Multiaccess Fading Channels–Part II: Delay-Limited Capacities

Stephen V. Hanly, *Member, IEEE*, and David N. C. Tse, *Member, IEEE*

***Abstract*—In multiaccess wireless systems, dynamic allocation of resources such as transmit power, bandwidths, and rates is an important means to deal with the time-varying nature of the environment. In this two-part paper, we consider the problem of optimal resource allocation from an information-theoretic point of view. We focus on the multiaccess fading channel with Gaussian noise, and define two notions of capacity depending on whether the traffic is delay-sensitive or not. In Part I, we have analyzed the *throughput capacity region* which characterizes the long-term achievable rates through the time-varying channel. However, the *delay* experienced depends on how fast the channel varies. In the present paper, Part II, we introduce a notion of *delay-limited capacity* which is the maximum rate achievable with delay *independent* of how slow the fading is. We characterize the *delay-limited capacity region* of the multiaccess fading channel and the associated optimal resource allocation schemes. We show that successive decoding is optimal, and the optimal decoding order and power allocation can be found explicitly as a function of the fading states; this is a consequence of an underlying polymatroid structure that we exploit.**

***Index Terms*—Delay-limited capacity, fading channels, multiaccess, power control, successive cancellation.**

I. Introduction

In Part I [15] of this paper, we studied the problem of optimal dynamic resource allocation for multiaccess fading channels from an information-theoretic point of view. We computed the Shannon capacity region of the multiaccess fading channel when the transmitters as well as the receiver have access to the channel state, and also characterized the optimal power- and rate-allocation schemes.

The Shannon capacity of a channel provides the ultimate limits on the rates that are achievable. The capacity itself is not dependent on any delay considerations, and is achievable in an asymptotic sense as delay tends to infinity. Thus when we focused on the Shannon capacity of the multiaccess fading channel in Part I, we found the Shannon limit over the set of all possible rate and power allocation strategies, with the objective of maximizing the long-term average rates. In Part II, we now consider limitations on rate and power allocation strategies that can be used due to delay constraints.

One way to think about this in the single-user case is to identify the Shannon capacity with a "long-term average of mutual information between the user and the receiver" in the channel. That is why we called the Shannon capacity of a fading channel its *throughput capacity* in Part I. On the other hand, there is a notion of "instantaneous mutual information" and this can fluctuate as a function of the fading state. Essentially, in the delay-limited case, we restrict ourselves to power control strategies such that the instantaneous mutual information is kept constant at all times. We call the maximum achievable rate the *delay-limited capacity* of the channel. Without such a restriction the throughput of the channel can be increased but at the expense of having the "instantaneous mutual information" fluctuating with the fading process, leading to delay at the time scale of the channel variations. While the single-user delay-limited power-control strategy is simply "channel inversion," the multiuser problem is more interesting as it involves tradeoffs between the powers allocated to each of the users to achieve desired rates.

There are many "delay-sensitive" applications such as voice and video, for which long delays cannot be tolerated. Unless the fading is fast on the time-scale of tolerable delay, the throughput capacity of Part I is not relevant for these applications. Our delay-limited capacity is the appropriate limit for these applications.

The notion of "delay limitedness" is implicit in many works. For example, papers on power control (see Gilhousen *et al.* [7], Hanly [9], Yates [16]) assume that a desired signal-to-interference ratio must be met for every fading state, and this means that the user's mutual information is kept constant in time. The formal notion of delay-limited capacity for multiaccess channels was defined in Hanly and Tse [10], where we considered the symmetric case with users having the same rate requirements. In the present paper, we focus on characterizing the entire delay-limited capacity region and the associated optimal power control schemes. As in Part I, we shall exploit the convex and polymatroid structure of this problem. Again, we find that the optimal solution is always successive decoding, and that the optimal power control can be explicitly characterized and has a greedy interpretation.

Part II is organized as follows. In Section II we introduce the Gaussian, multiaccess, flat fading model and present a coding theorem for the delay-limited capacity region when

Manuscript received December 19, 1996; revised May 3, 1998. This work was supported in part by the Air Force Office of Scientific Research under Grant F49620-96-1-0199. This work was also supported by the Australian Research Council large grant. Part of this work was performed when both authors were at the Mathematical Research Science Center, AT&T Bell Laboratories. The material in this paper was presented in part at the International Symposium on Information Theory, Cambridge, MA, August 1998.

S. V. Hanly is with the Department of of Electrical Engineering, University of Melbourne, Melbourne, Australia (e-mail: s.hanly@ee.mu. oz.au).

D. N. C. Tse is with the Department of Electrical Engineering and Computer Sciences, University of California, Berkeley, CA 94720 USA (e-mail: dtse@eecs.berkeley.edu).

Publisher Item Identifier S 0018-9448(98)07360-X.

transmitters and receiver can track the channel. This theorem implies that the extra benefit gained from the transmitters tracking the channel is fully realized in the ability to allocate transmit power based on the channel state. In Section III, we use Lagrangian techniques to show that the optimal power allocation can be obtained by solving a family of optimization problems over a set of parallel time-invariant multiaccess Gaussian channels, one for each fading state. Given the Lagrange multipliers, which can be interpreted as power prices, the problem is that of finding the optimal "power" allocation as a function of each fading state so as to minimize the total power cost. Here, we exploit the polymatroid structure of the optimization problem to obtain an explicit solution via a greedy algorithm. In Section IV, we turn to the problem of finding an appropriate set of power prices so that a target delay-limited rate vector can be met within given power constraints. We present an iterative algorithm which, if the target rates are achievable, is guaranteed to converge to the right power prices. Moreover, it also solves a *call admissions* problem by determining if a given set of target rates are indeed achievable.

In the remainder of the paper, we will extend the basic results in several directions. In Section V, we will present greedy power-allocation algorithms when additional power constraints are imposed. These results exploit further properties of polymatroids. In Section VI, we relax the delay limited requirement in two ways. First, we consider a multiple time-scale model, with slow and fast fading, and compute the optimal power control when we are delay-limited with respect to the slow fading. Secondly, we consider a frequency-selective fading channel, in which rates can be allocated to the different frequencies, but the sum rate over all frequencies must be constant for each fading state. Finally, in Section VII, we explore the implications of these information-theoretic results to systems with suboptimal coding and decoding.

In Hanly and Tse [10], the concept of delay-limited capacity is extended to take advantage of statistical multiplexing: it is not always necessary for power control to be used to ensure that "sufficient mutual information is available at every time instant"; this can also be a property of the averaging of the independent fading of a large number of users, even if no power control, or only decentralized power control is employed. In the present paper, however, we allow centralized power control and so do not consider statistical multiplexing of fading.

Our results also provide a link between information theory and the theory of networking. Clearly, the power prices (and in Part I, rate rewards) have the potential to be tuned by the network in order to provide control over the radio resources. This is indeed our approach in Section IV, in which a call admission problem is solved by the adaptation of power prices, using an algorithm reminiscent of max-min fair bandwidth allocation algorithms in data networks. In Part I, we employed similar iterative algorithms to control real-time radio resource allocation (see Section IV in Part I). More generally, there is an economic flavor to our results, as touched on in Section III, and more directly in Part I, Section V.

A word about notation: we will use boldface letters to denote vector quantities.

II. Delay-Limited Capacity

As in Part I, we focus on the uplink scenario where a set of M users communicate to a single receiver. Consider the discrete-time multiple-access Gaussian channel

$$Y(n) = \sum_{i=1}^{M} \sqrt{H_i(n)} X_i(n) + Z(n) \qquad (1)$$

where M is the number of users, $X_i(n)$ and $H_i(n)$ are the transmitted waveform and the fading process of the ith user, respectively, and $Z(n)$ is white Gaussian noise with variance σ^2. We assume that the fading processes for all users are jointly stationary and ergodic, and the stationary distribution has a continuous density and is bounded. User i is also subject to an average transmit power constraint of $\bar{P}_i$. We shall call $\boldsymbol{H}(n) = (H_1(n), H_2(n), \cdots, H_M(n))$ the joint fading process.

Suppose each source i codes over a blocklength of T symbols, where T is the delay, using a codebook $\mathcal{C}_i$ of size $2^{R_i T}$ (i.e., at rate R_i bits per channel use). Each codeword $\boldsymbol{x}$ of the ith user satisfies $\|\boldsymbol{x}\|_2^2 \leq T\bar{P}_i$. Fix a decoding scheme and assume the messages are chosen with equal probability. Let $p_e(T)$ be the probability of the event that any user is decoded incorrectly. The following is the definition of the throughput capacity region when both the transmitters and the receiver have access to the channel states. Characterizing this region was our focus in Part I.

Definition 2.1: The rate-tuple $\boldsymbol{R} = (R_1, \cdots, R_M)$ lies in the interior of the throughput capacity region $\mathcal{C}(\bar{\boldsymbol{P}})$ if and only if for every $\epsilon > 0$ there exists a delay T, codebooks, and a decoding scheme such that the probability of error $p_e(T)$ is less than ϵ. Moreover, the codewords can be chosen as a function of the realization of the fading processes.

The notion of throughput capacity defined above is a natural extension of that for time-invariant Gaussian channels, where rates are achieved with arbitrarily long coding delays. However, there is a subtle but important difference between time-varying and time-invariant Gaussian channels. In the time-invariant Gaussian channel, the delay is needed to average out the Gaussian noise to get small error probabilities, and this is typically quite short. Thus the capacity is not only an upper bound to the achievable performance; it is a useful upper bound in the sense that it is possible to achieve rates close to capacity with acceptable delay, even for real-time traffic. In typical time-varying wireless channels, on the other hand, the fading process is a complex superposition of different effects some of which can be quite slow. Thus the delay required to average out such fading effects may be much longer than the acceptable delay.

To this end, we define a second notion of capacity region for time-varying multiaccess channels. Let $\mathcal{H}$ be the set of all possible joint fading states of the users, $\mathcal{Q}$ be a given distribution on $\mathcal{H}$, and $\mathcal{A}(\mathcal{Q})$ be the set of all stationary, ergodic fading processes with stationary distribution $\mathcal{Q}$. We observe from Theorem 2.1 in Part I that the throughput capacity region of the multiaccess fading channel depends only on the stationary distribution of the joint fading processes and not on the correlation structure. The following definition of the *delay-limited* capacity region also has this characteristic.

Definition 2.2: A rate vector $(R_1, \cdots, R_M)$ lies in the interior of the *delay-limited* capacity region $\mathcal{C}_d(\bar{\boldsymbol{P}})$, if for every $\epsilon > 0$ there exists a coding delay T such that for every fading process in $\mathcal{A}(\mathcal{Q})$ there exists codebooks and a decoding scheme with $p_e(T) < \epsilon$. Moreover, the codewords can be chosen as a function of the realization of the fading processes.

Contrast this with Definition 2.1, where the coding delay can be chosen depending on the specific fading process, the coding delay here has to work *uniformly* for all fading processes with a given stationary distribution. Hence, rates in the delay-limited capacity region can be achieved with delays *independent* of the correlation structure of the fading. Thus the rates in the delay-limited capacity region are essentially those that can be achieved by coding that averages out the white noise but does not average over the fading process. It is an appropriate limit on the performance for traffic with stringent delay requirements and when the fading processes change relatively slowly (due to users at walking speed, for example). It should also be noted that the throughput capacity region contains the delay-limited capacity region.

In Definition 2.2, we only require that there be a codebook for every realization of every fading process. However, the proof of Theorem 2.3 below shows that we can provide a single codebook of unit power that we scale by the power control policy identified in the theorem. This codebook will work no matter what fading process is chosen (i.e., for any correlation structure). By "power control policy", we mean the following.

A *power control policy* $\mathcal{P} : \Re^M \rightarrow \Re^M$ is a mapping such that given a joint fading state $\boldsymbol{h} = (h_1, \cdots, h_M)$ for the users, $\mathcal{P}_i(\boldsymbol{h})$ can be interpreted as the transmitter power allocated to user i. Given power control policy $\mathcal{P}$, $\mathbb{E}_{\boldsymbol{H}}[\mathcal{P}_i(\boldsymbol{H})]$ is the average power usage for user i. We say a power control policy is *feasible* for a power constraint $\bar{\boldsymbol{P}}$ if $\mathbb{E}_{\boldsymbol{H}}[\mathcal{P}_i(\boldsymbol{H})] \leq \bar{P}_i$ for all i.

The following theorem provides a characterization of the delay-limited capacity region for the case when all the transmitters and the receiver know the current state of the channel.

