

## Ergodic Sum Capacity Maximization for CDMA: Optimum Resource Allocation

Onur Kaya, *Student Member, IEEE*, and Sennur Ulukus, *Member, IEEE*

**Abstract**—We solve for the optimum signature sequence and power allocation policies that maximize the information-theoretic ergodic sum capacity of a code-division multiple-access (CDMA) system subject to fading. We show that at most  $N$  users may transmit at any given channel state, where  $N$  is the processing gain; and those users who are transmitting should be assigned orthogonal signature sequences. We also show that the power allocation policy that maximizes the capacity together with the choice of these signature sequences is single-user water-filling over sets of channel states that are favorable to each user. That is, the capacity maximizing signaling scheme is shown to dictate that the users allocate their powers and signature sequences in such a way that they always avoid interference from each other.

**Index Terms**—Code-division multiple access (CDMA), fading channels, interference avoidance, iterative water-filling, power control, signature sequence optimization, sum capacity.

### I. INTRODUCTION

An important consideration in the design of wireless communication systems is the unavoidable presence of fading, caused by the nature of the system. To maximize the overall network capacity, one should therefore exploit the variations in the channel fade levels while allocating the available resources. Resource allocation for wireless systems can be viewed in several different contexts, including the signal-to-interference ratio (SIR)-based approaches and information theoretic approaches (see [2] for some references on both). We can also group the allocation schemes according to the type of resources that are allocated. Two important resources are the transmit powers and the transmit waveforms. Power control has been studied in SIR-based and information-theoretic contexts for fading and nonfading channels [2]–[5], whereas waveform optimization has been studied in these two contexts for nonfading channels only [6]–[8]. Throughout this correspondence, the objective of resource allocation is to maximize the information-theoretic ergodic (expected) sum capacity, and we consider allocating both powers and the waveforms as functions of the channel state information (CSI) in order to achieve this objective. The system of interest is assumed to be a code-division multiple-access (CDMA) network; thus, waveforms will simply be referred to as signature sequences.

The problem of power control in the context of capacity maximization for fading channels is first studied in [3] for a single-user channel, and it is shown that the optimum power allocation policy that maximizes the ergodic channel capacity subject to an average power constraint is a water-filling of power over the inverse of the fade levels. In this policy, more power is allocated to stronger channel states; and no power is allocated for channel states below some threshold.

Manuscript received July 17, 2003; revised October 21, 2004. This work was supported by the National Science foundation under Grants ANI-0205330 and CCR-0311311; and by ARL/CTA under Grant DAAD-19012-0011. The material in this correspondence was presented in part at IEEE Global Telecommunications Conference, San Francisco, CA, Dec. 2003.

The authors are with the Department of Electrical and Computer Engineering, University of Maryland, College Park, MD 20742 USA (e-mail: onurkaya@eng.umd.edu; ulukus@eng.umd.edu).

Communicated by G. Caire, Associate Editor for Communications.  
Digital Object Identifier 10.1109/TIT.2005.846413

Although we can define the capacity for a single-user channel, when we move onto multiple-access channels (MACs), the notion of capacity has to be replaced by a region of achievable rates [9], [10]. In such channels, it is customary to consider the maximum achievable sum of rates, i.e., the *sum capacity* as a figure of merit. For a multiuser scalar channel, where all users transmit with the same waveform, [4] finds the optimum power allocation policy which maximizes the ergodic sum capacity. There, it is shown that, to maximize the sum capacity, the users perform single-user water-filling over disjoint sets of channel states. That is, each user transmits only when its channel state (normalized by a factor) is greater than or equal to that of all other users. Since the channel fading is a continuous random variable, the equality of the two ratios corresponds to a zero probability event. Thus, for such a channel, at most one user transmits at a given channel state with probability 1.

For vector multiple-access channels, such as CDMA, [11] proposes an asymptotically optimal single-user water-filling strategy to maximize the ergodic sum capacity in the special case of a large system with random signature sequences. The more generalized version of the power control problem with arbitrary signature sequences and arbitrary number of users and processing gain is solved in [2], where the solution is shown to be a *simultaneous water-filling* of powers of users, and an iterative algorithm which performs a one-user-at-a-time single-user water-filling for each user, while the powers of all other users are fixed, is shown to converge to the optimal solution. The optimum power allocation in that case is shown to dictate more than one user to transmit simultaneously in certain regions of the channel state space, provided that the signature sequences satisfy some mild conditions.

The sum capacity of a CDMA network can also be optimized as a function of the signature sequences. When each user has an average power constraint, and there is no fading in the system, [6] shows that when the number of users is less than or equal to the processing gain, the optimal strategy is to allocate orthogonal signature sequences to all users, and when the number of users is greater than the processing gain, with all users having the same average power constraint, the optimal strategy is to allocate Welch bound equality (WBE) [12] sequences. [7] generalizes [6] to arbitrary (unequal) average power constraints, and gives the optimal signature sequence allocation as a function of the power constraints of the users. Specifically, for the case in which the number of users is greater than the processing gain, when a user has a *relatively larger power constraint* than the others, it is called “oversized,” and such users are allocated orthogonal signature sequences; whereas the “nonoversized” users are allocated the so-called generalized Welch-bound-equality (GWBE) sequences.

In this correspondence, we attack the problem of joint power and signature sequence optimization in order to maximize the ergodic sum capacity of a fading CDMA system. Specifically, we adapt the set of signature sequences and transmit powers of all users as a function of the CSI, in order to maximize the ergodic sum capacity. At each fading state, for any given arbitrary power allocation, results of [7] can be used to allocate the optimal sequences. Among those power allocations, with signature sequences chosen optimally, we find the best power allocation strategy.

We show that the optimal strategy is still a water-filling strategy for each user, and very strikingly, at each fading state, that strategy dictates that we allocate (at most)  $N$  orthogonal signature sequences to the users with best (at most)  $N$  channel states (scaled by a factor as in [4]). Moreover, the other users with worse channel states than the users with orthogonal sequences do not transmit at those particular channel states.

