and using

\[
g \left( \sigma_m^2 \right) \leq \frac{1}{1 + \frac{L-1}{\sigma_m^2}}
\]

(41)

by applying the first upper bound (40), we obtain

\[
\sigma^2_{m+1} \leq \frac{\alpha \sigma^2_m}{\sigma_m^2 + L-1} + \sigma^2
\]

(42)

For the unique fixed point variance we obtain the simple quadratic boundary equation

\[
\sigma^2_{\infty} = \frac{\alpha \sigma^2_{\infty}}{\sigma^2_{\infty} + L-1} + \sigma^2
\]

(43)

which has the solution

\[
\sigma^2_{\infty} = \frac{1}{2} \left( \alpha + \sigma^2 + (L-1) + \sqrt{(\alpha + \sigma^2 + (L-1))^2 + 4(L-1)\sigma^2} \right)
\]

(44)

Defining the *uncoded* load \( \alpha' = \alpha/(L-1) \), i.e., the load with respect to the information bit rate, we obtain

\[
\sigma^2_{\infty} \leq \frac{L-1}{2} \left( \alpha' + 1 + \frac{\sigma^2}{L-1} + \sqrt{\left( \alpha' + 1 + \frac{\sigma^2}{L-1} \right)^2 + 4\frac{\sigma^2}{L-1}} \right)
\]

(45)

Note that if we apply a (single-shot) MMSE filter [38], [13] to the received signal, the output variance of each user’s signal after filtering is given by [33], [32],

\[
\sigma^2_{\text{mmse}} = \frac{1}{2} \left( \alpha - 1 + \sigma^2 + \sqrt{(\alpha - 1 + \sigma^2)^2 + 4\sigma^2} \right)
\]

(46)

From (45) and (46) we see that the MMSE filter with load \( \alpha \) and noise variance \( \sigma^2 \) can directly be related to the performance of the iterative filter with load \( \alpha/(L-1) \) and noise variance \( \sigma^2/(L-1) \), which we express in

**Theorem 2:** Iterative cancellation of equal-power, equal-rate random CDMA users using rate \( 1/L \) repetition codes with \( L \to \infty \), has a per-chip capacity at least as large as linear minimum-mean square error filtering with equality for

\[
\alpha > 2(L-1) - 2\sigma^2/P.
\]

**Proof:** The variance (45) of an iterative receiver is at most as large as that of an MMSE filter with load \( \alpha/(L-1) \) and noise variance \( \sigma^2/(L-1) \). Assuming BPSK modulation for both, we obtain

\[
C = \frac{\alpha}{L} C_B \left( \frac{L^D}{\sigma_{\infty}} \right)
\]

and

\[
C_{\text{mmse}} = \frac{\alpha}{L-1} C_B \left( \frac{(L-1)P}{\sigma_{\infty}} \right)
\]

(47)

for iterative layering and MMSE filtering, respectively. Equality in (44) holds for iterative cancellation if equality holds in the first inequality of (40), i.e., if \( \sigma^2_{\infty} > P(L-1) \). From this and (44) we easily derive the condition of the theorem.

In fact, more can be said. Relating (10) and (37) to results applying statistical mechanics to CDMA detection [28], [17], we can show that the performance of the iterative canceler for low and moderate loads is essentially identical to that of (uncoded) individually optimal detection (APP) detection. Using [28, Eqns. 45 and 43] the signal-to-interference ratio \( \gamma^2 \) of the APP detector can be computed as solution to the implicit equation

\[
\gamma^2 = \left[ \sigma^2 + \alpha E \left[ 1 - \tanh \left( \frac{F}{E} (\gamma^2 + \gamma \xi) \right) \right] \right]^{-1}
\]

(48)

where \( E \) and \( F \) are given in [28, Eqn. 43], and are the mean and variance of the (individual) APP output signal. In general \( E > F \), however in our case \( E = F \), which follows from the definitions of \( m, q, E, F \) in [28, Eqn. 27 ff.]. The result (48) holds for loads \( \alpha < \alpha_s \approx 1.49 \), where \( \alpha_s \), the “spinodal value”, is the largest number for (35) has a unique solution – see also Figure 4. This result has been put on rigorous footing recently in [17] using the concept of “sparse spreading”. We can now relate (48) to the variance equation for the iterative
detector, which, from above, is given by
\[
\frac{\sigma^2}{L-1} = \frac{\sigma^2}{L-1} + \frac{\alpha}{L-1} E \left[ 1 - \tanh \left( \frac{L-1}{\sigma^2} \left( \frac{L-1}{\sigma^2} + \sqrt{\frac{L-1}{\sigma^2}} \xi \right) \right) \right]^2
\] (49)

From (48) and (49) we conclude the following

**Proposition 1:** Iterative cancellation of equal-power, equal-rate random CDMA using rate 1/L repetition codes has a performance which equals that of an optimal (uncoded detector) with system load \( \alpha/(L-1) \) and noise variance \( \sigma^2/(L-1) \) for all \( \alpha < (L-1)\alpha_s \), where \( \alpha_s \) is the spinodal value of (35) and \( \alpha_s \approx 1.49 \) for high SNR.

