Decentralized Rate Assignments in a Multi-Sector CDMA Network

Jennifer Price, *Student Member, IEEE,* and Tara Javidi, *Member, IEEE*

***Abstract*— As the demand for wireless broadband data services grows, it becomes increasingly important to address the issue of optimal resource allocation. Specifically, such allocation should address not only Quality of Service (QOS) requirements, but continually changing resource demands. In this paper, we consider a data-only multi-sector wideband CDMA network with arbitrary layout of sectors and base-stations. Optimal reverse-link rate allocation for such a network is examined. Using an optimization framework, we develop the conceptual basis for a distributed algorithm for finding the optimal rate assignments using information available to the mobiles. We then use this to propose practical algorithms at the bases as well as the mobiles. These algorithms provide a proportional fair transmission rate/interference control scheme. Finally, we examine the dynamic behavior of such a distributed algorithm in the presence of realistic physical conditions such as shadowing and mobility.**

***Index Terms*— CDMA networks, distributed optimization, rate control.**

I. Introduction

As the applications of wireless networks (such as cellular systems) spread to various fields, the demand for wireless broadband data services increases. Due to the inherent differences between the statistical behavior and Quality of Service (QoS) requirements of a data connection with those of a voice user, new challenges and problems in the design of future generations of wireless systems arise. More specifically, in such systems where throughput and fairness are essential elements of QoS provisioning, the need to revisit the question of optimal resource allocation becomes crucial.

In this paper we address the issue of decentralized reverse-link rate assignment in a data optimized (DO) multi-sector wideband CDMA network with variable rates. This is similar to the recently deployed commercial CDMA2000 1xEV-DO systems, which use a distributed feedback mechanism for the reverse-link rate assignments. It is our goal to address the issue of optimal rate assignments within the existing confines of CDMA2000 1xEV-DO. To the best of our knowledge, this work represents the first attempt to create and analyze an optimal, non-uniform and decentralized rate assignment scheme within this structure.

Manuscript received September 22, 2004; revised January 11, 2006; accepted May 31, 2006. The associate editor coordinating the review of this paper and approving it for publication was Q. Zhang. This work was supported in part by the NSF ADVANCE Cooperative Agreement No. SBE-0123552, and in part by the NSF CAREER Award No. CNS-0347961.

The authors are with the Department of Electrical and Computer Engineering, University of California/San Diego, La Jolla, CA 92093 USA (email: jenn-price@ucsd.edu; tara@ece.ucsd.edu).

Digital Object Identifier 10.1109/TWC.2006.04625.

The remainder of the paper is organized as follows. Section II provides a detailed discussion of the motivation for this work including an overview of the related literature, a description of the Reverse Traffic MAC Protocol for CDMA2000 1xEV-DO, and the notation and CDMA interference model used throughout this paper. Section III investigates the existence of feasible rate allocations in power controlled cellular systems. Section IV describes an optimal rate assignment problem, and provides an in-depth discussion on the notion of optimality and the inherent trade-off between throughput and fairness. In Section V, we present a set of distributed algorithms that, under the assumption of time-invariance of the network layout and channel quality, converge to the solution of the optimal rate assignment. The formulation and implementation of these algorithms are strongly motivated by the feedback structure of CDMA2000 1xEV-DO. Section VI provides numerical examples and simulation results, including algorithm performance under dynamically varying conditions. Finally, Section VII has our conclusions and areas of future work.

II. Background and Motivation

In data-only networks, most traffic is elastic; that is, it can tolerate variable transmission rates and delay [1], [2]. This attribute is particularly useful in a broadband wireless environment where the medium is a shared, limited resource. In such an environment, it is of extreme importance for data networks to have characteristics similar to modern IP networks. In particular, it is important for such networks to respond to randomly fluctuating demands and failures by continually adapting rates in a scalable and decentralized manner, so as to maintain efficient channel utilization and fairness amongst users.

In order to design such a rate-assignment scheme, we draw heavily from the well-known works by F. Kelly [1], [3], S. Low [4], [5], and Mo and Walrand [6]. In these works, the network is modeled as a collection of individual users and network components who are in constant negotiation with one another. When carefully designed, the interactions between users and network components result in a distributed network-wide optimal solution.

Using this same idea, we model the rate-assignment for reverse-link CDMA as a global optimization problem over some set of feasible rate assignments. Understanding the role of power control and its effects are crucial in determining a constraint describing such a feasibility region. Although power control in voice-based cellular systems has received a good deal of attention (see [7], [8], [9], [10]), its impact on the variable-rate assignment problem has only recently begun to receive attention. Our study of feasible rate allocations is strongly motivated by the work in [11], although that work is done in the context of single-sector while we focus on multi-sector CDMA.