This means that there are no users in the system which are allocated GWBE sequences and are yet transmitting with nonzero powers. Thus, in contrast to signature sequence optimization for nonfading channels, GWBE sequences are never used in transmissions; more precisely, they are used only with zero probability.

Our solution resembles [4] in the sense that there is an ordering of channel states that determines which users will transmit, but it also resembles the solution in [3] in that once we know which users will transmit at each channel state, all users will choose their powers as if they were alone in the system, i.e., they will perform single-user water-filling over favorable regions of the channel state space. This result shows that, when we have the opportunity to control both the signature sequences and the powers of the users, the users completely avoid each other, i.e., certain groups of users transmit on disjoint sets of channel states, and within each group of users that transmit at the same channel state, users place themselves orthogonal to each other in the signature sequence space, thus avoiding any possible interference.

Finally, we provide an iterative algorithm that is guaranteed to converge to the optimal power and signature sequence allocation. The algorithm performs a one-user-at-a-time water-filling, and converges to the optimum solution described above.

## II. PROBLEM DEFINITION

We consider a CDMA system with processing gain  $N$ , where all  $K$  users transmit to a single receiver. In the presence of fading and AWGN, the received signal vector is given by [13]

$$\mathbf{r} = \sum_{i=1}^K \sqrt{p_i h_i} b_i \mathbf{s}_i + \mathbf{n} \quad (1)$$

where  $\mathbf{s}_i = [s_{i1}, \dots, s_{iN}]^\top$ ,  $p_i$ ,  $h_i$ ,  $b_i$  are the unit energy signature sequence, transmit power, channel gain, and information symbol, respectively, of user  $i$ , and  $\mathbf{n}$  is a zero-mean Gaussian random vector with covariance  $\sigma^2 \mathbf{I}_N$ . The information symbol  $b_i$  is assumed to have unit energy, i.e.,  $E[b_i^2] = 1$ . We assume that the receiver and all of the transmitters have perfect knowledge of the channel states of all users represented as a vector  $\mathbf{h} = [h_1, \dots, h_K]^\top$ . We further assume that although the fading is slow enough to ensure constant channel gain in a symbol interval, it is fast enough so that within the transmission time of a block of symbols the long-term ergodic properties of the fading process can be observed [14].

For a given set of signature sequences and a fixed set of channel gains  $\mathbf{h}$ , the sum capacity  $C_{\text{sum}}(\mathbf{h}, \bar{\mathbf{p}}, \mathbf{S})$  is [10]

$$C_{\text{sum}}(\mathbf{h}, \bar{\mathbf{p}}, \mathbf{S}) = \frac{1}{2} \log \det \left[ \mathbf{I}_N + \sigma^{-2} \sum_{i=1}^K h_i \bar{p}_i \mathbf{s}_i \mathbf{s}_i^\top \right] \quad (2)$$

where  $\bar{p}_i$  is the average power of user  $i$ ,  $\bar{\mathbf{p}} = [\bar{p}_1, \dots, \bar{p}_K]$ , and  $\mathbf{S} = [\mathbf{s}_1, \dots, \mathbf{s}_K]$ . To maximize the above capacity for that particular  $\mathbf{h}$ , one can choose the signature sequences of the users for a given set of power constraints. An equivalent problem is solved in [7], in the no-fading case, i.e.,  $h_i = 1$ , for all  $i$ .

In the presence of fading, if the channel state is modeled as a random vector, the quantity  $C_{\text{sum}}(\mathbf{h}, \bar{\mathbf{p}}, \mathbf{S})$  is random as well, and the ergodic sum capacity is found as the expected value of  $C_{\text{sum}}(\mathbf{h}, \bar{\mathbf{p}}, \mathbf{S})$ . Instead of keeping the transmit power of user  $i$  fixed to  $\bar{p}_i$  as in (2), we can choose the transmit powers of the users  $p_i(\mathbf{h})$ ,  $i = 1, \dots, K$ , as a function of the channel state with the aim of maximizing the ergodic sum capacity of the system subject to average transmit power constraints for all users. Similarly, we can choose the signature sequences  $\mathbf{S}$  to

be a function of the channel state as well; let us denote it by  $\mathbf{S}(\mathbf{h})$  to show the dependence on the channel state. Therefore, our problem is to solve for the jointly optimum transmit powers and signature sequences as a function of the channel state in order to maximize the ergodic sum capacity of the system in the presence of fading. The problem can be stated as

$$\begin{aligned} \max_{\mathbf{p}(\mathbf{h}), \mathbf{S}(\mathbf{h})} \int \log \det \left[ \mathbf{I}_N + \sum_{i=1}^K \frac{h_i p_i(\mathbf{h})}{\sigma^2} \mathbf{s}_i(\mathbf{h}) \mathbf{s}_i(\mathbf{h})^\top \right] f(\mathbf{h}) d\mathbf{h} \\ \text{s.t.} \int p_i(\mathbf{h}) f(\mathbf{h}) d\mathbf{h} = \bar{p}_i, \quad p_i(\mathbf{h}) \geq 0 \end{aligned} \quad (3)$$

where  $f(\mathbf{h})$  denotes the probability density function of the channel state vector.