Theorem 2.3: Assume that the set of possible fading states $\mathcal{H}$ is bounded. The delay-limited capacity region $\mathcal{C}_d(\bar{\boldsymbol{P}})$ is given by

$$\mathcal{C}_d(\bar{\boldsymbol{P}}) = \bigcup_{\mathcal{P} \in \mathcal{F}} \bigcap_{\boldsymbol{h} \in \mathcal{H}} \mathcal{C}_g(\boldsymbol{h}, \mathcal{P}(\boldsymbol{h})) \tag{2}$$

where $\mathcal{F}$ is the set of all feasible power control policies satisfying the average power constraints, and $\mathcal{C}_g(\boldsymbol{h}, \boldsymbol{P})$ is the capacity region of the time-invariant Gaussian multiaccess channel, given by[1]

$$\mathcal{C}_g(\boldsymbol{h}, \boldsymbol{P}) = \left\{ \boldsymbol{R} : \boldsymbol{R}(S) \leq \frac{1}{2} \log \left(1 + \frac{\sum_{i \in S} h_i P_i}{\sigma^2} \right) \right. \\ \left. \text{for every } S \subset \{1, \cdots, M\} \right\}. \tag{3}$$

Proof: See Appendix A. □

[1] Here, as in Part I, for any vector $\boldsymbol{v}$ and any subset S we use the notation $\boldsymbol{v}(S)$ to denote $\sum_{i \in S} v_i$.

The intuitive content of the above theorem is that a rate vector $(R_1, \cdots, R_M)$ is achievable in the delay-limited sense if one can choose a feasible power control policy to coordinate the powers of the users such that sufficient mutual information is maintained between the transmitters and the receiver at *all* fading states. Note that this is essentially the information-theoretic version of the objective of standard power control algorithms in which power is allocated to satisfy the signal-to-interference requirements of all the users. Contrast this with the characterization, in Theorem 2.1 of Part I, for the throughput capacity region, where a rate vector $(R_1, \cdots, R_M)$ is achievable as long as there is a feasible power control policy to provide sufficient *long-term* average mutual information, averaged over all fading states. The "instantaneous" mutual information at each fading state, however, fluctuates.

III. Characterization of the Delay-Limited Capacity Region

In this section, we will characterize the optimal power control to achieve points on the boundary of the delay-limited capacity region $\mathcal{C}_d(\bar{\boldsymbol{P}})$. We shall show that successive decoding is always optimal and we shall provide greedy algorithms for obtaining the optimal power control. Using this characterization, we will also provide a necessary and sufficient condition for $\boldsymbol{R}$ to be inside the capacity region.

A. Lagrangian Characterization and Optimal Power Allocation

We first define the boundary surface of $\mathcal{C}_d(\bar{\boldsymbol{P}})$, which is essentially the set of optimal operating points on the capacity region.

Definition 3.1: The boundary surface of $\mathcal{C}_d(\bar{\boldsymbol{P}})$ is the set of those rates such that no component can be increased with the other components kept fixed, while remaining in $\mathcal{C}_d(\bar{\boldsymbol{P}})$.

The following lemma gives a Lagrangian characterization of the capacity region.

Lemma 3.2:

1) A rate vector $\boldsymbol{R}^*$ lies in $\mathcal{C}_d(\bar{\boldsymbol{P}})$ if and only if there exists a $\boldsymbol{\lambda} \in \Re_+^M$ and a power control policy $\mathcal{P}$ such that for every joint fading state $\boldsymbol{h}$, $\mathcal{P}(\boldsymbol{h})$ is a solution to the optimization problem

$$\min_{\boldsymbol{p}} \boldsymbol{\lambda} \cdot \boldsymbol{p} \text{ subject to } \boldsymbol{R}^* \in \mathcal{C}_g(\boldsymbol{h}, \boldsymbol{p}) \tag{4}$$

and

$$\mathbb{E}_{\boldsymbol{H}}[\mathcal{P}_i(\boldsymbol{H})] \leq \bar{P}_i, \qquad i = 1, \cdots, M$$

where $\bar{P}_i$ is the constraint on the average power of user i. Moreover, $\mathcal{P}$ is a power control policy which can achieve the rate vector $\boldsymbol{R}^*$.

2) A rate vector $\boldsymbol{R}^*$ lies on the boundary surface if and only if there exist $\boldsymbol{\lambda}$ as above but with all the average power constraints holding with equality.

Analogous to Lemma 3.10 of Part I, this lemma reduces the computation of the optimal power control to a family of optimization problems over a set of parallel time-invariant

Gaussian channels. As in the analysis of the throughput capacity region, the vector $\boldsymbol{\lambda}$ can be interpreted as a set of *power prices* reflecting the power constraints. The important difference is that in this case, we require that the rate vector $\boldsymbol{R}^*$ be in the Gaussian capacity region $\mathcal{C}_g(\boldsymbol{h}, \mathcal{P}(\boldsymbol{h}))$ for *all* fading states $\boldsymbol{h}$. This is consistent with the nature of delay-limited capacities.

Proof: Since any rate vector inside the capacity region is dominated by some point on the boundary surface, statement 2) would imply statement 1). Hence, we will focus on proving statement 2).

First note that since the capacity region $\mathcal{C}_d(\bar{\boldsymbol{P}})$ is convex, a point $\boldsymbol{R}^*$ is on the boundary surface of the region if and only if it is a solution to the optimization problem

$$\max_{\boldsymbol{R}} \boldsymbol{\mu} \cdot \boldsymbol{R} \text{ subject to } \boldsymbol{R} \in \mathcal{C}_d(\bar{\boldsymbol{P}}) \tag{5}$$

for some positive vector $\boldsymbol{\mu}$. Now consider the set

$$S \equiv \{(\boldsymbol{R}, \boldsymbol{P}) : \boldsymbol{R} \in \mathcal{C}_d(\bar{\boldsymbol{P}})\}.$$

By the concavity of the $\log$ function, it can readily be verified that S is a convex set. Thus $\boldsymbol{R}^*$ solves (5) if and only if there exist nonnegative Lagrange multipliers $\boldsymbol{\lambda}$ such that $(\boldsymbol{R}^*, \bar{\boldsymbol{P}})$ is a solution to the problem

$$\max_{(\boldsymbol{R},\boldsymbol{P}) \in S} \boldsymbol{\mu} \cdot \boldsymbol{R} - \boldsymbol{\lambda} \cdot \boldsymbol{P}.$$

Hence, $\boldsymbol{R}^*$ is on the boundary surface of $\mathcal{C}_d(\bar{\boldsymbol{P}})$ if and only if $\bar{\boldsymbol{P}}$ is a solution to the problem

$$\min_{\boldsymbol{P}} \boldsymbol{\lambda} \cdot \boldsymbol{P} \text{ subject to } \boldsymbol{R}^* \in C_d(\boldsymbol{P})$$

i.e., if and only if there exists a power control policy $\mathcal{P}^*$ which solves

$$\min_{\mathcal{P}} \boldsymbol{\lambda} \cdot \mathbb{E}_{\boldsymbol{H}}[\mathcal{P}(\boldsymbol{H})] \text{ subject to } \boldsymbol{R}^* \in \cap_h C_g(\boldsymbol{h}, \mathcal{P}(\boldsymbol{h}))$$

and

$$\mathbb{E}_{\boldsymbol{H}}[\mathcal{P}^*(\boldsymbol{H})] = \bar{\boldsymbol{P}}.$$

We note that this last optimization problem is equivalent to solving (4) for every fading state $\boldsymbol{h}$. This completes the proof. □

The vector $\boldsymbol{\mu}$ can be interpreted as the *rate rewards* and $\boldsymbol{\lambda}$ as the *power prices*. Thus a point on the boundary of the capacity region is achieved by maximizing the total revenue for a given rate reward vector $\boldsymbol{\mu}$. Appropriate power prices have to be chosen such that the average power constraints are satisfied.

The computation of the optimal power control is now reduced to solving the optimization problem (4). This is a linear program but one with an exponentially large number of constraints (in M). However, as in Part I, we exploit the polymatroid structure of the problem to provide a simple greedy solution to this problem. Recall the definition of *contra-polymatroids* in Definition 3.1 of Part I. It is straightforward to verify (Corollary 3.13 of Part I) that for a given rate vector $\boldsymbol{R}^*$ and fading state $\boldsymbol{h}$, the set of *received powers* that can support $\boldsymbol{R}^*$

$$\mathcal{G}(\boldsymbol{R}^*) \equiv \{\boldsymbol{q} : q_i = h_i p_i, \boldsymbol{R}^* \in \mathcal{C}_g(\boldsymbol{h}, \boldsymbol{p})\}$$

is a contra-polymatroid with rank function

$$f(S) = \exp(2\boldsymbol{R}^*(S)) - 1.$$

Applying Lemma 3.3 in Part I, the optimization problem (4) can be readily solved

$$p^*_{\pi(i)} = \begin{cases} \dfrac{\sigma^2}{h_{\pi(1)}}[\exp(2R^*_{\pi(1)}) - 1], & \text{if } i = 1 \\ \dfrac{\sigma^2}{h_{\pi(i)}}\left[\exp\left(2\displaystyle\sum_{k=1}^{i} R^*_{\pi(k)}\right)\right. \\ \quad \left. - \exp\left(2\displaystyle\sum_{k=1}^{i-1} R^*_{\pi(k)}\right)\right], & i = 2, \cdots, M \end{cases} \tag{6}$$

where the permutation π satisfies

$$\frac{\lambda_{\pi(1)}}{h_{\pi(1)}} \geq \cdots \geq \frac{\lambda_{\pi(M)}}{h_{\pi(M)}}. \tag{7}$$

This optimal point corresponds to successive decoding in the order given by π, with power allocated to the users such that the target delay-limited rate vector $\boldsymbol{R}^*$ is achieved. One can think of the successive decoding order π as a way to give priority to different users in the scheduling of resources; a user decoded later in the ordering is given higher priority than a user decoded earlier. This is because users need less transmit power to support their target rates when they are decoded later (note that $(\pi(i))_{i=1}^M$ is in *decreasing* priority order). The scheduling rule here depends on both the power prices $\boldsymbol{\lambda}$ and the current fading state. In fact, this rule is analogous to the classic $c - \mu$ rule in scheduling theory (see, e.g., [14]), as both arise from the polymatroid structure of the problem. The additional feature in our problem is that the scheduling priority is a dynamic function of the fading state. Another interesting aspect of the solution to the optimization problem (4) is that the solution depends on the power prices $\boldsymbol{\lambda}$ only via the decoding order. This will simplify our later analysis.

Note that when the power price vector $\boldsymbol{\lambda}$ is strictly positive, then with probability 1 the ordering is uniquely defined since the fading processes have a continuous stationary distribution. Thus with probability 1, the solution to the optimization problem (4) is unique. Let us then define $\boldsymbol{P}(\boldsymbol{R}^*, \boldsymbol{\lambda})$ to be the unique average power vector corresponding to the almost surely unique power-control policy which solves the optimization problem (4).

In the common case when the fading processes of the users are independent of each other, the average power vector $\boldsymbol{P}(\boldsymbol{R}^*, \boldsymbol{\lambda})$ has a simple form

$$P_i(\boldsymbol{R}^*, \boldsymbol{\lambda}) = (\exp(2R_i^*) - 1) \int_0^\infty \frac{\sigma^2}{h_i} \prod_{k \neq i} \left\{ P\left(h_k > \frac{\lambda_k}{\lambda_i} h_i\right) + P\left(h_k \leq \frac{\lambda_k}{\lambda_i} h_i\right) \exp(2R_k^*) \right\} f_i(h_i)\, dh_i. \tag{8}$$

This expression can be obtained by noting that the power allocated to user i depends only on which users have values $\frac{\lambda_k}{h_k}$ greater than that of user i. Note that due to the special structure of the optimal power control policy, the computation of the average power has been reduced from an M-dimensional integral to a one-dimensional integral.

Combining this with Lemma 3.2, we have the following characterization of the delay-limited capacity region:

Theorem 3.3: Assume the fading processes of users are independent of each other. Then the rate vector $\boldsymbol{R}$ lies in the delay-limited capacity region $\mathcal{C}_d(\bar{\boldsymbol{P}})$ if and only if there exists $\boldsymbol{\lambda} \in \Re_+^M$ such that

$$(\exp(2R_i)-1)\int_0^\infty \frac{\sigma^2}{h_i} \times \prod_{k\neq i}\left\{1+F_k\left(\frac{\lambda_k}{\lambda_i}h_i\right)(\exp(2R_k)-1)\right\} f_i(h_i)\,dh_i \le \bar{P}_i, \quad i=1,\cdots,M. \tag{9}$$

The power-allocation policy that achieves this rate $\boldsymbol{R}$ is given by (6). Moreover, $\boldsymbol{R}$ lies on the boundary surface if and only if there exists $\boldsymbol{\lambda}$ such that (9) holds with equality.

We can also consider a set $\mathcal{D}_d(\boldsymbol{R}^*)$: this is the set of average power vectors that can support target delay-limited rates $\boldsymbol{R}^*$, i.e.,

$$\mathcal{D}_d(\boldsymbol{R}^*) \equiv \{\boldsymbol{P} : \boldsymbol{R}^* \in \mathcal{C}_d(\boldsymbol{P})\}.$$

Note that $\mathcal{D}_d(\boldsymbol{R}^*)$ is the structure in the power space that plays the same role as the capacity region $\mathcal{C}_d(\bar{\boldsymbol{P}})$ in the rate space. The above results lead to an explicit characterization of the boundary surface of $\mathcal{D}_d(\boldsymbol{R})$, parameterized by $\boldsymbol{\lambda}$.