In this work, we show that under the assumption of time-invariance of network layout and channel quality and subject to a linear feasibility constraint, finding an optimal rate allocation reduces to a general utility optimization framework. Capitalizing on an economic interpretation of the problem, we propose a distributed algorithm at the mobiles and bases, whose equilibrium point coincides with the desired solution. In other words, we propose two sets of algorithms: 1) the transmission rate control algorithm implemented at each mobile to react to varying levels of interference, and 2) the base-stations' interference management algorithm which measures the interference and produces interference indicator signals. The rate control algorithm at the mobiles uses a combination of its desirable QoS (rate utility) and interference indication signals from the bases to self-tune its data flow. Conceptually, this is the same feedback structure used in the CDMA2000 1xEV-DO 3GPP2 standard, which is described in Section II-A below.

It should be noted that there is a myriad of work on the optimal resource allocation problem for wireless networks, covering everything from power control, to quality of service, to cross-layer design. The authors in [12] and [13] examine the rate-assignment problem for both single and multi-sector CDMA systems, but leave the issue of decentralized non-uniform rate-assignment unaddressed. The authors in [14] propose algorithms that jointly optimize both transmit power and spreading-gain (i.e. rate). The focus in these papers, however, is a data-voice (DV) system dealing with both real-time and non-real-time traffic.

As we have shown, there is a rich literature on optimal resource allocation in wireless networks. Our work, however, represents the first attempt to formulate and analyze the optimal resource allocation problem in terms of existing technologies, namely CDMA2000 1xEV-DO. Although all of the works above provide valuable insight into the resource allocation problem, none of them address the problem within the constraints of existing technologies. As such, it is our belief that this work provides valuable insight not only to the wireless resource allocation problem, but to the continuing development and improvement of deployable technologies.

A. Feedback Structure of CDMA2000 1xEV-DO Revision A

In order to motivate the modeling and design choices made throughout this paper, we provide here a macroscopic-level description of a reverse-link protocol outlined in the CDMA2000 1xEV-DO Revision A (hereafter referred to as CDMA2000) standards. For a more detailed description of this protocol, refer to the section on Subtype 3 Reverse Traffic Channel MAC Protocol in [15].

The network consists of two main components: the access terminals (also referred to as mobiles), and the access network (which is comprised of sectors/base stations). Each access terminal is responsible for choosing a transmission power and

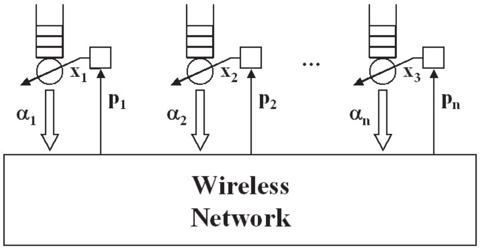

Fig. 1. Distributed feedback structure for rate assignments.

packet-size with which to transmit its data. This choice is based, in part, on feedback from the access network. The goal of the access network is to ensure the decodability of the CDMA signals. In other words, the terminals and network form a distributed feedback loop following the structure of Figure 1 where, as we will see later, α represents the choice of transmission power and packet size at an access terminal, while p represents the feedback from the access network.

In order to guarantee the decodability of the CDMA signals, the access networks has two goals: 1) ensure that the received pilot power from each terminal associated with a given sector is equalized at that sector, and 2) ensure that the total received power at each sector is below a pre-specified threshold. In order to meet the first goal, each sector transmits power-control signals to the terminals for whom it is the forward-link-serving sector. Based on the actual and desired values of received pilot power at the sector, these power-control signals simply indicate whether the terminal needs to increase or decrease its pilot power. In order to meet the second goal, each sector transmits a *Reverse Activity Bit* (RAB). This reverse activity bit simply indicates whether or not the desired total received power threshold has been violated at that sector.

Each access terminal has three tasks: 1) adjust its pilot power in response to the power-control signals from its forward-link-serving sector, 2) choose a transmission power based on the reverse activity bits from the access network, and 3) choose a packet-size based on the choice of transmission power. The transmission power is always chosen proportional to the pilot power. The ratio of transmission power to pilot power is referred to as $T2P$, and each value of $T2P$ corresponds to an allowable packet size. Thus, the function of the access terminal boils down to choosing its $T2P$ value based on the value of the RAB's from the access network.

Recall that the decodability of the CDMA signal is based on two factors: 1) the received pilot power from each terminal associated with a given sector is equalized at that sector, and 2) the total received power at each sector is below a pre-specified threshold. Since access terminals choose their transmission power proportional to their pilot power, these two conditions essentially guarantee an acceptable BER for all transmissions (a more detailed discussion of this can be found in Sections II-B and III below). It is worth noting that, in general, this may not be true for large packet sizes which require a higher pilot power for decodability. In order to account for this, the reverse-link channel does include an auxiliary pilot signal.