Note that Theorem 2 and Proposition 1 are complementary in that they address different load regimes. These are related to the co-existence of solutions and phase changes observed in [28]. In essence, for large \( \alpha \), the performance is given by Theorem 2, and is equivalent to MMSE, while in the low-load regime where Proposition 1 is valid, optimal APP performance can be achieved.

**Remark:** Spreading can be viewed as repetition coding of rate \( N \), and hence, the channel access mechanism can be viewed as applying repetition coding of rate \( 1/(NL) \), partitioned into an outer code of rate \( 1/L \) and an inner code of rate \( 1/N \), i.e., a two-level spreading/repetition coding. Only the outer code is processed optimally, while the inner code is used for cancellation only. Given a system load of \( K/(NL) \), it can be seen from (45), and (31), that the per-channel output variance is minimized for large \( L \), i.e., \( L \to \infty \) is optimal. However, the variance quickly saturates as \( L \) grows, suggesting a practically optimal partition between the inner and outer codes with large inner spreading.

We are grateful to an anonymous reviewer for suggesting this point of view, as well to a second reviewer for suggesting the link with the statistical mechanics results.

**VI. CONCLUSION**

Starting from iterative joint CDMA detector with unequal power and rate distributions we have examined conditions on power and rate distributions which allow such systems to achieve the Shannon bound for additive white Gaussian noise channels, and we have characterized optimal group sizes and rate and power values required. If such power or rate allocation methods are possible, it was demonstrated that the error control code used plays a secondary role in attaining high spectral efficiencies.

However, when both power and rates are fixed at equal levels for all users, linear interference cancellation will be operating under its worst-case conditions. In the absence of being able to maximize powers or rates, the error control codes are the only system variables. We demonstrated that in this case powerful error control codes are counter indicated, and that weak error control codes achieve significantly larger system capacities. In particular, we showed that simple repetition codes, resulting in a two-level partition of inner/outer spreading with simple cancellation processing, achieve the performance of a much more complex optimal a posteriori probability (APP) detector for uncoded transmission for low and medium system loads, and achieves a total system capacity which is tightly lower-bounded by that of MMSE filtering for high system loads or low values of the SNR.

**APPENDIX I**

**PROOF OF LEMMA 1**

We wish to minimize the average power
\[
P = \frac{\sigma^2}{\alpha} \sum_{i=1}^{N} \frac{\alpha_{i-1}}{\prod_{l=1}^{i} (\tau - \alpha_l)} \quad \text{given} \quad \sum_{i=1}^{N} \alpha_i = \alpha
\]

Denoting \( x_i = \frac{\alpha}{\tau} \), \( i = 1, \ldots, N \) we minimize
\[
P_{\min} = \min_{x_1, \ldots, x_N} \frac{\sigma^2}{\alpha} \sum_{i=1}^{N} \frac{x_i}{\prod_{l=1}^{i} (1 - x_l)} \quad \text{given} \quad \sum_{i=1}^{N} x_i = \frac{\alpha}{\tau}
\]

Note that
\[
\sum_{i=1}^{N} x_i \prod_{l=1}^{i} (1 - x_l) = \sum_{i=1}^{N} \frac{(x_i - 1)}{\prod_{l=1}^{i} (1 - x_l)} + \sum_{i=1}^{N} \frac{1}{\prod_{l=1}^{i} (1 - x_l)}
\]
\[
= \sum_{i=1}^{N} \frac{1}{\prod_{l=1}^{i} (1 - x_l)} - \sum_{i=2}^{N} \frac{1}{\prod_{l=1}^{i-1} (1 - x_l)} = \frac{1}{\prod_{l=1}^{N} (1 - x_l)} - 1
\]
\[
= \exp \left\{ -N \sum_{i=1}^{N} \frac{1}{N} \ln(1 - x_i) \right\} - 1
\]
\[
\geq \exp \left\{ -N \ln \left( 1 - \frac{1}{N} \sum_{i=1}^{N} x_i \right) \right\} - 1
\]
\[
= \exp \left\{ -N \ln \left( 1 - \frac{\alpha}{N\tau} \right) \right\} - 1 = \frac{1}{(1 - \alpha/(N\tau))^N} - 1
\]
where we have used the standard inequality \( \ln x \leq x - 1 \), and equality is achieved for \( \alpha_1 = \ldots = \alpha_N \).