Theorem 3.4: Assume the fading processes of users are independent of each other. Then the following equation gives an explicit parameterization of the boundary surface of the region $\mathcal{D}_d(\boldsymbol{R}^*)$ by $\boldsymbol{\lambda} \in \Re_+^M$:

$$P_i(\boldsymbol{\lambda}) = (\exp(2R_i^*)-1)\int_0^\infty \frac{\sigma^2}{h_i} \times \prod_{k\neq i}\left\{1+F_k\left(\frac{\lambda_k}{\lambda_i}h_i\right)(\exp(2R_k^*)-1)\right\} f_i(h_i)\,dh_i \quad i=1,\cdots,M. \tag{10}$$

It is important to note that the dimension of the boundary surface of $\mathcal{D}_d(\boldsymbol{R}^*)$ is $M-1$, which implies that the parameterization in Theorem 3.4 is onto, but not 1–1. Clearly, however, we can normalize λ to provide a 1–1 parameterization. An example of such a normalization is provided in Section III, Example 3).

The above results still leave open important questions: 1) how to check algorithmically if a target rate vector $\boldsymbol{R}$ is achievable, i.e., in the capacity region $\mathcal{C}_d(\bar{\boldsymbol{P}})$, and 2) how to find the appropriate power prices $\boldsymbol{\lambda}$ if $\boldsymbol{R}$ is indeed in the region. We will return to these questions in Section IV. But first, let us look at some special cases of Theorem 3.3.

B. Examples

1) **Single-User Channel:** When $M=1$, the delay-limited capacity $C_d(\bar{P})$ is given by

$$C_d(\bar{P}) = \frac{1}{2}\log\left(1+\frac{\bar{P}}{\sigma^2\int_0^\infty \frac{f(h)}{h}dh}\right). \tag{11}$$

The corresponding power control strategy inverts the channel [8]. We note that for some fading distributions, the delay-limited capacity may be zero. For example, for Rayleigh fading

$$f(h) = \frac{1}{a}\exp\left(-\frac{h}{a}\right)$$

and $\int_0^\infty \frac{f(h)}{h}dh = \infty$, so $C_d = 0$. The problem is that the channel is spending a lot of time close to zero. One approach to deal with this is to allow an event of outage when the channel gets too weak. (This is the approach taken by Ozarow *et al.* [12] and Cheng [3] for situations where there is no power control.) Thus even for these fading distributions, it is meaningful to consider the notion of delay-limited capacity during the times when the channel is reasonable, and declare an outage otherwise. This issue is investigated further in subsequent work by Caire *et al.* [2]. For many other distributions, such as the log-normal distribution for shadow fading, a nonzero delay-limited capacity is obtained even without the need of allowing outage.

2) **Symmetrical Case [10]:** Consider the case when there are M users, the fading of users are identical and independent, and their power constraints are the same. The symmetric delay-limited capacity C_d is the maximum common rate that can be achieved, and can be obtained by putting $\lambda_i = 1$ for all i in (9). Simplifying, we find that the capacity satisfies:

$$\sigma^2[\exp(2C_d)-1]\int_0^\infty [1+F(h)(\exp(2C_d)-1)]^{M-1}\frac{f(h)}{h}\,dh = \bar{P}.$$

The optimal power control policy has an interesting form. Namely, users are decoded in the order of decreasing channel strengths, with the strongest user decoded first and the weakest user decoded last. Powers are allocated accordingly. If channel strength is determined primarily by the distance to the base station, then this optimal decoding order results in the smallest possible transmit power for the furthest user to support the desired rate, as he only has to compete with the background noise and not the interference from any other user. This property is particularly appealing in terms of reducing intercell interference, as the furthest user will likely cause the most interference in an adjacent cell. Contrast this with the IS-95 CDMA scheme, in which the furthest user has to compete with all other users so that his received power has to be the same as that of the closest user.

3) **Two-User Capacity Region:** When $M=2$, the boundary of the delay-limited capacity region can be directly calcu-

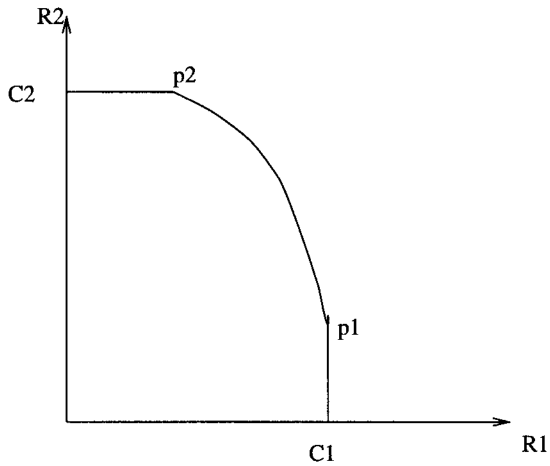

Fig. 1. A two-user delay-limited capacity region. The curved part is the boundary surface. The points p_1 and p_2 are the two extreme points of the surface. The point p_1 corresponds to giving absolute priority to user 1, i.e., decoding user 1 after user 2 at every fading state. At this point, user 1 gets rate $C_1 = \frac{1}{2}\log(1+\frac{\bar{P}_1}{A_1})$. And *vice versa* for point p_2. Note that all other points in the capacity region but not on the curved boundary are dominated by some point on the boundary.

lated by solving (9). Let $\lambda \equiv \frac{\lambda_1}{\lambda_2}$. Then the boundary is the following parametric curve as λ ranges from 0 to ∞ (see the bottom of this page), where

$$
\begin{aligned}
A_m &\equiv \int_0^\infty \frac{\sigma^2}{h} f_m(h)\,dh, \qquad m=1,2\\
B_1(\lambda) &\equiv \int_0^\infty \frac{\sigma^2}{h} F_2\left(\frac{h}{\lambda}\right) f_1(h)\,dh\\
B_2(\lambda) &\equiv \int_0^\infty \frac{\sigma^2}{h} F_1(\lambda h) f_2(h)\,dh.
\end{aligned}
$$

The parameter λ can be viewed as a prioritization between the two users. As $\lambda \to 0$, $B_1(\lambda) \to A_1$, and $B_2(\lambda) \to 0$ so $R_2(\lambda) \to \frac{1}{2}\log(1+\frac{\bar{P}_2}{A_2})$. This is the delay-limited capacity of user 2 when it is given strict priority over user 1 in all fading states (i.e., decoded last), and this is the best rate user 2 can get. Similarly, as $\lambda \to \infty$, $B_1(\lambda) \to 0$, and $B_2(\lambda) \to A_2$ so $R_1(\lambda) \to \frac{1}{2}\log(1+\frac{\bar{P}_1}{A_1})$. This is the delay-limited capacity of user 1 when it is given strict priority in all fading states, and this is the best rate user 1 can get. For λ in between these two extremes, the decoding order of users 1 and 2 changes depending on the fading state. See Fig. 1 for an illustration. Note that in this two-user case, we can parameterize the boundary surface of $\mathcal{C}_d(\bar{\boldsymbol{P}})$ by $\lambda \in \Re_+$. We will comment on whether this can be done in the general M-user case in Appendix B.

C. Extreme Points of Boundary Surface

We now extend the characterization of the interior points of the boundary surface of $\mathcal{D}_d(\boldsymbol{R})$ to include the extreme points.

Suppose $\mathcal{L}$ is a set of subsets of $E \equiv \{1,2,\cdots,M\}$ with the property that all subsets in $\mathcal{L}$ are nested. By this we mean that if $F_1, F_2 \in \mathcal{L}$ then $F_1 \subseteq F_2$ or $F_2 \subseteq F_1$. This nesting property enables us to define a new decoding rule. Let us use successive decoding, with the ordering determined by $\boldsymbol{\lambda}$ as before, except now all users in any set $F \in \mathcal{L}$ are decoded after users in F^c, for every fading state $\boldsymbol{h}$. Thus if $F_1 \subseteq F_2 \cdots \subseteq E$, then $(\boldsymbol{\lambda}_i)_{i\in F_1}$ is used to determine the ordering of users in F_1, and all these users provide interference to users in F_1^c. In particular, users in F_1 provide interference to the users in $F_2 \backslash F_1$. By (6), the powers of users in $F_2\backslash F_1$ are not affected by the ordering of users in F_1, so the ordering of users in $F_2\backslash F_1$ is entirely determined by $(\lambda_i)_{i\in F_2\backslash F_1}$. Inductively, $(\boldsymbol{\lambda}_i)_{i\in F_n\backslash F_{n-1}}$ is used to determine the ordering of users in $F_n\backslash F_{n-1}$, and all the users in F_n provide interference to the users in F_n^c. It is not difficult to show that all extreme points of the boundary surface of $D_d(\boldsymbol{R}^*)$ are obtained in this way. For the two-user example in Fig. 1, the extreme points are p_1 and p_2.

Let us also extend the notion of $P_i(\boldsymbol{R}^*, \boldsymbol{\lambda})$ in the following way.

Definition 3.5: Given $\boldsymbol{R}, \boldsymbol{\lambda} \in \Re_+^M$ and $\mathcal{L}$, a set of nested subsets of E, we denote the power vector characterized by $(\boldsymbol{R}, \boldsymbol{\lambda}, \mathcal{L})$ by $\boldsymbol{P}(\boldsymbol{R}, \boldsymbol{\lambda}, \mathcal{L})$.

Note that $\boldsymbol{P}(\boldsymbol{R}, \boldsymbol{\lambda})$ is not an extreme point of the boundary surface of $\mathcal{D}_d(\boldsymbol{R})$, but is still representable in this notation

$$\boldsymbol{P}(\boldsymbol{R}, \boldsymbol{\lambda}) \equiv \boldsymbol{P}(\boldsymbol{R}, \boldsymbol{\lambda}, \{E\}).$$

We shall have use for this extension in Section IV.

D. Further Remarks Concerning the Coding Theorem

We would like to remark on the decoding schemes to achieve points on the boundary of the delay-limited capacity region. Consider a channel in which the fading state $\boldsymbol{H}$ is fixed at level $\boldsymbol{h}$ for all time. It follows immediately from (6) that if users are allocated powers in $\mathcal{P}(\boldsymbol{h})$ then $\boldsymbol{R}^*$ is achievable by successive decoding. We conclude that if the fading is sufficiently slow that it does not change during the blocklength then the optimal solution is to do successive decoding with powers allocated as in (6). This separation of time-scales assumption may be quite reasonable if $\boldsymbol{H}(n)$ is a slow-fading process in relation to the tolerable coding delay (e.g., shadow fading). If $\boldsymbol{H}(n)$ changes during the blocklength then the optimal power control is still given by (6): it is *as if* successive decoding were being employed as far as power control was concerned, and we shall say that the optimal

$$
\begin{aligned}
R_1(\lambda) &= \frac{1}{2}\log\left[1 + \frac{B_2(\lambda)\bar{P}_1 - B_1(\lambda)\bar{P}_2 - A_1A_2 + \sqrt{(B_1(\lambda)\bar{P}_2 - B_2(\lambda)\bar{P}_1 + A_1A_2)^2 + 4A_1A_2B_2(\lambda)\bar{P}_1}}{2A_1B_2(\lambda)}\right]\\
R_2(\lambda) &= \frac{1}{2}\log\left[1 + \frac{B_1(\lambda)\bar{P}_2 - B_2(\lambda)\bar{P}_1 - A_1A_2 + \sqrt{(B_2(\lambda)\bar{P}_1 - B_1(\lambda)\bar{P}_2 + A_1A_2)^2 + 4A_1A_2B_1(\lambda)\bar{P}_2}}{2A_2B_1(\lambda)}\right]
\end{aligned}
$$

solution is of "successive decoding type." If we try to do successive decoding, we face the problem that the optimal ordering of the users may change during the blocklength, if the fading changes. This situation does not arise in the non-delay-limited case; successive decoding is optimal as shown in Part I. It may be possible to extend successive decoding techniques to deal with fading in the delay-limited case (an open problem). In practice, it may be sufficient to update the successive decoding order at the start of each code period, and make an allowance for the fading that occurs within the blocklength. We would then sacrifice optimality for ease of decoding.

IV. An Iterative Algorithm for Resource Allocation

In the previous section, we have characterized the structure of the optimal power allocation and used it for an implicit characterization of the delay-limited capacity region $\mathcal{C}_d(\bar{\boldsymbol{P}})$. The power prices $\boldsymbol{\lambda}$ play a central role as a mechanism through which resource is allocated to the different users. To achieve a target delay-limited rate vector $\boldsymbol{R}^*$, we have shown that a simple optimal power control can be obtained, for a *given* power price vector $\boldsymbol{\lambda}$. Since the power prices reflect the power constraints on the users, a natural question then is how an appropriate power price vector can be computed for given power constraints. More specifically, we will be concerned with the following problem.

- Is a target delay-limited rate vector $\boldsymbol{R}^*$ achievable under a given average power constraint $\bar{\boldsymbol{P}}$? If so, what is an appropriate power price vector?