B. Interference Model for CDMA Systems

Throughout this paper, we use the following notations. There are a total of N mobiles with chip bandwidth W. Each node has pilot power P_i^0 and pilot rate R_b. In addition, each node transmits data traffic with rate x_i (hereafter referred to as MAC rate) and transmit power P_i. The spreading gain for user i is defined as $s_i = \frac{W}{x_i}$, while the ratio of user i's transmit power to its pilot power is α_i. Note that α_i is exactly the $T2P$ parameter in CDMA2000, and $P_i = \alpha_i P_i^0$.

There are a total of L sectors in the system, where one or more sectors may be associated with a particular base station. The *tracking sector* for mobile i, denoted $b(i)$, is the sector associated with the mobile's forward link, and whose base station transmits power control signals to the mobile. M_l, $l = 1, \ldots, L$, is the set of mobiles which are being tracked by sector l, i.e. $i \in M_l$ iff $l = b(i)$. M_l, $l = 1, \ldots, L$ are assumed to be mutually disjoint. The channel power gain from mobile i to sector l is denoted by g_{il} and incorporates both path gain and antenna gain. We further assume that if $i \in M_l$, then $g_{il} > \epsilon$.

Now, consider mobile i which is tracked by sector $l = b(i)$. We define the ratio of pilot chip energy to interference power as:

$$\left(\frac{E_{CP}}{N_t}\right)^l (i) = \frac{\frac{W}{R_b} P_i^0 g_{il}}{Y + N_0 W + \sum_{j \neq i} P_j g_{jl}} \tag{1}$$

where N_0 is the thermal noise density, and Y is the total received power from overhead channels (including auxiliary pilot). Since Y is fixed and below the traffic power level, we can neglect this term in our analysis. The signal-to-interference ratio of mobile i's pilot signal at base station l can then be written as $PSINR^l(i) = \frac{E_{CP}}{N_t}^l (i) (\frac{R_b}{W})$.

Similarly, we can define the ratio of transmit energy per chip to interference power as:

$$\left(\frac{E_b}{N_t}\right)^l (i) = \frac{\frac{W}{x_i} P_i g_{il}}{Y + N_0 W + \sum_{j \neq i} P_j g_{jl}} \tag{2}$$

and the signal-to-interference ratio of mobile i's data signal at base station l to be $SINR^l(i) = \frac{E_b}{N_t}^l (i) (\frac{x_i}{W})$.

Notice that in an interference limited system such as CDMA, the relationship between SINR and information rate given by the Shannon equation can be approximated by a linear one (i.e. $log(1+q) \approx q$) [11]. This means that an increase in MAC rate translates into an increase in information rate if and only if $\frac{E_b}{N_t}$ is kept the same. It then follows from Eqn (2) that an increase in the MAC rate of a mobile must be matched by an equal increase in power. The consequence of this is that the parameter α_i is not only the ratio of transmission power to pilot power, but also approximates the ratio of MAC rate to pilot rate! (Because of this, we refer to the parameter α_i as user i's rate for the remainder of this paper).

Finally, we introduce a measure of total received power.

Definition 1: The *rise over thermal* (ROT) is defined as

$$Z_l = \frac{\sum_{i=1}^N P_i g_{il}}{N_0 W} \tag{3}$$

and indicates the ratio between the total power received from all the mobiles at the base station l and the thermal noise [15].

III. Feasible Rate-Power Pairs

In a wireless system, the definition of what constitutes a feasible rate-power pair is highly dependent upon both the application and design of a particular system. Constraints may include minimum or maximum power constraints, interference limits, minimum rate guarantees, QoS metrics, or maximum delay requirements. Since the focus of this work is a DO network, we choose to focus on BER, as determined by $\frac{E_b}{N_t}$ (similar to CDMA2000). Thus, we say a vector of rates $(\alpha_1, \ldots, \alpha_N)$ is a feasible solution if there exists a feasible vector of transmission powers $(P_1, \ldots, P_N)$ that guarantees an acceptable BER for each user, i.e. that satisfies the condition

$$\frac{s_i P_i g_{il}}{N_0 W + \sum_{j \neq i} P_j g_{jl}} \geq \gamma \ \forall i \in M_l \ and \ \forall l \tag{4}$$

where γ is the pre-defined target $\frac{E_b}{N_t}$ value based on the desired BER. Since we assume a linear relationship between rate and power, this is equivalent to satisfying the condition $\frac{E_{CP}}{N_t} \geq \gamma$ for all users. Again, this is similar to the specifications in CDMA2000, where the allowable values of $T2P$ are designed to maintain a constant value of $\frac{E_b}{N_t}$ across all packet sizes.