## III. JOINT SIGNATURE SEQUENCE AND POWER ALLOCATION

In order to jointly optimize the powers and signature sequences, we first fix power distributions of all users over all fading states. Then, the corresponding optimal signature sequence set at every channel state will consist of a combination of orthogonal and GWBE sequences [7]. This is due to the fact that the signature sequences at a fading state  $\mathbf{h}$  can be chosen independently of the signature sequences at any other state, since once the powers are fixed, there are no constraints relating  $\mathbf{S}(\mathbf{h})$  to  $\mathbf{S}(\bar{\mathbf{h}})$  for  $\mathbf{h} \neq \bar{\mathbf{h}}$ . Since the optimum signature sequences at each channel state depend only on powers  $\mathbf{p}(\mathbf{h})$  and the channel state  $\mathbf{h}$ , we can express the capacity at each channel state only as a function of the powers, and optimize the ergodic capacity in terms of the power allocation. Let us define the *signature-sequence-optimized-sum-capacity* at channel state  $\mathbf{h}$  for a given power control policy  $\mathbf{p}(\mathbf{h})$  by

$$C_{\text{opt}}(\mathbf{h}, \mathbf{p}(\mathbf{h})) \triangleq \max_{\mathbf{S}(\mathbf{h})} C_{\text{sum}}(\mathbf{h}, \mathbf{p}(\mathbf{h}), \mathbf{S}(\mathbf{h})) \quad (4)$$

where  $C_{\text{sum}}(\mathbf{h}, \mathbf{p}(\mathbf{h}), \mathbf{S}(\mathbf{h}))$  is the argument of the expectation in the objective function of (3), i.e., it is the function in (2) where  $\bar{\mathbf{p}}$  is replaced by  $\mathbf{p}(\mathbf{h})$  and  $\mathbf{S}$  is replaced by  $\mathbf{S}(\mathbf{h})$ . For a fixed  $\mathbf{h}$ , it can be shown using majorization theory that  $C_{\text{opt}}(\mathbf{h}, \mathbf{p}(\mathbf{h}))$  is a concave function of the power vector at channel state  $\mathbf{h}$ ,  $\mathbf{p}(\mathbf{h})$  [15, Proposition 2.2]. Then, the problem in (3) can be written only in terms of the powers as

$$\begin{aligned} \max_{\mathbf{p}(\mathbf{h})} \int C_{\text{opt}}(\mathbf{h}, \mathbf{p}(\mathbf{h})) f(\mathbf{h}) d\mathbf{h} \\ \text{s.t.} \int p_i(\mathbf{h}) f(\mathbf{h}) d\mathbf{h} = \bar{p}_i, \quad p_i(\mathbf{h}) \geq 0. \end{aligned} \quad (5)$$

First consider the case when  $K \leq N$ . For any fixed channel state, the optimal choice of signature sequences for a given power control policy  $\mathbf{p}(\mathbf{h})$  is an orthogonal set [6], [7]. Noting that the received power levels are  $p_i(\mathbf{h}) h_i$ , (5) is equivalent to solving  $K$  independent Goldsmith–Varaiya problems [3] (see also [2]), the solution to which is a single user water-filling for each user. More precisely, the optimal solution  $\mathbf{p}^*(\mathbf{h})$  is the unique solution satisfying the Karush–Kuhn–Tucker (KKT) conditions, and is given by

$$p_i^*(\mathbf{h}) = \left( \frac{1}{\lambda_i} - \frac{\sigma^2}{h_i} \right)^+, \quad i = 1, \dots, K \quad (6)$$

where  $\lambda_i$  is solved by plugging (6) into (5).

One remarkable observation is that in obtaining  $C_{\text{opt}}(\mathbf{h}, \mathbf{p}(\mathbf{h}))$ , it is possible to adopt a channel nonadaptive signature sequence allocation policy, i.e., each user can be assigned a designated signature sequence, which it can use at all channel states, as long as the signature sequences in this set are orthogonal. A channel adaptive scheme will also perform

equally well as long as the signature sequences we choose at each  $\mathbf{h}$  are from an orthogonal set.

When  $K > N$ , it has been shown in [7], for a nonfading channel, that given the power constraints of all users, one can group the users into two sets  $L$  and  $\bar{L}$ , of oversized and nonoversized users, respectively. Users  $i \in L$  are assigned orthogonal sequences, and users  $i \in \bar{L}$  are assigned GWBE sequences. For a channel with fading, at a certain channel state  $\mathbf{h}$ , and for a certain arbitrary power distribution of users which assigns powers  $p_1, \dots, p_K$  to channel state  $\mathbf{h}$ , let us define the matrix  $\mathbf{D} \triangleq \text{diag}(p_1 h_1, \dots, p_K h_K)$ , and define  $\mu_i$  to be the eigenvalues of the matrix  $\mathbf{S} \mathbf{D} \mathbf{S}^\top$ . Then the signature sequences that maximize the sum capacity for any fixed  $\mathbf{h}$  satisfy [16],

$$\mathbf{S} \mathbf{D} \mathbf{S}^\top \mathbf{s}_i = \mu_i \mathbf{s}_i, \quad i = 1, \dots, K \quad (7)$$

clearly with repetitions of some of the  $\mu_i$ 's (since there are only  $N$  eigenvalues of  $\mathbf{S} \mathbf{D} \mathbf{S}^\top$ ), where the optimal  $\mu_i$ 's are given by [7]

$$\mu_i(\mathbf{h}) = \begin{cases} \frac{\sum_{j \in \bar{L}(\mathbf{h})} p_j h_j}{N - |\bar{L}(\mathbf{h})|}, & i \in \bar{L}(\mathbf{h}) \\ p_i h_i, & i \in L(\mathbf{h}) \end{cases} \quad (8)$$

In the fading case with channel adaptive powers, as suggested by the results in [2]–[4], it is likely that some users will have powers equal to zero at some channel states, and they will not contribute to  $C_{\text{sum}}$  at those channel states. Although the concept of oversized users is defined for users with nonzero average power constraints, since users which are allocated zero power at state  $\mathbf{h}$  will not contribute to the sum capacity, we can add them to the set of nonoversized users at channel state  $\mathbf{h}$ ,  $\bar{L}(\mathbf{h})$ , and we can assume that we assign arbitrary sequences for those users without changing the solution. Note however that, while finding the set of oversized users, we will disregard the users with zero power. Using the optimum eigenvalue assignment in (8) at each state, the objective function of the problem (5) can be expressed in the alternative form

$$\int \left[ \sum_{i \in L(\mathbf{h})} \log \left( 1 + \frac{p_i(\mathbf{h}) h_i}{\sigma^2} \right) + (N - |\bar{L}(\mathbf{h})|) \log \left( 1 + \frac{\sum_{i \in \bar{L}(\mathbf{h})} p_i(\mathbf{h}) h_i}{\sigma^2 (N - |\bar{L}(\mathbf{h})|)} \right) \right] f(\mathbf{h}) d\mathbf{h}. \quad (9)$$

For a given channel state  $\mathbf{h}$ , let the set of users that will transmit with nonzero powers be  $\bar{K}(\mathbf{h})$ . Then the number of users in  $\bar{K}(\mathbf{h})$  cannot exceed  $N$ , as stated by the following theorem.