In the case of independent fading processes, this problem is equivalent to checking if there exists $\boldsymbol{\lambda}$ such that inequalities (9) can be satisfied. From a networking point of view, a solution to this problem serves the dual purposes of *call admissions* and *resource allocation*. It determines if a set of users with specified rate requirements is supportable and if so allocates appropriate amount of resources via the selection of the power prices.

An equivalent formulation is the optimization problem

$$\inf_{\boldsymbol{\lambda}>0} \max_{1\le i\le M} \frac{P_i(\boldsymbol{R}^*, \boldsymbol{\lambda})}{\bar{P}_i} \tag{12}$$

where $P_i(\boldsymbol{R}^*, \boldsymbol{\lambda})$ is the average power of the ith user under the optimal power control which minimizes the total power cost $\boldsymbol{\lambda}\cdot\boldsymbol{P}$ while achieving rates $\boldsymbol{R}^*$. (In the case of independent fading, $P_i(\boldsymbol{R}^*, \boldsymbol{\lambda})$ is given by the explicit expression (8).) By Lemma 3.2, the target rate vector $\boldsymbol{R}^*$ is achievable with power constraints $\bar{\boldsymbol{P}}$ if and only if the solution to (12) is no greater than 1. This optimization problem can also be interpreted as finding a solution for *fair* average powers for the users, weighted by the power constraints of the users.

We will provide an iterative algorithm that solves the problem (12). If the infimum in (12) is achieved at a positive $\boldsymbol{\lambda}^*$, the algorithm will converge to it. If this is not the case, then a solution achieving the infimum in (12) must be an extreme point of the boundary surface of $\mathcal{D}_d(\boldsymbol{R}^*)$ (the set of average power vectors that can support $\boldsymbol{R}^*$). More generally, we can represent all points on the boundary surface of $\mathcal{D}_d(\boldsymbol{R}^*)$, including extreme points, by $\boldsymbol{P}(\boldsymbol{R}^*, \boldsymbol{\lambda}, \mathcal{L})$, where $\mathcal{L}$ is a set of nested subsets of users giving absolute priority rules that hold irrespective of the fading state. This was discussed in Section III-C. In general, our algorithm provides the parameters $\boldsymbol{\lambda}^*$ and $\mathcal{L}^*$, and provably converges to the point $\boldsymbol{P}^* \equiv \boldsymbol{P}(\boldsymbol{R}^*, \boldsymbol{\lambda}^*, \mathcal{L}^*)$ such that $\boldsymbol{P}^*$ is an optimal solution to (12).

First, it is necessary to develop some notation. Since we assume that $\boldsymbol{R}^*$ is fixed throughout this section, we shall simplify notation and set

$$\boldsymbol{P}(\boldsymbol{\lambda}) \equiv \boldsymbol{P}(\boldsymbol{R}^*, \boldsymbol{\lambda})$$
$$\boldsymbol{P}(\boldsymbol{\lambda}, \mathcal{L}) \equiv \boldsymbol{P}(\boldsymbol{R}^*, \boldsymbol{\lambda}, \mathcal{L}).$$

We call $P_i(\boldsymbol{\lambda})$ the average power of user i at power prices $\boldsymbol{\lambda}$, where it is understood that this is the average power to achieve the rate vector $\boldsymbol{R}^*$ and minimize the total power cost $\boldsymbol{\lambda}\cdot\boldsymbol{P}$. Also without loss of generality, we can assume that the average power constraint $\bar{P}_i$ is 1 for all users, by appropriate rescaling of the fading processes. Hence, our problem is

$$\inf_{\boldsymbol{\lambda}>0} \max_i P_i(\boldsymbol{\lambda}).$$

We propose the following iterative algorithm for solving this problem. The basic idea is that at any iteration of the algorithm, we balance the required average powers of all users as much as possible by increasing the power prices of the users with larger average powers. This will result in lowering the required power of such users by giving them higher priority in the decoding order in more of the fading states. However, perfect balancing is not always possible since the required power of a user cannot be lowered beyond giving him highest priority (i.e., last in the decoding order) at *every* fading state.

Algorithm 4.1:

- **Initialization**: Start with an arbitrary positive $\boldsymbol{\lambda}^{(1)}$. Set $k = 1$.
- **Step** k: Increase the power price of the user with the largest average power $P_i(\boldsymbol{\lambda}^{(k)})$ until its power equals that of another user, keeping the power prices of other users fixed. Then increase the power prices of *both* users by the same factor until the average power of one of them equals that of a third user. Repeat the process and consider two cases.
  1) The process continues until there are no more users left. In this case, let the final value of the power prices be $\boldsymbol{\lambda}^{(k+1)}$ and go to step $k+1$.
  2) The process terminates when the powers of a subset U of users whose prices are being increased do not meet the power of any of the other users, even when the prices of that subset are increased to infinity. In this case, perfect balancing of powers between the two subsets is impossible, even when absolute priority is given to the users in subset U. Partition the users into U and L, the subset of remaining users. The users in U from this step on will always be given absolute priority over users in L. The power prices of each user i in L will be fixed at $\lambda_i^{(k)}$ and will not be further adjusted in the algorithm. The algorithm

is now recursively applied to U, using their current power prices as initialization.

After a finite number of iterations of this algorithm, the users will be partitioned into subsets $L_1, L_2, \cdots, L_K$, and H, where users in L_i is given absolute priority over users in L_j for $i > j$ and users in H given the highest priority, and such that no further partitioning of H will take place. Let

$$\mathcal{L}^* = \left\{ H, H \cup L_K, H \cup L_{K-1} \cup L_K, \cdots, H \bigcup_{i=1}^{K} L_i, E \right\}$$

be the absolute priority nesting corresponding to this partitioning of the users. We have the following convergence theorem under a mild condition on the fading distributions.[2]

Theorem 4.2: Assume that there exists a positive lower bound ϵ to the fading gains of all the users. If $\boldsymbol{\lambda}^{(n)}$ is the vector of power prices at iteration n, then

$$P_i^* \equiv \lim_{n\to\infty} P_i(\boldsymbol{\lambda}^{(n)}, \mathcal{L}^*)$$

exists for all i, and $P^* \equiv \max_i P_i^*$ is the optimal value for the problem (12), i.e.,

$$P^* = \inf_{\boldsymbol{\lambda}>0} \max_i P_i(\boldsymbol{\lambda}).$$

Moreover, $P_i^* = P^*$ for every user i in H.

Proof: First, we observe that for any j, the power allocation of the users in the subset L_j does not change after the iteration when the subset L_j is created. To see that, fix a subset L_j, and let $(\lambda_i^*)_{i\in L_j}$ be the power prices of the users in L_j when L_j is created. Let $H_j \equiv L_{j+1} \cup \cdots \cup L_K \cup H$; this is the subset of users which are given higher priority than users in L_j at all fading states. The rest of the users (in $L_1, \cdots, L_{j-1}$) will be given lower priority than users in L_j at all fading states. The optimal power allocation to users in L_j at fading state $\boldsymbol{h}$ is given by (6)

$$\begin{aligned}\mathcal{P}_{\pi(i)}(\boldsymbol{h}) = \frac{\sigma^2}{h_{\pi(i)}} \Bigg[&\exp\left\{ 2\left(\sum_{k\in H_j} R_k^* + \sum_{k=1}^{i} R_{\pi(k)}^* \right) \right\} \\ &- \exp\left\{ 2\left(\sum_{k\in H_j} R_k^* + \sum_{k=1}^{i-1} R_{\pi(k)}^* \right) \right\} \Bigg], \\ &i = 1, \cdots, |L_j|\end{aligned}$$

where π is an ordering of users in the subset L_j satisfying

$$\frac{\lambda_{\pi(1)}^*}{h_{\pi(1)}} \geq \cdots \geq \frac{\lambda_{\pi(|L_j|)}^*}{h_{\pi(|L_j|)}}. \tag{13}$$

The key point is that the power allocation to users in L_j only depends on the power prices of users in L_j, which remain fixed after the iteration when L_j is created, but do not depend on the power prices of the users of higher priority in H_j, which will be changed in future iterations (see Section III-C). Thus the power allocation to users in L_j stays fixed once L_j is created.

[2]This assumption is technical and can probably be removed with a more elaborate argument.

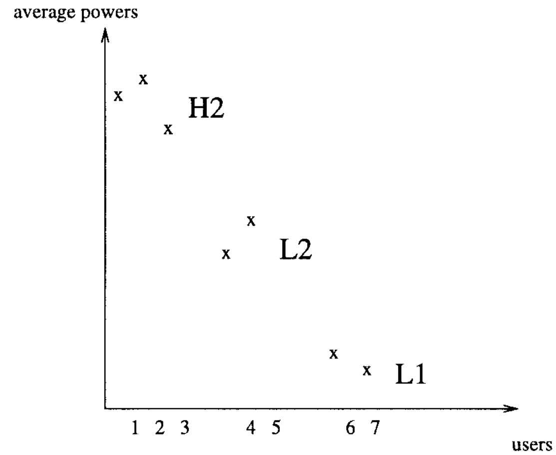

Fig. 2. The average powers of the users at the start of an iteration of the resource-allocation algorithm. Users are currently partitioned into subsets L_1, L_2, and H_2. Users in H_2 have the largest average powers, and are decoded after users in L_2, which are, in turn, decoded after users in L_1, at every fading state. The power prices of the users in H_2 will be adjusted in future iterations to further balance their powers; the power prices and average powers of users in L_1 and L_2 will stay fixed.

Second, we see from the definition of the algorithm that for each j, the minimum of the average powers of users in H_j (high-priority users) must monotonically increase after the iteration when L_j is formed.

It can also be seen that for each j, when the partitioning into H_j and L_j occurs, the minimum of the average powers in H_j must be greater than the maximum of the average powers in L_j. Combining this with the two observations above, we conclude at any iteration after L_j is created, the average power of any user in H_j must be greater than that of any user in $L_1, \cdots, L_j$. A typical situation is shown in Fig. 2. In particular, at any iteration after all of $L_1, L_2, \cdots, L_K$ are created, the average powers of any user in H must be greater than the average power of any user in $\cup_{j=1}^{K} L_j$.

Now, let us investigate the limiting behavior of the powers of the users in H, the final set of users for which no further splitting occurs. First we observe that the ratio of the power prices of the users in H must remain bounded. This follows from the assumption that the fading gains h_i are bounded, since then (13) implies that whenever the ratio of the prices of one subset of users to those of another subset exceeds a certain threshold, strict priority will be given to users in the former subset over the latter. But this would contradict the fact that there will be no further splitting of users in H.

Let

$$T : \frac{\boldsymbol{\lambda}^{(k)}}{\min_i \lambda_i^{(k)}} \mapsto \frac{\boldsymbol{\lambda}^{(k+1)}}{\min_i \lambda_i^{(k+1)}}$$

be the mapping representing one step of the algorithm updating the power prices of the users in H, normalized by the minimum prices. (Note that the average power of the users depends only on the normalized prices.) From the continuity of the fading distribution, it can be verified that the average powers are continuous functions of the normalized prices and so is T.

We now claim that T has a unique fixed point, to which any sequence of power prices generated by the algorithm converges. To see uniqueness, we note that $\boldsymbol{\lambda}$ provides a *tradeoff* between powers of users. No point on the boundary surface of $\mathcal{D}_d(R^*)$ strictly dominates another. This is true also for powers within H, if we give strict priority to H. However, the existence of more than one fixed point of T would provide such strict dominance of powers in H of one point over the other. This is so, since a fixed point of T must have that property that the powers of all users in H are the same.

To show the existence of a fixed point of T, we use the Lyapunov technique. Define the Lyapunov function V by

$$V(\lambda) = \max_{i \in H} P_i(\lambda, \mathcal{L}^*) - \min_{i \in H} P_i(\lambda, \mathcal{L}^*)$$

for any normalized price vector λ, which has the following properties:

- V is a continuous function of λ;
- for any λ, $V(T(\lambda)) \le V(\lambda)$ with equality if and only if λ is a fixed point of T.

Consider now any particular sequence of power prices, $(\boldsymbol{\lambda}^{(k)})_{k=1}^{\infty}$ generated by the algorithm. By assumption, the sequence is contained in a compact set, and therefore has accumulation points. Let $(\boldsymbol{\lambda}^{(n_k)})_{k=1}^{\infty}$ be a subsequence of points converging to an accumulation point λ^*. By the continuity of V, we have that $V(\boldsymbol{\lambda}^{(n_k)}) \downarrow V(\lambda^*)$, as $k \uparrow \infty$, and by a sandwich argument, $V(\boldsymbol{\lambda}^{(n_k+1)}) \downarrow V(\lambda^*)$ as well. By the continuity of T, $\boldsymbol{\lambda}^{(n_k+1)} \equiv T(\boldsymbol{\lambda}^{(n_k)}) \to T(\boldsymbol{\lambda}^*)$ so it follows that $V(T(\boldsymbol{\lambda}^*)) = V(\boldsymbol{\lambda}^*)$, i.e., $\boldsymbol{\lambda}^*$ is the unique fixed point of T. Hence the algorithm converges to the unique fixed point of T, and that for any sequence of power prices $\lambda^{(k)}$ generated by the algorithm, the corresponding powers $(P_i(\lambda^{(k)}, \mathcal{L}^*)_{i \in H})$ converge to a common value P^*.