We are still left with the problem of what constitutes a feasible power vector. Again, this depends on the application and also the design of the system. In this paper, we consider three types of feasible power vectors: an instantaneous transmission power limitation, a total received power limitation, and a linearized received power limitation. Our goal is to use these definitions of feasible power vectors, combined with considering the condition from Eqn (4) as an equality, to construct rate feasibility regions which are independent of power. In Appendix I, we show that this results in the following equation:

$$\underline{P} = \gamma N_0 W \mathbf{R} (I - \gamma \mathbf{G}\mathbf{R})^{-1} \underline{\mathbf{1}}_L \tag{5}$$

where $\mathbf{G}$ and $\mathbf{R}$ are defined below. To obtain the above closed form, we follow the work in [11] and look at an auxiliary variable we call effective rate:

Definition 2: For each user i, we define the quantity *effective rate* to be $r_i = \frac{1}{\gamma + s_i} = \frac{R_b \alpha_i}{\gamma R_b \alpha_i + W}$.

Effective rate is a unitless quantity that is used merely for tractability and simplicity of equations introduced later. Similarly, we introduce matrix $\mathbf{R}$ to allow for a closed form matrix relationship between power and rate.

Definition 3: We define the *normalized rate* of user i tracked by sector k to be given as $R_{ik} = \frac{r_i \Psi_{ik}}{g_{ik}}$, where Ψ_{ik} is an indicator function for $k = b(i)$. The matrix of normalized rates is defined as $\mathbf{R} = [R_{ik}]$, with dimension $N \times L$.

Define the gain matrix $\mathbf{G}$ of dimension $L \times N$ such that $G_{li} = g_{il}$. Now, consider the $L \times L$ matrix $\mathbf{GR}$. Assuming perfect power control, the element lk of matrix $\mathbf{GR} = [\rho_{lk}] = [\sum_{i \in M_k} \frac{g_{il}}{g_{ik}} r_i]$ represents the total effective rate of sector k as it is perceived at the base l. We note that this is closely related to the work in [7], [16] when all rates are equal. As we will see in the following sections, the existence of a non-negative power vector that satisfies Eqn (4) with equality is dependent upon the invertibility of the matrix $(I - \gamma \mathbf{GR})$. Thus, we introduce the following assumption and proposition:

Assumption 1: $\rho(\gamma\mathbf{GR}) < 1$, where $\rho(\cdot)$ is the spectral radius.

Proposition 1: The matrix $(I - \gamma\mathbf{GR})$ is invertible (with an inverse which is non-negative component-wise) if and only if Assumption 1 holds.

Proof:

First, assume that $\rho(\gamma\mathbf{GR}) < 1$. This implies that there exists a matrix norm $\|\cdot\|$ such that $\|\gamma\mathbf{GR}\| < 1$ (see [17], Lemma 5.6.10, page 297). Now we use the following fact: a matrix A is invertible if there exists a norm $\|\cdot\|$ such that $\|I - A\| < 1$, with inverse given by $A^{-1} = \sum_{k=0}^{\infty}(I-A)^k$ (see [17], Corollary 5.6.16, page 301). This implies that the matrix $(I - \gamma\mathbf{GR})$ is invertible, and that the inverse is non-negative component-wise.

Now, assume that the matrix $(I - \gamma\mathbf{GR})$ is invertible, with inverse given by $(I - \gamma\mathbf{GR})^{-1} = \sum_{k=0}^{\infty}(\gamma\mathbf{GR})^k$. This implies that $lim_{k->\infty}(\gamma\mathbf{GR})^k = 0$, which in turn implies that $\rho(\gamma\mathbf{GR})^k < 1$ (see [17], Theorem 5.6.12, page 298). ■

Now, using the above notions, we seek to translate three different feasibility conditions on the power vectors into rate feasibility regions. Note that in considering Eqn (4) as an equality, we assume a power control mechanism that guarantees that $\frac{E_{CP}}{N_t}$, and hence $\frac{E_b}{N_t}$, is equal to γ. (This is the same concept as the power control mechanism used in CDMA2000, where received pilot powers are equalized at the forward-link-serving sector).