**Theorem 1:** Let  $\bar{K}(\mathbf{h})$  be a subset of  $\{1, \dots, K\}$ , such that  $\forall i \in \bar{K}(\mathbf{h}), p_i^*(\mathbf{h}) > 0$ , where  $\mathbf{p}^*(\mathbf{h})$  is the maximizer of (9). Then,  $|\bar{K}(\mathbf{h})| \leq N$ , almost surely.

*Proof:* By concavity of  $C_{\text{opt}}(\mathbf{h}, \mathbf{p}(\mathbf{h}))$ , it is clear that the function in (9) is concave, and the maximization in (5) is over an affine set of constraints. Therefore, a power vector  $\mathbf{p}^*(\mathbf{h})$  achieves the global optimum of the maximization problem if and only if it satisfies the KKT conditions. Then, writing the KKT conditions for the objective function in (9), it is easy to show that

$$\frac{h_i}{\mu_i(\mathbf{h}) + \sigma^2} \leq \lambda_i, \quad \forall \mathbf{h} \quad (10)$$

where  $\mu_i(\mathbf{h})$  is given by (8), and equality holds if  $p_i(\mathbf{h}) > 0$ . Now, let us assume that the number of nonzero components in  $\mathbf{p}^*(\mathbf{h})$  is  $|\bar{K}(\mathbf{h})| > N$ , for a given  $\mathbf{h}$ . Then, some users must share some of the available dimensions, i.e., not all users can be made orthogonal to each other. In fact, we can find at most  $N - 1$  sequences that are

orthogonal to all other sequences in the system, or equivalently, at least  $|\bar{K}(\mathbf{h})| - N + 1$  users will have the same

$$\mu_i(\mathbf{h}) = \sum_{j \in \bar{L}(\mathbf{h})} h_j p_j / (N - |\bar{L}(\mathbf{h})|).$$

Then, substituting this into (10), we get  $h_i/\lambda_i = h_j/\lambda_j$  for  $i \neq j$ ,  $i, j \in \bar{K}(\mathbf{h})$  for at least  $|\bar{K}(\mathbf{h})| - N + 1$  users. Note that as the channel fading is assumed to be a continuous random variable, this event has zero probability, and at most one user with GWBE sequences (one with highest  $h_i/\lambda_i$  ratio, as in [4]) may transmit, with probability 1. But this contradicts the assumption that  $|\bar{K}(\mathbf{h})| > N$ , which establishes our main result, i.e.,  $|\bar{K}(\mathbf{h})| \leq N$  almost surely.  $\square$

This result may be viewed as a generalization of [4] to a vector channel with a unit rank constraint on the covariance matrices of the inputs; [4] showed that in scalar MAC (i.e., when  $N = 1$ ), at most one user may transmit at a channel state with probability 1. An important implication of Theorem 1 is that, since the optimal power allocation dictates that at most  $N$  users transmit with positive powers at any given channel state, orthogonal sequences should be assigned to those users that are transmitting with positive powers. That is, although we allowed for allocating GWBE sequences to some of the users, the solution implies that there is at most one such user, and the problem reduces to the orthogonal case. The optimal power allocation is again single-user water-filling, similar to the solution given in (6), i.e.,

$$p_i^*(\mathbf{h}) = \begin{cases} \left( \frac{1}{\lambda_i} - \frac{\sigma^2}{h_i} \right), & i \in \bar{K}(\mathbf{h}) \\ 0, & \text{otherwise.} \end{cases} \quad (11)$$

Here, one needs to be careful about the transmit regions. Unlike the case where the actual number of users  $K \leq N$ , the users in the set  $\bar{K}(\mathbf{h})$  change with  $\mathbf{h}$ , thus, a channel adaptive allocation of the orthogonal sequences is necessary. Our convention is that we assign a sequence from an orthogonal set to a user wherever its power is positive.

To specify the optimal power allocation completely, let us define  $\gamma_i = h_i/\lambda_i$ . Then, the probability that  $\gamma_i = \gamma_j$ , for  $i \neq j$  is zero. Therefore, we can always find a unique order statistics  $\{\gamma_{[i]}\}_{i=1}^K$  such that  $\gamma_{[1]} > \dots > \gamma_{[K]}$ , for each given  $\mathbf{h}$ . Let us now place  $\sigma^2$  in that ordering, assuming that at least one of the  $\gamma_{[i]}$ 's is larger than  $\sigma^2$ . Define  $\gamma_{[K+1]} = 0$ . Then, for some  $n \in \{1, \dots, K\}$ , let

$$\gamma_{[1]} \geq \dots \geq \gamma_{[n]} > \sigma^2 \geq \gamma_{[n+1]} \geq \dots \geq \gamma_{[K+1]} \quad (12)$$

where the equalities are included for the sake of consistency of the indices, and do not affect the solution (note the strict inequality just before  $\sigma^2$ ).

First, let  $n \leq N$ . Then, we see that (11) gives positive powers for all  $n$  users, and thus all  $n$  users with highest  $\gamma_i$ 's will transmit with the nonzero powers given in (11). When  $n > N$ , there are more than  $N$  users satisfying the positivity constraints  $\gamma_i > \sigma^2$ . However, we know from our derivation that only the user with the highest  $\gamma_i$  from the set we intend to assign GWBE sequences may transmit. Therefore, a total of  $N$  users with the highest  $\gamma_i$ 's transmit at this channel state.