**APPENDIX II**

**PROOF OF LEMMA 2**

Since
\[
\left( \frac{\tau}{\tau - \alpha/N} \right)^N = \exp \left\{ N \ln \frac{\tau}{\tau - \alpha/N} \right\}
\]
we consider the function
\[
f(N) = N \ln \frac{\tau}{\tau - \alpha/N}
\]
Using the standard inequality \( \ln x \leq x - 1 \), we have
\[
f'(N) = \ln \frac{\tau}{\tau - \alpha/N} - \frac{\alpha}{N\tau - \alpha} \leq 0
\]
from which Lemma 2 follows.
APPENDIX III
PROOF OF LEMMA 3

We consider the following variational problem

\[
\min_{\alpha_1, \ldots, \alpha_J} \sum_{j=1}^{J} \alpha_j \text{ given } g(\{\alpha_j\}, J, b)
\]

\[
= \sum_{j=1}^{J} \alpha_j \ln \left(1 + \frac{1}{\sum_{i=1}^{j} \alpha_i + b}\right)
\]

\[
= C, \quad b = \sigma^2 / P \quad (50)
\]

The problem (50) is equivalent to

\[
\max_{\alpha_1, \ldots, \alpha_J} \sum_{j=1}^{J} \alpha_j \ln \left(1 + \frac{1}{\sum_{i=1}^{j} \alpha_i + b}\right)
\]

provided \(\sum_{j=1}^{J} \alpha_j = A\) \quad (51)

We first note that the solution to both (50) and (51) is monotonically increasing, i.e., \(\alpha_1 < \ldots < \alpha_K\).

This is shown as follows: Let \(\alpha_1, \ldots, \alpha_{J+1}\), and consider the the minimization with respect to \(\alpha_j, \alpha_{j+1}\), assuming that \(\alpha_j + \alpha_{j+1} = A\) and all other \(\{\alpha_i\}\) remain fixed. Then, letting \(\alpha_1 = \alpha\) it is sufficient to consider the maximization over \(\alpha\) of

\[
g(\alpha_1, \alpha_2) = g(\alpha)
\]

\[
= \alpha \ln \left(1 + \frac{1}{\alpha + B}\right)
\]

\[
+ (A - \alpha) \ln \left(1 + \frac{1}{A + B}\right)
\]

where \(B\) is constant. We observe that

\[
g'(\alpha) = \ln \left(1 + \frac{1}{\alpha + B}\right) - \ln \left(1 + \frac{1}{A + B}\right)
\]

\[
- \frac{(\alpha + B)(\alpha + B + 1)}{\alpha (\alpha + B + 1)}
\]

\[
g''(\alpha) < 0, \quad g'(0) > 0, \quad g'(A/2) < 0
\]

Then maximum of \(g(\alpha)\) is attained for some \(0 < \alpha < A/2\), from which the statement follows.

Now we obtain lower and upper bounds for \(g(\{\alpha_j\}, J, b)\) that work well for large \(J\) and well-behaved \(\{\alpha_j\}\).

The following statement holds: With \(\alpha_0 = \max_j \alpha_j\) the inequalities

\[
\int_{0}^{u_1} \ln \left(1 + \frac{1}{u + \alpha_0 + b}\right) \, du \leq g(\{\alpha_j\}, J, b)
\]

\[
\leq \int_{0}^{u_1} \ln \left(1 + \frac{1}{u + b}\right) \, du \quad (52)
\]

hold, where \(u_1 = \sum_{j=1}^{J} \alpha_j\).

This part is shown as follows: For a non-decreasing function \(\alpha(x), x \geq 0\) with \(\alpha(0) \geq 0\) define the functional

\[
g(\alpha(x), J, b) = \int_{0}^{J} \alpha(z) \ln \left(1 + \frac{1}{\int_{0}^{J} \alpha(u) \, du + b}\right) \, dz
\]

Also, for a non-negative sequence \(\alpha_1, \ldots, \alpha_J\) let \(\pi(x), 0 \leq x \leq J\) be the step-wise function such that \(\pi(j) = \alpha_j, j - 1 < x \leq j, j = 1, \ldots, J\). Then for \(j = 1, \ldots, J\) we have

\[
\sum_{i=1}^{j} \alpha_i = \int_{0}^{\pi(x)} dx,
\]

and

\[
\alpha_j \ln \left(1 + \frac{1}{\sum_{i=1}^{j} \alpha_i + b}\right) = \int_{j-1}^{j} x \ln \left(1 + \frac{1}{\int_{0}^{x} \alpha(u) \, du + b}\right) \, dx
\]

\[
\leq \int_{j-1}^{j} \pi(x) \ln \left(1 + \frac{1}{\int_{0}^{\pi(x)} \alpha(u) \, du + b}\right) \, dx
\]