A. Power-Limited Feasibilty Region: A Maximum Transmission Power Constraint

The most common condition for feasibility is to limit the transmission power of each mobile i to $P_i^{\max}$ at all times. Thus, we introduce our first feasibility region:

Definition 4: A vector of rates $\underline{\alpha} = (\alpha_1, \ldots, \alpha_N)$ is said to be power-limited feasible if there exists a power vector $(P_1, \ldots, P_N)$ that satisfies the following conditions:

C1. $\frac{s_i P_i g_{ib(i)}}{N_0 W + \sum_{j\neq i} P_j g_{jb(i)}} = \gamma \ \forall i$

C2. $0 \leq P_i \leq P_i^{\max} \ \forall i$

Theorem 1: Under Assumption 1, a vector of rates $\underline{\alpha} = (\alpha_1, \ldots, \alpha_N)$ is a *power-limited feasible rate assignment* iff the vector of effective rates $\underline{\mathbf{r}} = (r_1, \ldots, r_N)$ satisfies the following alternative condition:

AC1. $\mathbf{R}(I - \gamma\mathbf{GR})^{-1}\underline{\mathbf{1}}_L \leq \frac{1}{\gamma N_0 W}\underline{\mathbf{P}}^{\max}$,

where $\underline{\mathbf{1}}_M$ is a vector of dimension M whose elements are all 1, and $\underline{\mathbf{P}}^{\max}$ is a vector of size N with elements $P_i^{\max}$. We denote by Δ_P the *power-limited feasibility region*, i.e. the set of all rate vectors $\underline{\alpha}$ that satisfy AC1.

Proof: See Appendix I. ■

B. ROT-Controlled Feasibility Region: A Maximum Interference Constraint

In this section we consider an alternative feasibility condition that limits the total received power at each base-station (similar to the CDMA2000 standards). Recall the definition of rise over thermal given by Eqn (3) in Section II-B. Using this, we define our second feasibility region:

Definition 5: A vector of rates $\underline{\alpha} = (\alpha_1, \ldots, \alpha_N)$ is said to be ROT-controlled feasible if there exists a power vector $(P_1, \ldots, P_N)$ that satisfies the following conditions:

C1. $\frac{s_i P_i g_{ib(i)}}{N_0 W + \sum_{j\neq i} P_j g_{jb(i)}} = \gamma \ \forall i$

C3. $Z_l \leq K$, where K is a fixed value.

Theorem 2: Under Assumption 1, a vector of rates $\underline{\alpha} = (\alpha_1, \ldots, \alpha_N)$ is an *ROT-controlled feasible rate assignment* iff the vector of effective rates $\underline{\mathbf{r}}$ satisfies

AC2. $(I - \gamma\mathbf{GR})^{-1}.\underline{\mathbf{1}}_L \leq (1 + K)\underline{\mathbf{1}}_L$

We denote by Δ_Z the *ROT-controlled feasibility region*, i.e. the set of all rate vectors $\underline{\alpha}$ which satisfy Condition AC2.

Proof: See Appendix I. ■

Remark: Note that if $K \leq \frac{\epsilon P_i^{\max}(s_i^{min} + \gamma)}{\gamma N_0 W}$ for $\forall i$, then Condition C3 implies Condition C2.

We believe that the advantage of limiting ROT rather than individual power is three fold: 1) it provides a significant reduction in the number of constraints, 2) it guarantees that the variation in instantaneous transmitted power for each user is small, and 3) it provides a higher level of robustness to random channel variations such as fading and shadowing. The first advantage is clear, since limiting the total power results in a reduction in the number of constraints from N to L. To see the second and third advantages, we combine Conditions C1 and C3. This yields:

$$P_i = \frac{\gamma N_0 W r_i (1 + Z_{b(i)})}{g_{ib(i)}}. \tag{6}$$

or, in matrix form,

$$\underline{P} = \gamma N_0 W \mathbf{R}(\underline{1_L} + \underline{Z}) \tag{7}$$

Now, the second advantage is a direct result of Eqn (7). The third advantage can also be intuitively explained using Eqn (7): variations in gains (manifested by perturbations in $\mathbf{R}$) cause variations in the required power which are magnified by $(\underline{1_L} + \underline{Z})$. As a result, limiting Z_l by some value K increases the robustness of the power control mechanisms. We will return to this idea in Section VI.

C. Linearized Feasibility Region: A Simplex Subset of the ROT-Controlled Region

Searching for a rate vector over the above feasibility regions requires computationally expensive matrix operations. Our goal in this section is to propose linear constraints on effective rates (as seen later, allowing for a decentralized solution). Conceptually, this is motivated by the notion of effective bandwidths (see [18], [19], [20]).

Theorem 3: A vector of rates $\underline{\alpha} = (\alpha_1, \ldots, \alpha_N)$ is an ROT-controlled feasible rate assignment if the vector of effective rates $\underline{\mathbf{r}}$ satisfies

LC1. $\sum_{i=1}^{N} \frac{g_{il}}{g_{ib(i)}} r_i \leq \frac{K}{\gamma(1+K)} \ \forall\ l = 1, \ldots, L$

We denote by Δ_L the *linearized feasibility region*, i.e. the set of all rate vectors $\underline{\alpha}$ which satisfy Condition LC1.