Finally, we can summarize the jointly optimal power and signature sequence allocation policy as

$$\begin{aligned} p_i^*(\mathbf{h}) &= \begin{cases} \left( \frac{1}{\lambda_i} - \frac{\sigma^2}{h_i} \right), & \text{iff } i \in \Omega \\ 0, & \text{otherwise} \end{cases} \\ \mathbf{s}_i^*(\mathbf{h})^\top \mathbf{s}_j^*(\mathbf{h}) &= 0, \quad i \neq j, \quad \forall i, j \in \Omega \\ \Omega &= \{i : \gamma_{[i]} > \sigma^2, \quad i \leq \min\{K, N\}\}. \end{aligned} \quad (13)$$

#### IV. ITERATIVE POWER AND SEQUENCE OPTIMIZATION

We found in the preceding section that the optimal power control strategy is a water-filling over some favorable channel states for each user. However, in order to obtain the optimal power levels one should also compute the Lagrange multipliers  $\lambda_i$ , from the average power constraints. It turns out that the power allocation of each user still depends in a complicated fashion to those of the other users through these  $\lambda_i$ . In this section, we provide an iterative method to obtain the jointly optimal power and signature sequence allocation, and hence the  $\lambda_i$ .

In [2], we have shown that for fixed signature sequences  $\mathbf{S}$ , the optimal single-user update that maximizes the sum capacity as a function of  $p_k(\mathbf{h})$  is given by

$$p_k(\mathbf{h}, \mathbf{S}) = \left( \frac{1}{\lambda_k} - \frac{1}{h_k \mathbf{s}_k^\top \mathbf{A}_k^{-1} \mathbf{s}_k} \right)^+ \quad (14)$$

where the interference covariance matrix  $\mathbf{A}_k$  is defined as

$$\begin{aligned} \mathbf{A}_k &= \sigma^2 \mathbf{I}_N + \sum_{i \neq k} h_i p_i(\mathbf{h}) \mathbf{s}_i \mathbf{s}_i^\top \\ &= \sigma^2 \mathbf{I} + \mathbf{S} \mathbf{D} \mathbf{S}^\top - h_k p_k(\mathbf{h}) \mathbf{s}_k \mathbf{s}_k^\top. \end{aligned} \quad (15)$$

We can find and fix the optimal signature sequences at each state for a given power allocation using results of [7]. Then, plugging these sequences in (15), multiplying both sides by the optimal signature sequence  $\mathbf{s}_k^*$ , and noting that the signature sequences that maximize the sum capacity for a fixed set of power constraints satisfy (7), we get

$$\mathbf{A}_k \mathbf{s}_k^* = (\sigma^2 + \mu_k - h_k p_k) \mathbf{s}_k^* \quad (16)$$

where  $\mu_k$  are given by (8). Therefore,

$$\mathbf{s}_k^{*\top} \mathbf{A}_k^{-1} \mathbf{s}_k^* = \frac{1}{\sigma^2 + \mu_k - h_k p_k}. \quad (17)$$

This shows that we can represent the base level for the water-filling in (14) as a function of the power levels in the previous iteration. Substituting this in (14), we get the optimal power allocation at the  $n+1$ th step,  $p_k^{n+1}(\mathbf{h})$  for user  $k$ , with optimal sequences and fixed powers  $\{p_i(\mathbf{h})\}_{i \neq k}$  from the previous iteration

$$p_k^{n+1}(\mathbf{h}) = \left( \frac{1}{\lambda_k^{n+1}} - \frac{\sigma^2 + \mu_k^n(\mathbf{h}) - h_k p_k^n(\mathbf{h})}{h_k} \right)^+, \quad \forall \mathbf{h} \quad (18)$$

where we use  $\{p_1^{n+1}(\mathbf{h}), \dots, p_{k-1}^{n+1}(\mathbf{h}), p_k^n(\mathbf{h}), \dots, p_K^n(\mathbf{h})\}$  to compute  $\mu_k^n(\mathbf{h})$ . Combining this with (8) gives us the power update at each step. It is easy to observe that, once the eigenvalues  $\mu_k^n(\mathbf{h})$  are determined using the power levels from the previous iteration, we can use (18) to solve for  $k$ th user's power by water-filling. Note that, the Lagrange multiplier  $\lambda_k^{n+1}$  is chosen to satisfy the average power constraint of user  $k$  at each iteration, and can be obtained by plugging (18) into the constraint in (3). The water-filling algorithm automatically obtains the value of  $\lambda_k^{n+1}$  as it is the inverse of the "water level."

The proposed algorithm may be interpreted in two ways. First, it may be seen as an iteration from a set of powers to another set of powers as given by (18). Therefore, one may run this algorithm starting with an arbitrary power distribution, to obtain the capacity-maximizing power distribution when the algorithm converges. The signature sequences may then be assigned to the users after the algorithm converges: at each channel state, the users that have nonzero powers (there will be at most  $N$  such users) are assigned signature sequences from an orthogonal set. Second, the algorithm may be seen as an iteration from powers to

signature sequences, and then back to powers again. Specifically, for a given set of powers, the optimal sequences may be found using (7) and (8), i.e., as in [7]; corresponding to these sequences, base levels for the water-filling in (14) can be computed using (16) and (17), and new powers may be found using (14) as in [2].