Thus we have proved that for every user $i \in E$

$$P_i^* \equiv \lim_{n \to \infty} P_i(\boldsymbol{\lambda}^{(n)}, \mathcal{L}^*)$$

exists and

$$P_i^* = P^* = \max_{k \in E} P_k^* \text{ for every } i \in H.$$

Also, for any power price vector $\boldsymbol{\lambda}$, $\max_{i \in H} P_i(\boldsymbol{\lambda}) \ge P^*$. To see this, assume this is *not* the case. Then there exists $\boldsymbol{\lambda}$ and n such that

$$P_i(\boldsymbol{\lambda}) < P_i(\boldsymbol{\lambda}^{(n)}, \mathcal{L}^*) \text{ for every } i \in H.$$

This is impossible since under $\mathcal{L}^*$, users in H are already given the highest priority over other users at all fading states and hence $\boldsymbol{P}(\boldsymbol{\lambda}^{(n)}, \mathcal{L}^*)$ achieves the minimum total average power cost

$$\sum_{i \in H} \lambda_i^{(n)} P_i$$

for users in H. Thus $P^* = \inf_{\boldsymbol{\lambda} > 0} \max_i P_i(\boldsymbol{\lambda})$ and the proof is complete. □

For the reader who is familiar with flow-control problems in virtual circuit networks, this algorithm may be reminiscent of *fair bandwidth-allocation* algorithms. Here, the objective is to find a fair average power requirements for the users, weighted by their power constraints. Users in the set H correspond to users whose routes pass through the *bottleneck node*, and have the maximum (weighted) power requirement. In fact, it can be shown that by applying the algorithm recursively to balance the power requirements of users in the subsets $L_1, \cdots, L_k$ defined above, one can in fact compute a *min-max fair* solution (see [1] for a corresponding algorithm for bandwidth allocation).

V. Auxiliary Constraints on Transmit Power

The constraints on the transmit powers we considered so far are on their long-term *average* value, and under power control, the transmit power will vary depending on the fading state. In practice, one often wants to have some shorter term constraints on the transmit power as well. These constraints may be due to regulations, or as a way of imposing a more stringent limit on how much interference a mobile can cause to adjacent cells. To model such auxiliary constraints, we consider the following feasible set of power controls:

$$\mathcal{F}_p \equiv \{\mathcal{P} : \mathbb{E}_{\boldsymbol{H}}[\mathcal{P}_i(\vec{H})] \le \bar{P}_i \text{ and } \mathcal{P}_i(\boldsymbol{h}) \le \hat{P}_i \ \ \forall i \text{ and } \boldsymbol{h} \in \mathcal{H}\}$$

where $\mathcal{H}$ is the set of all possible joint fading states of the users. Thus in addition to the average power constraints, we also have a constraint $\hat{P}_i$ on the transmit power of the ith user in every state. We shall now concentrate on the problem of computing the optimal power control subject to these constraints.

We focus on the capacity region

$$\mathcal{C}_d^p(\bar{\boldsymbol{P}}, \hat{\boldsymbol{P}}) \equiv \bigcup_{\mathcal{P} \in \mathcal{F}_p} \bigcap_{\boldsymbol{h} \in \mathcal{H}} \mathcal{C}_g(\boldsymbol{h}, \mathcal{P}(\boldsymbol{h})).$$

Lemma 5.1: A rate vector $\boldsymbol{R}^*$ lies on the boundary of $\mathcal{C}_d^p(\bar{\boldsymbol{P}}, \hat{\boldsymbol{P}})$ if and only if there exists a $\boldsymbol{\lambda} \in \Re^M$ and a power control policy $\mathcal{P}$ such that for every joint fading state $\boldsymbol{h}$, $\mathcal{P}(\boldsymbol{h})$ is a solution to the optimization problem

$$\min \boldsymbol{\lambda} \cdot \boldsymbol{p} \text{ subject to } \boldsymbol{R}^* \in \mathcal{C}_g(\boldsymbol{h}, \boldsymbol{p}) \text{ and } p_i \le \hat{P}_i \ \forall i \qquad (14)$$

and

$$\mathbb{E}_{\boldsymbol{H}}[\mathcal{P}_i(\boldsymbol{H})] = \bar{P}_i, \qquad i = 1, \cdots, M$$

where $\bar{P}_i$ is the constraint on the average power of user i. Moreover, $\mathcal{P}$ is a power control policy which can achieve the rate vector $\boldsymbol{R}^*$.

The proof of this result is similar to that of Lemma 3.2, and is the analogue of Lemma 6.1 in Part I, and will not be given here.

To solve the optimization problem (14), we first prove a few results about contra-polymatroids.

Definition 5.2: The rank function f of a contra-polymatroid is said to be **strictly** supermodular if for any subset S, T such that neither is a subset of the other[3]

$$f(S) + f(T) < f(S \cup T) + f(S \cap T).$$

[3] Clearly, if one is a subset of the other, equality must hold.

The following lemma is motivated by a result of Hanly and Whiting [11], which was proved in the context of multiaccess capacity regions.

Lemma 5.3: Let $\mathcal{D}(f)$ be a contra-polymatroid with a strictly supermodular rank function f. Consider any vector $\boldsymbol{x} \in \mathcal{D}(f)$, and let $S_1, \cdots, S_J$ be the subsets corresponding to the constraints of $\mathcal{D}(f)$ that are tight at $\boldsymbol{x}$, i.e., these constraints hold with equality at $\boldsymbol{x}$. Then there exists an ordering π such that

$$S_{\pi(1)} \subset \cdots \subset S_{\pi(J)}$$

i.e., they are nested.

Proof: Take any two tight constraints corresponding to subsets S_i and S_j. Suppose neither is a subset of the other. Then

$$\begin{aligned} \boldsymbol{x}(S_i \cup S_j) &= \boldsymbol{x}(S_i) + \boldsymbol{x}(S_j) - \boldsymbol{x}(S_i \cap S_j) \\ &\le f(S_i) + f(S_j) - f(S_i \cap S_j) \\ &< f(S_i \cup S_j) \end{aligned}$$

a contradiction, since $\boldsymbol{x} \in \mathcal{D}(f)$. Hence, the subsets corresponding to the tight constraints must be nested. □

Now let a_i's be positive constants, and consider the optimization problem

$$\min \boldsymbol{\lambda} \cdot \boldsymbol{x} \text{ subject to } \boldsymbol{x} \in \mathcal{D}(f) \text{ and } x_i \le a_i \ \forall i \tag{15}$$

where the vector $\boldsymbol{\lambda}$ satisfies

$$\lambda_1 \ge \cdots \ge \lambda_M \ge 0.$$

We will refer to the constraints $x_i \le a_i$ as *peak* constraints. To motivate the algorithm for solving this problem, we first observe that the algorithm given in Lemma 3.3 of Part I (which we applied in (6) to solve the same problem but without the peak constraints) can be viewed as a greedy algorithm.

- **Initialization**: Set $x_i^{(0)} = 0$ for all i. Set $k = 1$.
- **Step** k: Increase the value of x_k until a constraint becomes tight. Goto Step $k+1$
- After M steps, optimal solution is reached.

With this interpretation, the following greedy algorithm for problem (15) can be viewed as a natural generalization to the case when there are peak constraints:

Algorithm 5.4:

- **Initialization**: Set $x_i^{(0)} = a_i$ for all i. If $\boldsymbol{x}^{(0)} \notin \mathcal{D}(f)$ then stop. Else set $k = 1$.
- **Step** k: Decrease the kth component of $\boldsymbol{x}$ until a constraint becomes tight. Go to Step $k+1$
- Stop after M steps.

Theorem 5.5: If $\boldsymbol{x}^{(0)} \notin \mathcal{D}(f)$, then the optimization problem (15) has an empty feasible region. Otherwise, Algorithm 5.4 terminates at an optimal solution to (15).

Proof: The first statement follows from the easily verified fact that if $\boldsymbol{x}, \boldsymbol{y}$ are two vectors such that $\boldsymbol{y}_i \le \boldsymbol{x}_i \ \forall i$ and $\boldsymbol{x} \notin \mathcal{D}(f)$, then $\boldsymbol{y} \notin \mathcal{D}(f)$.

Now suppose $\boldsymbol{x}^{(0)} \in \mathcal{D}(f)$ and the Algorithm 5.4 terminates at the point $\boldsymbol{x}^*$. We first show that $\boldsymbol{x}^*$ is a vertex of the feasible region. At each step k of the algorithm, either the kth component cannot be decreased, in which case the constraint $x_k \le a_k$ is tight, or it can be decreased until a constraint of $\mathcal{D}(f)$ corresponding to some subset S becomes tight. In any case, at each stage of the algorithm, we are having an additional linearly independent constraint becoming tight. Moreover, since we are always decreasing the components of $\boldsymbol{x}$, subset constraints that become tight will remain tight. Hence, at termination, there are M linearly independent tight constraints, and $\boldsymbol{x}$ is a vertex of the feasible region.

Let $S_1, S_2, \cdots, S_J$ be the subset constraints that are tight at $\boldsymbol{x}^*$. By Lemma 5.3, we can without loss of generality assume that $S_1 \subset \cdots \subset S_J$. Let us now identify the tight peak constraints. Consider the partition of the base set E into $S_1, S_2 - S_1, S_3 - S_2, \cdots, S_J - S_{J-1}$. Since the tight constraints are all linearly independent, it follows that in each subset $S_j - S_{j-1}$, at most $|S_j - S_{j-1}| - 1$ elements can correspond to tight peak constraints. But since there are $M - J$ tight peak constraints, in fact exactly $|S_j - S_{j-1}| - 1$ elements correspond to peak constraints.

Now, the optimization problem of interest is a linear programming problem. Thus to verify the optimality of $\boldsymbol{x}^*$, it suffices to show that the objective function cannot decrease along any of the M edges of the polyhedron that emanate from $\boldsymbol{x}^*$. Each edge is obtained by relaxing precisely one of the tight constraints. We consider the following two cases.

1) Suppose we relax a tight constraint $x_k \le a_k$, where $k \in S_j - S_{j-1}$ for some j. Let $m \in S_j - S_{j-1}$ be such that the corresponding peak constraint is *not* tight. The edge can be seen to be along the half-line

$$x_k + x_m = x_k^* + x_m^*, \quad x_k \le x_k^*, \quad x_i = x_i^*, \qquad i \ne k, m.$$

We first note that $k > m$. For the purpose of contradiction, suppose instead that $k < m$. The point

$$(x_1^*, \cdots, x_k^* - \epsilon, \cdots, x_m^* + \epsilon, \cdots, x_M^*)$$

is in the feasible region, which means that in the kth step of the algorithm, the kth component can be further decreased beyond x_k^*. This is a contradiction. Hence, $k > m$. Since the coefficients of the objective function satisfy $\lambda_k \le \lambda_m$, it follows that the objective function cannot decrease moving along the edge.

2) Suppose we relax a subset constraint corresponding to S_j for some j. If $j < J$, let $k \in S_j - S_{j-1}$ and $m \in S_{j+1} - S_j$ correspond to peak constraints that are not tight at $\boldsymbol{x}^*$. In this case, the edge can be seen to be along the half-line

$$x_k + x_m = x_k^* + x_m^*, \quad x_k \ge x_k^*, \quad x_i = x_i^*, \qquad i \ne k, m.$$

Since $\lambda_k \ge \lambda_m$, it follows that the objective function cannot decrease along this edge. On the other hand, if $j = J$, let $k \in S_J - S_{J-1}$ be the component corresponding to a peak constraint that is not tight. The corresponding edge is along the half-line

$$x_k \ge x_k^*, \quad x_i = x_i^*, \qquad i \ne k.$$

Clearly, the objective function cannot decrease along this edge. Hence we conclude that indeed $\boldsymbol{x}^*$ is an optimal solution. □

At each step k, Algorithm 5.4 has to check when a constraint becomes tight. This is equivalent to the *membership problem*: given a point $\boldsymbol{x}$, check if $\boldsymbol{x}$ is in $\mathcal{D}(f)$ or not. For general contra-polymatroids, there is no known efficient combinatorial algorithm to solve this problem (checking every constraint of $\mathcal{D}(f)$ requires complexity exponential in M.) However, for the special case of contra-polymatroids with generalized symmetric rank functions, a very simple test exists. This result is due to Federgruen and Groenevelt [5].