Proof: See Appendix II. ■

We note that these linear constraints guarantee ROT feasibility but result in a strictly smaller feasibility region. This, in general, results in a loss of optimality.

Remark: Notice that Condition LC1 (which can be written as $\|\gamma\mathbf{GR}\|_\infty \leq \frac{K}{(1+K)}$) guarantees the validity of Assumption 1, since $\rho(\gamma\mathbf{GR}) \leq \|\gamma\mathbf{GR}\|_\infty$.

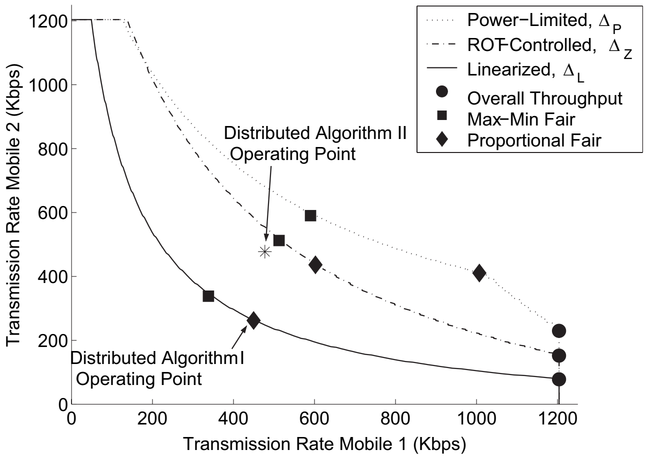

Fig. 2. Power limited, ROT-controlled, and linearized feasibility regions and optimal points of transmission rates.

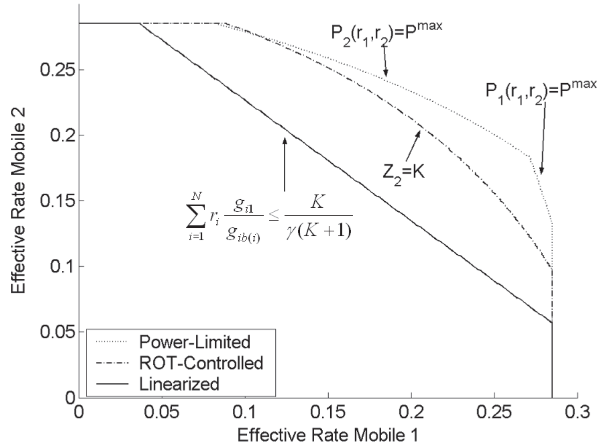

Fig. 3. Power limited, ROT-controlled, and linearized feasibility regions of effective rates.

Figure 2 shows the power-limited (Δ_P), ROT-controlled (Δ_Z), and linearized (Δ_L) feasibility regions of rates for a simple scenario involving 2 bases placed 2500 m apart, and 2 randomly positioned mobiles - one of which is assigned to each base. In order to show a meaningful feasibility region, we ignore any antenna gain factors for this scenario. Figure 2 shows the feasibility regions Δ_P, Δ_Z, and Δ_L. Figure 3 shows the same feasibility regions, but in terms of effective (rather than actual) transmission rates. We see here the reason for the introduction of effective rates - the non-convex regions shown in Figure 2 become convex as a function of effective rate. Additionally, in Figure 3 we identify the active constraint associated with each portion of the feasibility region boundaries.

IV. Rate Assignment: Overall Throughput versus Fairness

In this section we introduce the notion of "optimal" rate assignment. As shown above, one can identify a feasibility region Δ (power-limited Δ_P, ROT-controlled Δ_Z, or a linearized approximation Δ_L) consisting of rate vectors which can be served in a cellular structure, while various constraints on transmission power are met. Our goal here is to underline the desirable properties of various rate assignments and to specify the trade-off between fairness and overall throughput. Ultimately, after defining a reasonable measure of fairness/throughput, we seek to distinguish an "optimal" rate vector chosen from the feasibility region. Mathematically, we seek to define an appropriate objective function of transmission rates, $F(\underline{\alpha}) := F(\alpha_1, \ldots, \alpha_N)$, and solve the following:

$$(\alpha_1^{opt}, \ldots, \alpha_N^{opt}) := \arg\max_{\underline{\alpha} \in \Delta} F(\underline{\alpha}). \tag{8}$$

A. Overall Throughput

The overall throughput of a system with rate vector $\underline{\alpha}$ is defined to be the sum of the rates assigned to the mobiles, i.e.