We will now show that (18) and, equivalently, the sequential signature sequence and power update algorithm indeed converges to the global optimum of the sum capacity function. To see this, first observe that for fixed signature sequences, the update (14) is the best one-user-at-a-time power update and is guaranteed to give a nondecreasing sequence of sum capacity values [2]. Similarly, for fixed powers, the signature sequence update will increase (or keep constant) the value of the sum capacity. The sum capacity is upper-bounded, therefore, it is guaranteed that the sequence of nondecreasing sum capacity values obtained through these iterations have a limit. Moreover, the algorithm terminates if and only if the update (18) yields a fixed point  $\mathbf{p}(\mathbf{h})$ . Since the fixed point is characterized by  $\mathbf{p}^{n+1} = \mathbf{p}^n$ , it is easy to see that the fixed point of the update (18) actually satisfies the KKT conditions for our original problem. Since the convergence point  $\mathbf{p}(\mathbf{h})$  satisfies the KKT conditions, it achieves the global optimum of the sum capacity, proving the convergence of the sequential algorithm.

Note that we have incorporated the eigenvalues of  $\mathbf{S} \mathbf{D} \mathbf{S}^\top$  in the power iteration (18) rather than including the signature sequences explicitly. This implementation is very useful, since it does not require us to compute the signature sequences at intermediate steps. Finally, it is useful to point out that, although the power allocation policy that maximizes the sum capacity is unique, the signature sequence selection that is jointly optimal with this power allocation is not, for two reasons: first, because of the arbitrariness of the optimal sequences for users with zero powers; and second, because of the fact that even for users with nonzero powers, there are infinitely many sets of orthogonal sequences.

#### V. NUMERICAL EXAMPLES

First, we simulate a system where the number of users is equal to the processing gain:  $K = N = 3$ . In all of our simulations, we pick  $\sigma^2 = 1$ , the average power of each user to be 1, the initial power distribution uniformly, and the probability distribution of the channel to be uniform on the intervals shown in the figures. In this case, by our arguments in Section III, we expect the optimal signature sequences to be three orthogonal sequences. Fig. 1 shows the convergence of our algorithm, together with the convergence of the iterative water-filling algorithm in [2] for fixed sequences. When we optimize the powers and signature sequences jointly, we see that the sum capacity achieved is identical to that of a system with fixed orthogonal sequences, meaning channel adaptive and nonadaptive sequence selections give us the same capacity value. The power allocation strategy corresponding to the orthogonal signature sequences found by the algorithm is independent one-user-at-a-time water-filling for each user. Fig. 1 shows the sum capacity versus per user iterations, where one full cycle of the algorithm is equivalent to  $K = 3$  iterations. In this case, while the algorithm stops after  $18/3 = 6$  cycles for a threshold value of  $10^{-7}$  on the mean squared difference in power, i.e.,  $E[(p_i^{n+1}(\mathbf{h}) - p_i^n(\mathbf{h}))^2] < 10^{-7}$  for all users  $i$ , it converges to the optimum sum capacity value in practically one cycle of iterations (one iteration for each user). The capacity achieved by a randomly generated signature sequence matrix  $\mathbf{S}$  containing unit-norm sequences is also given for comparison; as expected,

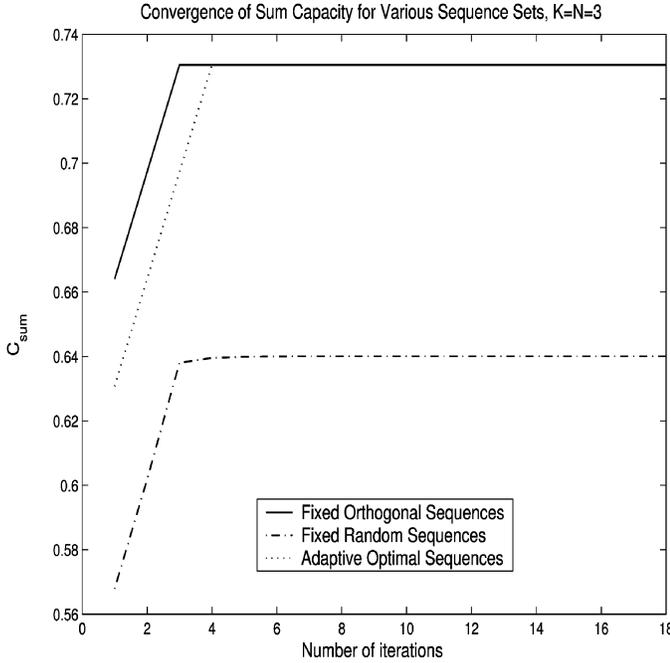


Fig. 1. Convergence of sum capacity for  $K = N = 3$ .

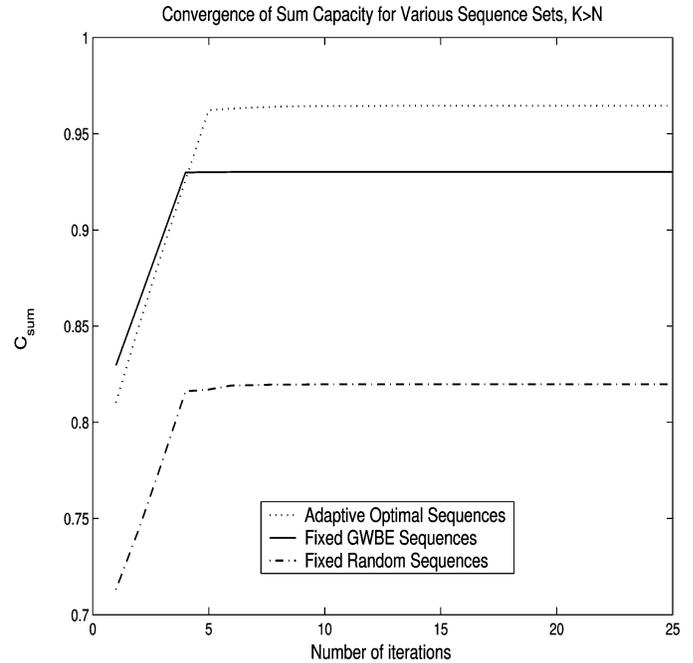


Fig. 2. Convergence of sum capacity for  $K = 4, N = 3$ .

the sum capacity for that matrix  $\mathbf{S}$  is inferior to the orthogonal sequences case.