Lemma 5.6 [5]: Suppose f is generalized symmetric, i.e., $f(\cdot) = g(\boldsymbol{y}(\cdot))$ for some convex increasing function g and vector $\boldsymbol{y}$. Given any $\boldsymbol{x}$, let σ be a permutation on E such that

$$\frac{x_{\sigma(1)}}{y_{\sigma(1)}} \leq \cdots \leq \frac{x_{\sigma(M)}}{y_{\sigma(M)}}.$$

Then $\boldsymbol{x} \in \mathcal{D}(f)$ if and only if

$$\sum_{i=1}^{m} x_{\sigma(i)} \geq g\left(\sum_{i=1}^{m} y_{\sigma(i)}\right) \qquad \forall m = 1, \cdots, M.$$

This lemma implies that one only needs to check M constraints to determine if $\boldsymbol{x}$ is a member of $\mathcal{D}(f)$, instead of $2^M - 1$. Combining this lemma with Algorithm 5.4, we can in fact compute explicitly the value to which the kth component must be decreased in the kth step of the algorithm. Thus in the case when $f(\cdot) \equiv g(\boldsymbol{y}(\cdot))$, the algorithm now becomes

- **Initialization**: Set $x_i^{(0)} = a_i$ for all i. If $\boldsymbol{x}^{(0)} \notin \mathcal{D}(f)$ then stop. Else set $k = 1$.
- **Step** k: Let $\sigma^{(k)}$ be a permutation on

  $$\{1, \cdots, k-1, k+1, \cdots, M\}$$

  such that

  $$\frac{x_{\sigma^{(k)}(1)}}{y_{\sigma^{(k)}(1)}} \leq \cdots \leq \frac{x_{\sigma^{(k)}(k-1)}}{y_{\sigma^{(k)}(k-1)}} \leq \frac{x_{\sigma^{(k)}(k+1)}}{y_{\sigma^{(k)}(k+1)}} \leq \cdots \leq \frac{x_{\sigma^{(k)}(M)}}{y_{\sigma^{(k)}(M)}}$$

  Then set

  $$x_i^{(k)} = \begin{cases} x_i^{(k-1)}, & \text{if } i \neq k \\ \max\limits_{j \neq k}[f(S_j \cup \{k\}) - \boldsymbol{x}(S_j)], & i = k \end{cases}$$

  where

  $$S_j \equiv \{\sigma^{(k)}(1), \cdots, \sigma^{(k)}(j)\}$$

  (noting that the element $\sigma^{(k)}(k)$ does not exist.)
  Go to step $k+1$.
- Stop after M steps.

Lemma 5.6 implies that at step k of the algorithm, the subset constraints that can become tight are the ones corresponding to the subsets $S_j \cup \{k\}$, for $j = 1, 2, \cdots, k-1, k+1, \cdots, M$. The value that the kth component should be decreased to is determined by the first of these constraints becoming tight. The complexity of this algorithm is $O(M^2)$.

By observing that the set of feasible received powers $\boldsymbol{q}$ that support a given rate vector $\boldsymbol{R}^*$ is a contra-polymatroid with generalized symmetric rank function, we can immediately apply the above simplified form of Algorithm 5.4 to solve the optimization problem (14). This gives an efficient way to compute the optimal power allocation at a fading state, for given power prices $\boldsymbol{\lambda}$. Moreover, the polymatroid theory yields a result of independent interest: an efficient membership test for the Gaussian capacity region. More concretely, given rate vector $\boldsymbol{R}$ and power constraint $\boldsymbol{P}$, to check the exponentially large number of constraints

$$\boldsymbol{R}(S) \leq \frac{1}{2}\log\left(1 + \frac{\boldsymbol{P}(S)}{\sigma^2}\right), \qquad S \subset E$$

one needs only to sort $\frac{P_i}{R_i}$'s in ascending order, and check the M nested constraints corresponding to that ordering.

It should be noted that unlike the optimal power-control schemes for the previous problems we considered in this paper (Parts I and II), the optimal solution for this problem cannot in general be achieved by by successive decoding of the M users. Due to the auxiliary constraints, the optimal solution is not necessarily on a vertex of the capacity region. However, Rimoldi and Urbanke [13] show that each user can be split into at most two "virtual users," such that the resulting point *can* be achieved by successive decoding of at most $2M$ virtual users. Their procedure for calculating the power levels that define the splitting is *greedy*, a fact that again arises from the generalized symmetric polymatroid structure of the Gaussian multiaccess capacity region.

VI. Multiple Time-Scale Fading and Frequency-Selective Fading Channels

The notions of throughput capacity and delay-limited capacity for fading channels can be viewed as two ends of a spectrum. If we look upon a fading channel as a set of parallel channels, one for each fading state, then the throughput capacity is the maximum total rate that one can achieve by an arbitrary allocation of rates and powers over the parallel channels, subject to a power constraint. The delay-limited capacity, on the other hand, is the maximum total rate subject to the constraint of a common rate for each of the parallel channels. Thus one can consider other notions of capacities where the rate-allocation policy is not as stringent as in the delay-limited case, but not completely arbitrary as in the throughput capacity. In this section, we will look at two applications of this idea: fading with multiple time-scale dynamics, and frequency-selective fading.

Consider first the situation when the fading processes have two components, one slow and one fast. The slow fading might be due to shadowing, for example, and the fast due to multipath. We assume that the fast fading is sufficiently fast to average out over the tolerable delay, but that we are delay-limited with respect to the slow fading. We define a notion of capacity in this context.

Let $\mathcal{S}$ be the set of joint slow states, and $\mathcal{H}$ the set of joint fading states. Let $(\boldsymbol{S}(n), \boldsymbol{H}(n))$ be the joint slow state and fading state process, with $\boldsymbol{H}(n)$ having stationary distribution $\mathcal{Q}$ on $\mathcal{H}$. However, *conditional on* $\boldsymbol{S} = \boldsymbol{s}$, $\boldsymbol{H}$ has stationary distribution $\mathcal{Q}_s$ on $\mathcal{H}$. A feasible power-allocation policy must satisfy $\mathbb{E}_{\boldsymbol{S},\boldsymbol{H}}[\mathcal{P}(\boldsymbol{S}, \boldsymbol{H})] \leq \bar{\boldsymbol{P}}$.

In order to define the notion of capacity we are interested in, we first consider the capacity of an associated channel. Consider a channel associated with a slow state s, and with an arbitrary power-allocation policy $\mathcal{P}(\boldsymbol{s},\boldsymbol{h})$. This channel has unit power, and fading process $F_i(n)\mathcal{P}_i(\boldsymbol{s},\boldsymbol{F}(n))$, where we assume that $\boldsymbol{F}(n)$ has stationary distribution $\mathcal{Q}_s$ on $\mathcal{H}$. We denote the throughput capacity region for this associated channel by $\mathcal{C}(\boldsymbol{s},\mathcal{P}(\boldsymbol{s},\cdot))$. Now we return to the original channel with both slow and fast fading, and define the **delay-limited capacity with respect to slow fading**.

Definition 6.1:

$$\mathcal{C}_{\rm ds}(\bar{\boldsymbol{P}}) \equiv \bigcup_{\mathcal{P}\in\mathcal{F}} \bigcap_{\boldsymbol{s}\in\mathcal{S}} \mathcal{C}(\boldsymbol{s},\mathcal{P}(\boldsymbol{s},\cdot))$$

where $\mathcal{F}$ is the set of power control policies $\mathcal{P}(\boldsymbol{s},\boldsymbol{h})$ for which $\mathbb{E}_{S,H}[\mathcal{P}(\boldsymbol{S},\boldsymbol{H})] = \bar{\boldsymbol{P}}$.

Alternatively, the definition of the *delay- limited capacity with respect to slow fading*, can be formulated in the same manner as in Section II. Definition 6.1 then becomes a theorem, with a proof along the same lines as the proof of Theorem 2.3. In some sense, this alternative approach is more robust. However, the approach we have taken is completely watertight provided one assumes a separation of time scales, namely, that the coding time scale is much shorter than that of the slow fading.

From the point of view of a parallel channel decomposition of the fading channel, our approach here corresponds to partitioning the parallel channels into subsets each associated with a slow fading state. In the above definition of delay-limited capacity for multiple time-scale fading channels, one is allowed to do rate allocation among the channels within each subset, but subject to the constraint that the total rate in each subset (slow state) is the same.

We now consider the problem of resource allocation; the tuning of power prices and rate rewards to achieve a particular delay-limited bit-rate vector. The dual set $\mathcal{D}_{\rm ds}(\boldsymbol{R}^*)$ is defined as usual

$$\mathcal{D}_{\rm ds}(\boldsymbol{R}^*) \equiv \{\bar{\boldsymbol{P}} : \boldsymbol{R}^* \in \mathcal{C}_{\rm ds}(\bar{\boldsymbol{P}})\}.$$

In this section, we shall limit ourselves to the characterization of the extreme points of $\mathcal{D}_{\rm ds}(\boldsymbol{R}^*)$.

As in Section III, we characterize any point on the boundary of $\mathcal{D}_{\rm ds}(\boldsymbol{R}^*)$ by solving the following problem, for every slow state $\boldsymbol{s}$

$$\min \sum_i \lambda_i^* \mathbb{E}_{\boldsymbol{H}|\boldsymbol{S}=\boldsymbol{s}}[\mathcal{P}_i(\boldsymbol{s},\boldsymbol{H})]$$
$$\text{s.t.} \sum_{i\in\mathcal{L}} R_i^* \le \mathbb{E}_{\boldsymbol{H}|\boldsymbol{S}=\boldsymbol{s}}\left[\frac{1}{2}\log\left(1+\frac{\sum_{i\in\mathcal{L}}\mathcal{P}_i(\boldsymbol{s},\boldsymbol{H})H_i}{\sigma^2}\right)\right].$$

Since the delay-limited capacity region is convex, there exist Lagrange multipliers $\boldsymbol{\mu}^*(\boldsymbol{s})$ for which $(\boldsymbol{R}^*,\mathcal{P}(\boldsymbol{s},\cdot))$ solves

$$\max_{\boldsymbol{R},\mathcal{P}} \sum_{i=1}^M (\mu_i^*(\boldsymbol{s})R_i - \lambda_i^* \mathbb{E}_{\boldsymbol{H}|\boldsymbol{S}=\boldsymbol{s}}[\mathcal{P}_i(\boldsymbol{s},\boldsymbol{H})])$$
$$\text{s.t.} \sum_{i\in\mathcal{L}} R_i \le \mathbb{E}_{\boldsymbol{H}|\boldsymbol{S}=\boldsymbol{s}}\left[\frac{1}{2}\log\left(1+\frac{\sum_{i\in\mathcal{L}}\mathcal{P}_i(\boldsymbol{s},\boldsymbol{H})H_i}{\sigma^2}\right)\right]$$

which is equivalent to solving

$$\max_{\mathcal{R}(\boldsymbol{s},\boldsymbol{h}),\mathcal{P}(\boldsymbol{s},\boldsymbol{h})} \sum_{i=1}^M (\mu_i^*(\boldsymbol{s})\mathcal{R}_i(\boldsymbol{s},\boldsymbol{h}) - \lambda_i^*\mathcal{P}_i(\boldsymbol{s},\boldsymbol{h}))$$
$$\text{s.t.} \sum_{i\in\mathcal{L}} \mathcal{R}_i(\boldsymbol{s},\boldsymbol{h}) \le \frac{1}{2}\log\left(1+\frac{\sum_{i\in\mathcal{L}}\mathcal{P}_i(\boldsymbol{s},\boldsymbol{h})h_i}{\sigma^2}\right) \tag{16}$$

for each $\boldsymbol{s},\boldsymbol{h}$. The appropriate $\boldsymbol{\mu}^*(\boldsymbol{s})$ is determined by the condition

$$\mathbb{E}_{\boldsymbol{H}|\boldsymbol{S}=\boldsymbol{s}}[\mathcal{R}_i(\boldsymbol{s},\boldsymbol{H})] = R_i^*. \tag{17}$$

A greedy algorithm for solving (16) was presented in Theorem 3.14 in Part I. Moreover, an iterative procedure for computing $\boldsymbol{\mu}^*(\boldsymbol{s})$ was provided in Algorithm 5.3 of Part I: we start with an arbitrary $\boldsymbol{\mu}(\boldsymbol{s})$ and update it until (17) holds.

In this section, we have found the minimal cost power control policy to obtain a consistent mutual information vector $\boldsymbol{R}^*$ over every slow fading state. With this power control, we can obtain any rate strictly below $\boldsymbol{R}^*$ in a delay-limited fashion with respect to slow fading. A very important observation is that to obtain this solution we do not need to know the statistics of the slow fading at all. This is because we have prescribed the delay-limited rates $\boldsymbol{R}^*$ as a constraint, but not the long-term average power consumption. The average power used is obtainable from the solution to the power control problem, but we do not need to know it *a priori*. Moreover, the algorithm that we use to determine $\boldsymbol{\mu}(\boldsymbol{s})$ does not need to know explicitly the conditional distribution of the fading process given the slow state, but rather it adapts to changes in these statistics.

Another important point is that successive decoding is optimal under our assumption of a separation of time scales between the coding time scale, and the slow-fading time scale. Given any slow state, we use successive decoding to achieve $\boldsymbol{R}^*$, as in Section III; in this case, the decoding order is a function of the slow state $\boldsymbol{s}$.