$$F(\alpha_1, \ldots, \alpha_N) = \sum_{i=1}^{N} R_b \alpha_i. \tag{9}$$

We can generalize this by considering the case when users have a limited number of packets to send, and thus benefit only from a limited transmission rate. This can be modeled as $F(\alpha_1, \ldots, \alpha_N) = \sum_{i=1}^{N} U_i(\alpha_i)$, where

$$U_i(\alpha_i) := \begin{cases} R_b \alpha_i & \text{if } \alpha_i < \alpha_i^{\max} \\ R_b \alpha_i^{\max} & \text{if } \alpha_i \geq \alpha_i^{\max} \end{cases} \tag{10}$$

where $R_b \alpha_i^{\max} \leq W$ is user i's maximum transmission rate.

Note that in the case of a single-cell system, this problem can be converted to the throughput maximization problem studied in [21] where the nature of the solution is shown to be an opportunistic scheduling. The multi-sector case is more complicated, and we do not attempt to provide an analytic solution for it in this paper. The numerical solution is shown in Figure 2 for various definitions of Δ and manifest an opportunistic solution where the user with the best channel gets r^{max}, while the other is given the left-over rate (the filled circles).

B. Max-Min Fairness

As seen in Figure 2, the maximum overall throughput might be achieved only at the cost of specific users. In other words, maximum overall throughput and fairness are traded-off in most scenarios. There are several ways to address the issue of fairness in resource allocation. One of the most common ways to address fairness is the max-min fair criterion [3], [1], [22].

Definition 6: A vector of rates $(\alpha_1^*, \alpha_2^*, \ldots, \alpha_N^*)$ is *max-min fair* if it maximizes the minimum assigned rate in the network (hence, the term max-min); i.e. for any feasible rate vector $(\alpha_1, \alpha_2, \ldots, \alpha_N)$ we have

$$\min_j \alpha_j^* \geq \min_j \alpha_j. \tag{11}$$

This corresponds to the solution to Eqn (8) where $F(\alpha_1, \ldots, \alpha_N) = \min_j \alpha_j$.

Remark: Our definition of max-min fair rate assignment is less specific than the max-min fairness as defined in [1]. In fact, if a rate vector is max-min fair in the context of [1], it is max-min fair in our context but the reverse is not true. When the max-min solution defined above is unique, however, our solution is also max-min fair as defined in [1].

Theorem 4 below shows that there always exists a max-min fair rate for which $\alpha_1 = \ldots = \alpha_N$.

Theorem 4: Consider the feasibility region Δ $(= \Delta_P, \Delta_Z, \Delta_L)$. If the vector $(\alpha_1^*, \alpha_2^*, \ldots, \alpha_N^*)$ is a max-min fair solution over Δ, then the vector $(\tilde{\alpha}_1, \tilde{\alpha}_2, \ldots, \tilde{\alpha}_N)$ such that $\tilde{\alpha}_i = \min_{j \in M_{b(i)}} \alpha_j^*$ for $\forall i$ is also a max-min fair solution. Similarly, the vector of all equal rates $(\hat{\alpha}, \hat{\alpha}, \ldots, \hat{\alpha})$, $\hat{\alpha} = \min_{1 \le j \le N} \alpha_j^*$ is a max-min fair solution.

Proof:

Consider mobile i such that $i \in M_l$ and $\alpha_i^* > \min_{j \in M_l} \alpha_j^*$. Mobile i can reduce its rate to $\tilde{\alpha}_i = \min_{j \in M_l} \alpha_j^*$ while transmitting at the same power. This will result in a lower rate transmission but with better bit error rate (BER) for mobile i without changing the SINR of other mobiles. This would imply that the new rate vector $(\alpha_1^*, \alpha_2^*, \ldots, \tilde{\alpha}_i, \ldots, \alpha_N^*)$ is both a feasible rate allocation and satisfies Eqn (11) (this proves that all of the rate feasibility regions defined in the last section are coordinate convex [11]). Hence the above vector is a max-min fair solution.

Similarly, every user whose rate is not the minimum rate can reduce its rate to the minimum allocated rate (notice that this minimum rate, by definition, is maximized). ■

It is easy to construct examples where the max-min fair solution (in the context of Definition 6) is not unique. For instance, the overall max-min fair solution is not unique if there is no inter-cell interference between neighboring cells (hence, many decoupled cells) with asymmetric loads (in terms of number of mobiles and their locations). In such cases as seen in the above proof, reducing the rates while maintaining the same power, in effect, results in an overall consumption of less bandwidth, i.e. the rate assignment $(\hat{\alpha}, \hat{\alpha}, \ldots, \hat{\alpha})$ might be too inefficient. In other words, when more than one solution exists, a more interesting question remains to find the most efficient max-min fair solution, e.g. the one achieving the best overall throughput among all max-min fair solutions.