The convergence plots for a more interesting case where  $K = 4, N = 3$  are given in Fig. 2. Here, we again compare the capacity achieved by our algorithm to some fixed random sequences, and we see that we get a higher capacity. We also compare our result to a fixed set of WBE sequences, which are the optimum sequences for a fixed channel state and equal average received powers. The iterative water-filling with sequence optimization again achieves a better sum capacity. The algorithm again stops after  $25/4 \cong 6$  cycles, converging to sum capacity in practically one cycle of iterations. Also remarkably, the transmit strategy is such that at most three of the four users transmit together (on a region with nonzero probability, after eliminating the states where the channel states of any two users are equal), and they are allocated orthogonal sequences. Fig. 3(a)–(d) further illustrates the details of the power and signature sequence allocation.

Fig. 3(a) and (b) pertains to a plane in four-dimensional channel state space, where we pick  $h_3 = h_4 = 0.4$ , and observe the power distribution of users 1 and 2 as a function of their channel states. The gray levels correspond to the amount of power allocated, lighter colors indicating more power. Clearly, the users perform single-user water-filling for the chosen channel states, and their powers do not depend on fading states and powers of each other. As  $h_3 = h_4 = 0.4$ , from (13) we expect that users 1 and 2 would transmit when their channels are better than 0.4, with orthogonal sequences, and hence the single-user water-filling. Note that, according to the notion in [7], users 1 and 2 are oversized whenever their channel gains are better than 0.4.

Fig. 3(c) and (d) corresponds to a case where we pick the maximum possible values for the channel states  $h_3$  and  $h_4$ , i.e.,  $h_3 = h_4 = 0.9$ , so that except for the degenerate equality cases, users 3 and 4 will always be oversized on the plane of channel states we consider. Then, the remaining user, according to our results, should transmit if and only if it has the next best channel (note that since channels are all taken to be identically distributed, the  $\lambda_i$ 's are the same for all users and they do not affect the ordering). This is what is observed in Fig. 3(c) and (d),

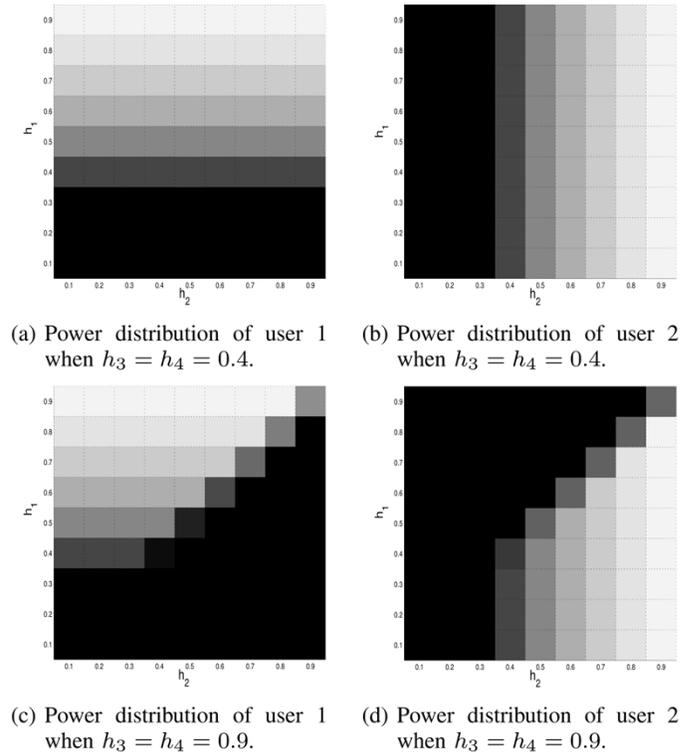


Fig. 3. Cross sections of power distributions for users 1 and 2. (a) Power distribution of user 1 when  $h_3 = h_4 = 0.4$ . (b) Power distribution of user 2 when  $h_3 = h_4 = 0.4$ . (c) Power distribution of user 1 when  $h_3 = h_4 = 0.9$ . (d) Power distribution of user 2 when  $h_3 = h_4 = 0.9$ .

the stronger of users 1 and 2 perform single-user water-filling, and the weaker one remains silent, as in the Knopp–Humblet [4] solution. The *arbitrariness* in powers in the equal channels case is again observed, and is consistent with our previous arguments.

## VI. CONCLUSION

For a CDMA system subject to fading, we showed that the ergodic sum capacity is maximized by allocating orthogonal signature sequences to  $\min(N, K)$  of the users with *favorable* channel states, and allocating powers to those users by a single-user water-filling strategy over some partitions of channel state space. In each partition, a group of users perform orthogonal transmissions, thus, the users avoid any interference from each other in order to maximize the sum capacity. We also proposed an iterative signature-update/power-water-filling algorithm to find the optimal allocation of signature sequences and powers numerically, and proved its convergence to the globally optimum solution.