The characterization of the extreme points of $\mathcal{C}_{\rm ds}(\bar{\boldsymbol{P}})$ is slightly more complicated, and we do not attempt it here. Clearly, the calculation of the capacity region requires explicit knowledge of the statistics of the fading, including the slow fading.

Similar reasoning can be applied to the analysis of the delay-limited capacity of frequency-selective fading channels, as defined in Section VII of Part I. Under an assumption that the product of the delay spread and the Doppler spread is small, one can look upon the frequency-selective fading channel as a time-varying channel where, at each fading state, a frequency response is specified for each user, representing the multipath. Thus it can be viewed as a set of parallel channels, each one jointly specified by the fading state and the frequency. In order to be delay-limited in this channel, each user can allocate rates over the different frequencies but the total rate summed over the frequencies must be the same for each fading state. Thus the resulting optimization problem is identical to the one studied in the present section for multiple time-scale fading processes, and hence the optimal power allocation for given delay-limited rates can be obtained from our theory. This ability of being able to perform dynamic power allocation over different frequencies is an advantage of a wideband system over a narrowband system, especially for delay-sensitive traffic.

VII. Power Control for Suboptimal Systems

In the previous sections, we have focussed on optimal power control from an information-theoretic point of view. We will now demonstrate that the ideas can also be applied, in a straightforward manner, to characterize optimal power control laws for situations when successive decoding is done but nonideal single-user codes are used so that one is not operating at information-theoretic limits.

Consider the multiaccess scenario with M users where the mth user has a desired signal-to-interference ratio SIR of α_m. Here, the interference is the sum of the background noise (with power σ^2) and that caused by the users whose signals have not yet been decoded. In general, the SIR requirement of a user depends on the coding scheme, the data rate, and the error probability requirement, but we assume that the SIR captures the quality of service requirement of the user. We now ask what is the optimal power control law which maintains the SIR requirements of the users? Focus first on a time-invariant multiaccess Gaussian channel where user m has a transmit power of P_m and a path gain h_m. For a given successive decoding order π, let $\mathcal{F}(\pi, \boldsymbol{\alpha}, \boldsymbol{h})$ be the set of transmit power vectors $\boldsymbol{P} = (P_1, \cdots, P_M)$ which can support the given SIR vector $\boldsymbol{\alpha} = (\alpha_1, \cdots, \alpha_M)$. It is given by

$$\mathcal{F}(\pi, \boldsymbol{\alpha}, \boldsymbol{h}) = \left\{ \boldsymbol{p} : \frac{h_{\pi(m)} p_{\pi(m)}}{\sigma^2 + \sum\limits_{i<m} h_{\pi(i)} p_{\pi(i)}} \geq \alpha_{\pi(m)} \ \forall m \right\}.$$

Thus if successive decoding is used, the set of transmit power vectors that can support a given set of SIR requirements $\boldsymbol{\alpha}$ is given by

$$\bigcup_{\pi} \mathcal{F}(\pi, \boldsymbol{\alpha}, \boldsymbol{h}). \tag{18}$$

Further, if we allow time sharing between different successive decoding orders, then the set of feasible power vectors is enlarged to the convex hull of (18). Call this polytope $\mathcal{F}(\boldsymbol{h}, \boldsymbol{\alpha})$.

If we let $R_m \equiv \frac{1}{2}\log(1+\alpha_m)$, i.e., the single-user capacity that can be achieved with a SIR of α_m, then we observe that the set $\mathcal{F}(\boldsymbol{h}, \boldsymbol{\alpha})$ is the same as

$$\mathcal{G}(\boldsymbol{h}, \boldsymbol{R}) \equiv \{\boldsymbol{p} : \boldsymbol{R} \in \mathcal{C}_g(\boldsymbol{h}, \boldsymbol{p})\}$$

i.e., the set of transmit power vectors such that the rate vector $\boldsymbol{R}$ is in the multiaccess Gaussian capacity region. To see this, note that the only vertex of $\mathcal{F}(\pi, \alpha, \boldsymbol{h})$ is the power vector in which the SIR's of all users are satisfied with equality. This corresponds to the vertex of $\mathcal{G}(\boldsymbol{h}, \boldsymbol{R})$ where the successive decoding order is π. Thus the polytopes $\mathcal{F}(\boldsymbol{h}, \boldsymbol{\alpha})$ and $\mathcal{G}(\boldsymbol{h}, \boldsymbol{R})$ have the same set of vertices, and hence must be identical.

With this identification, we can now apply the machinery developed earlier to characterize the optimal power control law to maintain the SIR requirements at all times in a fading channel, subject to transmit power constraints. We allow successive decoding at each fading state, where both the order and the powers can vary with the fading states. Using results in Section III, we see that the optimal successive decoding order at fading state $\boldsymbol{h}$ is in increasing $\frac{\lambda_m}{h_m}$, where $\boldsymbol{\lambda}$ are power prices independent of the fading state, chosen to meet the average power constraints. (Ties can be broken arbitrarily.) For independent fading processes, the boundary of the set of feasible SIR's supportable by given average power constraints $\bar{P}_i$'s consists of vectors $\boldsymbol{\alpha}$ satisfying

$$\alpha_i \int_0^\infty \frac{\sigma^2}{h_i} \prod_{k \neq i} \left\{ 1 + F_k\left(\frac{\lambda_k}{\lambda_i} h_i\right) \alpha_k \right\} f_i(h_i)\, dh_i = \bar{P}_i, \qquad i = 1, \cdots, M$$

for some power prices $\boldsymbol{\lambda}$. Finally, for a given set of SIR requirements and average transmit power constraints, the algorithm given in Section IV can be used to determine feasibility and to compute a set of appropriate power prices if feasible.

VIII. Conclusion

In this paper we have shown that any point on the delay-limited capacity region is achievable by solutions of "successive decoding type." Successive decoding is indeed optimal under the separation of time-scales assumption of Section III-D. Given a set of delay-limited rates, we have used a Lagrangian characterization of all the possible optimal power vectors to get an explicit parameterization in terms of certain "power prices." Any such optimal solution is obtained by choosing an appropriate set of power prices, and then solving a family of power control problems over a set of parallel time-invariant Gaussian multiple-access channels, one for each fading state. We have exploited the polymatroid structure of the multiaccess Gaussian capacity region to provide a simple greedy solution to each of these power control problems, despite the fact that there are an exponentially large number of constraints. It is also shown that the Lagrange multipliers associated with the power constraints (the power prices) can be computed by simple iterative procedures. We have also addressed the issues of peak power constraints, and extensions of the delay-limited concept to multiple time-scale fading processes, frequency-selective fading, and suboptimal coding schemes.

It is interesting to compare the structure of the optimal schemes for achieving throughput capacities and those for achieving delay-limited capacities. While successive decoding is optimal in both cases, the throughput-optimal schemes maintain the *same* decoding order at all fading states. However, the rates of the users are dynamically adjusted depending on the state, and indeed it is possible that a user may be allocated no rate in some states. For optimal delay-limited schemes, on the other hand, the rates are fixed at all fading states, and the successive decoding order is adjusted to maintain those target rates with the least power cost.

Appendix A
Proof of Theorem 2.3

Let $f(\boldsymbol{h})$ be the equilibrium probability density of being in fading state $\boldsymbol{h}$. Without loss of generality, assume that the fading of all users is bounded by 1. For each k, let

$$I_k = \{0, \frac{1}{k}, \frac{2}{k}, \cdots, 1\}^M$$

be a partition of the fading state space $[0,1]^M$.

First, suppose that $\boldsymbol{R}$ is in the interior of $\cap_{\boldsymbol{h}\in\mathcal{H}}\mathcal{C}_g(\boldsymbol{h}, \mathcal{P}(\boldsymbol{h}))$ for some power control policy $\mathcal{P}$. Let user i generate a random codebook of $2^{R_i T}$ codewords of length T by selecting each symbol at random from a $N(0,1)$ distribution. User i then transmits in time n, the nth symbol of the appropriate codeword, scaled by $\sqrt{\mathcal{P}_i(\boldsymbol{H}(n))}$. Such a set of codewords then satisfies the power constraint $\bar{\boldsymbol{P}}$. Given this set of codebooks, let $p(T)$ be the conditional probability of decoding any user incorrectly, using maximum-likelihood decoding, under the assumption that the decoder is given the realization

$$\boldsymbol{H} = (\boldsymbol{H}(1), \boldsymbol{H}(2), \cdots, \boldsymbol{H}(T)).$$

For S a subset of $\{1, 2, \cdots, M\}$, let $p(S,T)$ be the conditional probability of decoding any user in S incorrectly, conditional on correctly decoding the users in S^c. The union bound implies

$$p(T) \le \sum_S p(S,T).$$

As shown in Gallager [6]

$$p(S,T) \le \exp(\rho T \boldsymbol{R}(S)) \sum_{\boldsymbol{h}} f(\boldsymbol{h}) \cdot \sum_{\boldsymbol{y}} \sum_{(\boldsymbol{x}_j); j\in S^c} Q_j(\boldsymbol{x}_j \mid \boldsymbol{h}) \times \left[\sum_{(\boldsymbol{x}_i); i \in S} Q_i(\boldsymbol{x}_i \mid \boldsymbol{h}) p(\boldsymbol{y} \mid \boldsymbol{x}, \boldsymbol{h})^{1/(1+\rho)} \right]^{1+\rho}$$

for any $\rho > 0$, where $Q_i(\boldsymbol{x}_i \mid \boldsymbol{h})$ is the conditional probability density of $\boldsymbol{x}_i$ being the codeword of user i, conditional on the fading being $\boldsymbol{h}$. In our case, we obtain

$$p(S,T) \le \sum_{\boldsymbol{h}} f(\boldsymbol{h}) \cdot \exp\left(-\rho \left[-T\boldsymbol{R}(S) + \frac{1}{2} \sum_{n=1}^{T} \cdot \log \left(1 + \frac{\sum_{i\in S} \mathcal{P}_i(\boldsymbol{h}(n))}{\sigma^2 (1+\rho)} \right) \right] \right).$$

By assumption, $\exists \epsilon$ such that

$$\forall \boldsymbol{h} \in \mathcal{H} \; \boldsymbol{R}(S) \le \log \left(1 + \frac{\sum_{i\in S} \mathcal{P}_i(\boldsymbol{h})}{\sigma^2} \right) - \epsilon.$$

Thus

$$p(S,T) \le \exp\left(-\rho T (\epsilon - \log(1+\rho))\right)$$

and hence

$$p(T) \le \exp(M \ln 2 - \rho T(\epsilon - \log(1+\rho))). \tag{19}$$

By taking ρ sufficiently small, we have $\epsilon - \log(1+\rho) > 0$ and it follows that $p(T) \to 0$ as $T \uparrow \infty$. Moreover, we have in (19) a bound that decays in T at a rate *independent* of the correlation structure of the fading process. It follows that $\boldsymbol{R} \in \mathcal{C}_d(\bar{\boldsymbol{P}})$.

To prove the converse, suppose that $\boldsymbol{R}$ is an interior point of $\mathcal{C}_d(\bar{\boldsymbol{P}})$. Recall that we have partitioned the fading state space into cubes $(E_j)_{j=1}^{K^M}$. We consider a sequence of Markov processes defined on $\mathcal{H}$ of the following form. Consider a Markov chain on the "coarse" states E_j with transition probabilities $t(E_j, E_k)$. We use such a chain to define a Markov process on $\mathcal{H}$: conditional on the chain being in coarse state E, we select a fading state for the process by using the stationary distribution conditional on the fading being in E. The process remains in this state for an exponential time $\tau(E) \equiv$ Exponential$(\lambda(E))$ and then selects a new coarse state according to t. We assume that the Markov process has the required stationary distribution on $\mathcal{H}$, by choosing appropriate $(\lambda(E))_{E\in I_k}$. By scaling all $\lambda(E_j)$ by a constant, we can speed up or slow down the rate of fading while retaining the required stationary distribution.