and therefore

\[
g(\{\alpha_j\}, J, b) \leq \int_{0}^{\pi(x)} \ln \left(1 + \frac{1}{\int_{0}^{x} \alpha(u) \, du + b}\right) \, dx =
\]

\[
= \int_{0}^{J} \ln \left(1 + \frac{1}{\int_{0}^{x} \alpha(y) \, dy + b}\right) \, d \left(\int_{0}^{x} \pi(y) \, dy\right)
\]

\[
= \int_{0}^{u_1} \ln \left(1 + \frac{1}{u + b}\right) \, du =
\]

\[
= (u_1 + b + 1) \ln (u_1 + b + 1) - (u_1 + b) \ln (u_1 + b) - (b + 1) \ln (b + 1) + b \ln b
\]

where

\[
u_1 = \int_{0}^{J} \pi(y) \, dy = \sum_{j=1}^{J} \alpha_j
\]

from which the right-hand side of inequalities (52) follows. On the other hand, we have

\[
\alpha_j \ln \left(1 + \frac{1}{\sum_{i=1}^{j} \alpha_i + b}\right)
\]

\[
\geq \int_{j-1}^{j} \pi(x) \ln \left(1 + \frac{1}{\int_{0}^{\pi(x)} \alpha(u) \, du + b}\right) \, dx
\]

from which the left-hand side of inequalities (52) follows.

For a given \(\sum_{j=1}^{J} \alpha_j\) and large \(J\) both bounds (52) are close to each other if \(\alpha_0 \ll b\). Note that the function

\[
f(\alpha) = \alpha \ln \left(1 + \frac{1}{\alpha + B}\right), \quad B \geq 0
\]

increases monotonically with \(\alpha \geq 0\).
Appendix IV

Variance-Transfer Curves

The figure below shows the VT curves of a collection of rate $R = 1/2$ LDPC codes. As can be seen, the shape of the curves can be manipulated somewhat by the choice of degree profiles, and thus quite different system loads are achievable with equal-power systems. It also becomes evident that using a strong LDPC code approaches the limiting, step-function VT curve of an ideal Shannon-type code. Furthermore, note that increasing the number of iterations of the LDPC code can only improve performance and load, since a VT curve associated with a higher iteration number majorizes that with a lower iteration number.

![Variance-Transfer Curves](image)

Fig. 5. Illustration of variance transfer curves of various strong practical-sized LDPC codes of length $N = 5000$, and code rate $R = 0.5$, as well as of a weaker convolutional code, confirming the observation that the more powerful codes exhibit a more step-like transfer function.

References


[27] A. Sanderovich, M. Peleg and S. Shamai, “LDPC Coded
Christian Schlegel received the Dipl. El. Ing. ETH degree from the Federal Institute of Technology, Zurich, in 1984, and the M.S. and Ph.D. degrees in electrical engineering from the University of Notre Dame, Notre Dame, IN, in 1986 and 1989.

He held academic positions at the University of South Australia, the University of Texas and the University of Utah. In 2001 he was named iCORE Professor for High-Capacity Digital Communications at the University of Alberta, Canada, a 5 Mio$ chaired research program. His interests are in the area of error control coding and applications, multiple access communications, digital communications, as well as analog and digital implementations of communications system. He is the author of the research monographs “Trellis Coding” and “Trellis and Turbo Coding” by IEEE/Wiley, as well as “Coordinated Multiple User Communications”, co-authored with Professor Alex Grant, published by Springer. Dr. Schlegel received an 1997 Career Award, and a Canada Research Chair in 2001.

Dr. Schlegel is associate editor for coding theory and techniques for the IEEE Transactions on Communications, and a guest editor of the IEEE Proceedings on Turbo Coding. He served as the technical program co-chair of ITW 2001 and ISIT 05. He was also the general chair of the CTW 05, as well as member of numerous technical conference program committees.

Marat Burnashev received his M.S. and Ph.D. degrees from the Moscow Institute of Physics and Technology in 1971, and 1974, respectively, and the Dr. Math. Science degree from the Steklov Mathematical Institute in 1985.

From 1974 to 1986 he was a research staff member at the Institute for Problems of Information Transmission at the USSR Academy of Sciences, where he held a Leading Researcher position since 1986. He was also an Assistant Professor at the Moscow Institute of Physics and Technology from 1978 to 1980, and a Full Professor at the Moscow Institute of Communications from 1986-1988.

Professor Burnashev has held numerous visiting and adjunct positions over the years at the Politecnico di Torino, the University of Notre Dame, Tokyo University, the Universities of North Carolina and Washington, Bielefeld University, Universite du Maine, University de Paris, the Universities of Utah and California, Riverside, the EPFL, Lausanne, and the University of Alberta, Canada.