## REFERENCES

- [1] O. Kaya and S. Ulukus, "Jointly optimal power and signature sequence allocation for fading CDMA," in *Proc. IEEE Global Communications Conf.*, San Francisco, CA, Dec. 2003, pp. 1872–1876.
- [2] —, "Optimum power control for CDMA with deterministic sequences in fading channels," *IEEE Trans. Inf. Theory*, vol. 50, no. 10, pp. 2449–2458, Oct. 2004.
- [3] A. J. Goldsmith and P. P. Varaiya, "Capacity of fading channels with channel side information," *IEEE Trans. Inf. Theory*, vol. 43, no. 6, pp. 1986–1992, Nov. 1997.
- [4] R. Knopp and P. A. Humblet, "Information capacity and power control in single-cell multiuser communications," in *Proc. IEEE Int. Conf. Communications*, Seattle, WA, Jun. 1995, pp. 331–335.
- [5] R. D. Yates, "A framework for uplink power control in cellular radio systems," *IEEE J. Select. Areas Commun.*, vol. 13, no. 7, pp. 1341–1347, Sep. 1995.
- [6] M. Rupp and J. L. Massey, "Optimum sequence multisets for synchronous code-division multiple-access channels," *IEEE Trans. Inf. Theory*, vol. 40, no. 4, pp. 1261–1266, Jul. 1994.
- [7] P. Viswanath and V. Anantharam, "Optimal sequences and sum capacity of synchronous CDMA systems," *IEEE Trans. Inf. Theory*, vol. 45, no. 7, pp. 1984–1991, Nov. 1999.
- [8] P. Viswanath, V. Anantharam, and D. N. C. Tse, "Optimal sequences, power control and capacity of synchronous CDMA systems with linear MMSE multiuser receivers," *IEEE Trans. Inf. Theory*, vol. 45, no. 5, pp. 1968–1983, Sep. 1999.
- [9] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. New York: Wiley Interscience, 1991.
- [10] S. Verdú, "Capacity region of Gaussian CDMA channels: The symbol-synchronous case," in *Proc. Allerton Conf. Communications, Control and Computing*, Monticello, IL, Oct. 1986, pp. 1025–1034.
- [11] P. Viswanath, D. N. C. Tse, and V. Anantharam, "Asymptotically optimal waterfilling in vector multiple access channels," *IEEE Trans. Inf. Theory*, vol. 47, no. 1, pp. 241–267, Jan. 2001.
- [12] L. R. Welch, "Lower bounds on the maximum cross-correlation of signals," *IEEE Trans. Inf. Theory*, vol. IT-20, no. 3, pp. 397–399, May 1974.
- [13] S. Verdú, *Multiuser Detection*. Cambridge, U.K.: Cambridge Univ. Press, 1998.
- [14] E. Biglieri, J. Proakis, and S. Shamai (Shitz), "Fading channels: Information-theoretic and communications aspects," *IEEE Trans. Inf. Theory*, vol. 44, no. 6, pp. 2619–2692, Oct. 1998.
- [15] P. Viswanath and V. Anantharam, "Optimal sequences for CDMA under colored noise: A Schur–saddle function property," *IEEE Trans. Inf. Theory*, vol. 48, no. 6, pp. 1295–1318, Jun. 2002.
- [16] S. Ulukus and R. D. Yates, "Iterative construction of optimum signature sequence sets in synchronous CDMA systems," *IEEE Trans. Inf. Theory*, vol. 47, no. 5, pp. 1989–1998, Jul. 2001.

## A Distribution Dependent Refinement of Pinsker's Inequality

Erik Ordentlich, *Member, IEEE*, and  
Marcelo J. Weinberger, *Senior Member, IEEE*

**Abstract**—Given two probability distributions  $Q$  and  $P$ , let  $\|Q - P\|_1$  and  $D(Q||P)$ , respectively, denote the  $L_1$  distance and divergence between  $Q$  and  $P$ . We derive a refinement of Pinsker's inequality of the form  $D(Q||P) \geq c(P)\|Q - P\|_1^2$  and characterize the best  $P$ -dependent factor  $c(P)$ . We apply the refined inequality to large deviations and measure concentration.

**Index Terms**—Divergence, Hoeffding's inequality,  $L_1$  distance, measure concentration, Pinsker's inequality, Sanov's theorem.

## I. PRELIMINARIES

Let  $\mathcal{A}$  denote the finite set  $\{1, 2, \dots, a\}$ . For two probability distributions  $Q$  and  $P$  on  $\mathcal{A}$ , let

$$\|Q - P\|_1 = \sum_{k=1}^a |Q(k) - P(k)|$$

denote the variational, or  $L_1$ , distance between  $Q$  and  $P$ , and let

$$D(Q||P) = \sum_{k=1}^a Q(k) \log \frac{Q(k)}{P(k)} \quad (1)$$

denote the divergence between  $Q$  and  $P$ , where throughout  $\log(\cdot)$  denotes the natural logarithm.

For  $0 \leq p_1, p_2 \leq 1$ , let

$$d(p_1||p_2) = p_1 \log \frac{p_1}{p_2} + (1 - p_1) \log \frac{1 - p_1}{1 - p_2} \quad (2)$$

denote the binary divergence, whereas for  $(p_1, p_2) \notin [0, 1]^2$  we set  $d(p_1||p_2) = \infty$ .

The following conventions implied by continuity are adopted: for  $c > 0$ ,  $c/0 = \infty$ ,  $c/\infty = 0$ ,  $c\infty = \infty$ ,  $\log \infty = \infty$ ,  $e^{-\infty} = 0$ . Additionally, in (1) and (2), it is assumed that  $0 \log(0/0) = 0$  and  $0 \log 0 = 0$ .

For  $p \in [0, 1/2)$ , we define [4]

$$\varphi(p) = \frac{1}{1 - 2p} \log \frac{1 - p}{p} \quad (3)$$

and, by continuity, set  $\varphi(1/2) = 2$ .

For a probability distribution  $P$  on  $\mathcal{A}$ , we define

$$\pi_P = \max_{A \subseteq \mathcal{A}} \min\{P(A), 1 - P(A)\}. \quad (4)$$

Note that  $\pi_P \leq 1/2$  for any  $P$ .

Finally, throughout we take the minimum of a function over an empty set to be  $\infty$ .

## II. REFINED PINSKER'S INEQUALITY

We refine Pinsker's inequality ([2, Problem 3.17], [1, Lemma 12.6.1]) relating the  $L_1$  distance to the divergence as follows.

Manuscript received October 24, 2003; revised December 22, 2004.

The authors are with the Hewlett-Packard Laboratories, Palo Alto, CA 94304 USA (e-mail: eord@hpl.hp.com; marcelo@hpl.hp.com).

Communicated by A. B. Nobel, Associate Editor for Pattern Recognition, Statistical Learning and Inference.

Digital Object Identifier 10.1109/TIT.2005.846407