For each $T = 1, 2, \cdots$, let $\boldsymbol{H}^{(T)}$ be such a fading process with the following properties. We assume a random variable $\boldsymbol{H}(0)$ on $\mathcal{H}$ with the stationary distribution of the processes we require. We assume all fading processes start with $\boldsymbol{H}^{(T)}(0) \equiv \boldsymbol{H}(0)$, $T = 1, 2, \cdots$. The initial sojourn time in state $\boldsymbol{H}(0)$ of fading $\boldsymbol{H}^{(T)}$ is given by $\tau_T(\boldsymbol{H}(0))$, where $\tau_T(E_j) \asymp$ Exponential$(r_T \lambda(Ej))$ and independent of $\boldsymbol{H}(0)$ for all j. The constant r_T gives the "rate of fading" for process $\boldsymbol{H}^{(T)}$. Let δ be a fixed, positive constant. By choosing an appropriate decreasing sequence $(r_T)_{T=1}^{\infty}$, $r_T \downarrow 0$, we can ensure that for all j

$$P(\forall T, \tau_T(E_j) > T) > 1 - \delta. \tag{20}$$

Since $R \in \mathcal{C}_d(\bar{\boldsymbol{P}})$, we can choose for each T and each user i a code of size $2^{R_i T}$. A codeword from user i's codebook consists of T symbols. Let $\boldsymbol{X}^{(T)}(n) n = 1, 2, \cdots, T$ denote a random, independent selection of codewords for the users, for which the probability of error in channel $\boldsymbol{H}^{(T)}$ goes to zero with T. Let $p(T)$ be the probability of error for $\boldsymbol{X}^{(T)}$ under fading $\boldsymbol{H}^{(T)}$. We note that $\boldsymbol{X}^{(T)}$ may be random; say, with dependence on $\boldsymbol{H}^{(T)}$, although we do not require this. Let $\Omega(E)$ be the subset of the sample space on which $\boldsymbol{H}(0) \in E$ and $\forall T, \tau_T(E) > T$. Let Q be uniform on $[0,T]$, and independent of all other variables. Define

$$\begin{aligned} V_i(E,T) &\equiv \mathbb{E}\left[\left(X_i^{(T)}\right)^2(Q) \mid \Omega(E)\right] \\ W_i(E,T) &\equiv \mathbb{E}\left[\left(X_i^{(T)}\right)^2(Q) \mid [\boldsymbol{H}(0) \in E] - \Omega(E)\right] \\ \boldsymbol{Z}(E,T) &\equiv \boldsymbol{V}(E,T) P(\forall T, \tau_T(E) > T \mid \boldsymbol{H}(0) \in E] \\ &\quad + \boldsymbol{W}(E,T) P(\exists T : \tau_T(E) \le T \mid \boldsymbol{H}(0) \in E). \end{aligned}$$

Then the power constraint is that $\forall T$

$$\sum_E f(E) \boldsymbol{Z}(E,T) \le \bar{\boldsymbol{P}}.$$

By Assumption (20), we have

$$\sum_E f(E) \boldsymbol{V}(E,T) \le \frac{\bar{\boldsymbol{P}}}{1-\delta}.$$

Taking limits along a convergent subsequence, we have that

$$\boldsymbol{V}(E,T) \to \boldsymbol{V}(E) \quad \text{as } T \uparrow \infty$$
$$\sum_E f(E)\boldsymbol{V}(E) \le \frac{\bar{\boldsymbol{P}}}{1-\delta}.$$

Now let us define a new fading process $\boldsymbol{H}$ by

$$\boldsymbol{H}(n) \equiv \sum_E I[\boldsymbol{H}(0) \in E]\boldsymbol{h}(E)$$

where $\boldsymbol{h}(E)$ is the upper corner of the cube E. Note that conditional on $\boldsymbol{H}(0)$, the fading process is deterministic. Let $q(T \mid \Omega(E))$ be the conditional probability of error for code $X^{(T)}$ in this new fading channel, conditional on the event $\Omega(E)$. By construction

$$p(T) \ge P(\Omega(E))q(T \mid \Omega(E)).$$

By assumption, $p(T) \to 0$ as $T \uparrow \infty$, and hence

$$q(T \mid \Omega(E)) \to 0, \qquad \text{as } T \uparrow \infty.$$

But conditional on $\Omega(E)$, we have a a constant fading channel, and a sequence of codes satisfying the power constraint $\boldsymbol{V}(E)$. It follows that for all $E \in I_k$, $\boldsymbol{R} \in \mathcal{C}_g(\boldsymbol{h}(E), \boldsymbol{V}(E))$.

Define $\mathcal{F}_{k,\delta}$ to be those power-control policies that satisfy the power constraint $\frac{\bar{\boldsymbol{P}}}{1-\delta}$ and are piecewise-constant on each cubic element in I_k. Set

$$\mathcal{P}_{k,\delta}(\boldsymbol{h}) \equiv \sum_{E \in I_k} \boldsymbol{V}(E)I[\boldsymbol{h} \in E]$$

and note that $\mathcal{P}_{k,\delta} \in \mathcal{F}_{k,\delta}$. For any power control $\mathcal{P}$, define

$$\bar{\mathcal{C}}_g^{(k)}(\boldsymbol{h}, \mathcal{P}(\boldsymbol{h})) \equiv \{\boldsymbol{R} : \boldsymbol{R} \in \mathcal{C}_g(\bar{\boldsymbol{h}}^{(k)}, \mathcal{P}(\boldsymbol{h}))\}$$

where $\bar{\boldsymbol{h}}_i^{(k)} = \frac{1}{k}\lceil k\boldsymbol{h}_i \rceil$ for all $i = 1, 2, \cdots, M$. We have shown that for any $\delta > 0$

$$\forall \boldsymbol{h} \in \mathcal{H}, \qquad \boldsymbol{R} \in \bar{\mathcal{C}}_g^{(k)}(\boldsymbol{h}, \mathcal{P}_{k,\delta}(\boldsymbol{h})).$$

It follows that

$$\boldsymbol{R} \in \cup_{\mathcal{P} \in \mathcal{F}_k} \cap_{\boldsymbol{h} \in \mathcal{H}} \bar{\mathcal{C}}_g^{(k)}(\boldsymbol{h}, \mathcal{P}(\boldsymbol{h})) \tag{21}$$

where $\mathcal{F}_k \equiv \mathcal{F}_{k,0}$. Now by the first part of the proof, we have that

$$\cup_{\mathcal{P} \in \mathcal{F}_k} \cap_{\boldsymbol{h} \in \mathcal{H}} \mathcal{C}_g(\boldsymbol{h}, \mathcal{P}(\boldsymbol{h})) \subseteq \cup_{\mathcal{P} \in \mathcal{F}} \cap_{\boldsymbol{h} \in \mathcal{H}} \mathcal{C}_g(\boldsymbol{h}, \mathcal{P}(\boldsymbol{h})) \subseteq \mathcal{C}_d(\bar{\boldsymbol{P}}).$$

We have shown in (21) that

$$\mathcal{C}_d(\bar{\boldsymbol{P}}) \subseteq \cup_{\mathcal{P} \in \mathcal{F}_k} \cap_{\boldsymbol{h} \in \mathcal{H}} \bar{\mathcal{C}}_g^{(k)}(\boldsymbol{h}, \mathcal{P}(\boldsymbol{h})).$$

But these lower and upper bounds converge as $k \uparrow \infty$, and hence

$$\mathcal{C}_d(\bar{\boldsymbol{P}}) = \cup_{\mathcal{P} \in \mathcal{F}} \cap_{\boldsymbol{h} \in \mathcal{H}} \mathcal{C}_g(\boldsymbol{h}, \mathcal{P}(\boldsymbol{h})).$$

This completes the proof.

APPENDIX B
PARAMETERIZATION OF BOUNDARY OF CAPACITY REGION

One unsatisfactory feature of Section III is that we are unable to provide an explicit parameterization of the boundary surface of $\mathcal{C}_d(\bar{\boldsymbol{P}})$. Theorem 3.3 suggests a parameterization of the boundary surface by $\boldsymbol{\lambda} \in \Re_+^M$, and we discuss this further below.

The following lemma shows that for any $\boldsymbol{\lambda} \in \Re_+^M$ there is at least one $\boldsymbol{R} \in \Re_+^M$ such that $(\boldsymbol{\lambda}, \boldsymbol{R})$ solves (9).

Lemma B.1: Define $x_i = \exp(2R_i) - 1$, and the transformation T by

$$T_i(\boldsymbol{x}) \equiv \frac{P_i}{\int_0^\infty \frac{\sigma^2}{h} \prod_{k \ne i} \left(1 + F_k\left(\frac{\lambda_k h}{\lambda_i}\right)x_k\right) f_i(h)\, dh}.$$

Then there exists a fixed point of T.

Proof: T is continuous, and

$$0 \le T_i(\boldsymbol{x}) \le \frac{P_i}{\sigma^2 \int_0^\infty \frac{f_i(h)}{h} dh}$$

and hence T is a mapping from

$$\prod_i \left[0, \frac{P_i}{\sigma^2 \int_0^\infty \frac{f_i(h)}{h} dh}\right]$$

to itself. By the Brouwer fixed-point theorem, there exists a fixed point for T in this set. □

It follows that (9) has a solution in $\boldsymbol{R}$ for any positive $\boldsymbol{\lambda}$. Even if a closed-form parameterization of $\mathcal{C}_d(\bar{\boldsymbol{P}})$ is not possible, it would be useful to have a computational procedure to find a solution to (9). Consider then the following algorithm, which we might use to try and find such a solution:

$$x \to T^n(x). \tag{22}$$

where x is the starting point of the algorithm, and $T^n(x)$ is the nth iterate. It is easy to show that T^2 satisfies the monotonicity property of Section IV in Part I. Thus if T has a unique fixed point then $T^n(x)$ will converge to it from any starting point x. We leave the problem of establishing the uniqueness of the fixed point of T open. It is equivalent to the following conjecture.

Conjecture 1:

1) Given $\bar{\boldsymbol{P}}$, the mapping $\boldsymbol{\mu} \to \boldsymbol{\lambda}(\bar{\boldsymbol{P}}, \boldsymbol{\mu})$ is invertible, implying that we can parameterize the boundary surface of $\mathcal{C}_d(\bar{\boldsymbol{P}})$ by $\boldsymbol{\lambda} \in \Re_+^M$ (note, we assume that μ and λ are appropriately normalized, so that these mappings are well-defined).
2) Given $\boldsymbol{R}^*$, the mapping $\boldsymbol{\lambda} \to \boldsymbol{\mu}(\boldsymbol{R}^*, \boldsymbol{\lambda})$ is invertible, implying that we can parameterize the boundary surface of $\mathcal{D}_d(\boldsymbol{R}^*)$ by $\boldsymbol{\mu} \in \Re_+^M$.

We also conjecture that the analogous results hold in Part I; that is, the maps $\boldsymbol{\lambda}(\bar{\boldsymbol{P}}, \boldsymbol{\mu})$ and $\boldsymbol{\mu}(\boldsymbol{R}^*, \boldsymbol{\lambda})$ are invertible in the throughput capacity case as well.

REFERENCES

[1] D. P. Bertsekas and R. G. Gallager, *Data Networks*, 2nd ed. Englewood Cliffs, NJ: Prentice Hall, 1992.

[2] G. Caire, G. Taricco, and E. Biglieri, "Power control for the fading channel: An information theoretic approach," in *Proc. 35th Allerton Conf.* (Monticello, IL, Sept. 1997).

[3] R. Cheng, "On capacity and signal waveform design for slowly fading CDMA channels," in *Proc. 31st Allerton Conf.* (Monticello, IL, 1993), pp. 11–20.

[4] J. Edmonds, "Submodular functions, matroids and certain polyhedra," in *Proc. Calgary Int. Conf. Combinatorial Structures and Applications* (Calgary, Alta, June 1969), pp. 69–87.

[5] A. Federgruen and H. Groenevelt, "The greedy procedure for resource allocation problems: Necessary and sufficient conditions for optimality," *Oper. Res.* , vol. 34, pp. 909–918, 1986.

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

[7] K. S. Gilhousen, I. M. Jacobs, R. Padovani, A. J. Viterbi, L. A. Weaver, and C. E. Wheatley, "On the capacity of a cellular CDMA system," *IEEE Trans. Veh. Technol.*, vol. 40, pp. 303–312, May 1991.

[8] A. Goldsmith and P. Varaiya, "Capacity of fading channel with channel side information," *IEEE Trans. Inform. Theory*, vol. 43, pp. 1986–1992, Nov. 1995.

[9] S. V. Hanly, "An algorithm for combined cell-site selection and power control to maximize cellular spread spectrum capacity," *IEEE J. Select. Areas Commun.* (Special Issue on the Fundamentals of Networking), vol. 13, Sept. 1995.

[10] S. V. Hanly and D. Tse, "Multi-access fading channels: Shannon and delay-limited capacities," in *Proc. 33rd Allerton Conf.* (Monticello, IL, Oct. 1995).

[11] S. V. Hanly and P. A. Whiting, "Constraints on capacity in a multi-user channel," presented at the International Symposium on Information Theory, Trondheim, Norway, 1994.

[12] L. H. Ozarow, S. Shamai, and A. D. Wyner, "Information-theoretic considerations for cellular mobile radio," *IEEE Trans. Veh. Technol.*, vol. 43, pp. 359–378, May 1994.

[13] 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. 1995.

[14] J. G. Shantikumar and D. D. Yao, "Multiclass queueing systems: Polymatroid structure and optimal scheduling control," *Oper. Res.*, vol. 40, no. 2, pp. 293–299, 1992.

[15] D. Tse and S. V. Hanly, "Multiaccess fading channels–Part I: Polymatroid structure, optimal resource allocation, and throughput capacities," this issue, pp. 2796–2815.

[16] R. Yates, "A framework for uplink power control in cellular radio systems," *IEEE J. Select. Areas Commun.* (Special Issue on the Fundamentals of Networking), vol. 13, Sept. 1995.