On the other hand, we believe that in situations where mobiles contribute to the interference of each sector in the network (i.e. when $g_{il} \neq 0 \; \forall \; i, l$), the max-min fair rate vector over Δ_P, Δ_Z, or Δ_L is unique. Although we have not been able to prove this rigorously for Δ_P or Δ_Z, we present the following theorem for the linearized region Δ_L.

Theorem 5: If the channel gains in a network are such that $g_{il} \neq 0 \; \forall \; i, l$, the max-min fair solution $(\alpha_1^*, \alpha_2^*, \ldots, \alpha_N^*)$ over Δ_L is unique and is such that $\alpha_1^* = \alpha_2^* = \ldots = \alpha_N^*$.

Proof:

From Theorem 4, we know that $\alpha_1^* = \alpha_2^* = \ldots = \alpha_N^*$ is always a max-min fair solution - we need only prove that it is unique. To do so, we use the following solidarity property from [22]:

> If $\underline{\alpha} \in \Delta_L$, then $\forall \; i$ there exists a vector $\underline{\delta}$ with $\delta_i < 0$ and $\delta_j > 0 \; \forall j \neq i$ such that $\underline{\alpha} + \underline{\delta} \in \Delta_L$.

It is clear that this property guarantees that the equal-rate solution is the unique max-min fair solution (since decreasing the rate of any user will increase the rate of all other users). It is left only to prove that Δ_L satisfies the solidarity property.

Assume we have a set of feasible effective rates $(r_1, \ldots, r_N)$. Let each user's new rate be $r_j + \delta_j$, where user i has $\delta_i < 0$ and all other users have $\delta_j > 0$. For each base, we have the following:

$$\begin{aligned} \sum_{j \neq i} \quad (r_j + \delta_j) \quad \frac{g_{jl}}{g_{jb(j)}} + (r_i + \delta_i) \frac{g_{il}}{g_{ib(i)}} \\ \leq \qquad \frac{K}{\gamma(1+K)} + \sum_{j \neq i} \delta_j \frac{g_{jl}}{g_{jb(j)}} + \delta_i \frac{g_{il}}{g_{ib(i)}} \end{aligned}$$

From this inequality, it is clear we can choose $\underline{\delta}$ as described above such that $\underline{r} + \underline{\delta}$ is also a set of feasible effective rates. Hence, Δ_L satisfies the solidarity property. ■

The (all equal) max-min solution is shown in Figure 2 for various notions of feasibility regions.

C. Proportional Fairness

To address the issue of fairness, other performance measures such as proportional fairness can be used [1], [6]. This is equivalent to a utility maximization (solution to Eqn (8) when $F(\underline{\alpha}) = \sum_{i=1}^{N} U_i(\alpha_i)$) where mobile i's utility is defined as:

$$U_i^2(\alpha_i) := \begin{cases} \log(R_b \alpha_i) & \text{if } \alpha_i < \alpha_i^{\max} \\ \log(R_b \alpha_i^{\max}) & \text{if } \alpha_i \geq \alpha_i^{max} \end{cases} \tag{12}$$

Hence, we seek to solve:

$$\max_{\underline{r} \in \Delta} \sum_{i=1}^{N} U_i^2\left(\frac{W r_i}{1 - \gamma r_i}\right) \tag{13}$$

where Δ can be either Δ_P, Δ_Z, or Δ_L. The diamonds in Figure 2 show the solutions to Eqn (13) considering various notions of feasibility regions. Note how the proportional fair solutions are the intermediate solutions between max-min fair and maximum total throughput.

V. Distributed Proportional Fair Rate Assignment

This section provides the last and somewhat most interesting results of our paper. Here, we use the results from previous sections to construct algorithms that solve the proportional fair rate assignment in a distributed manner. We also show that these distributed algorithms can be implemented in a practical manner at individual base stations and mobiles, following the structure of the CDMA2000 standards described earlier.

Consider the following problem:

<u>Problem (P)</u>

> Consider the model introduced in Section II-B and U_i^2 in Eqn (12). We seek to solve the following:
>
> $$\max_{\underline{r} \geq 0} \sum_{i=1}^{N} U_i^2\left(\frac{W r_i}{1 - \gamma r_i}\right) \tag{14}$$
>
> subject to
>
> $$\max_{l} \sum_{i=1}^{N} \frac{g_{il}}{g_{ib(i)}} r_i \leq \frac{K}{\gamma(1+K)} \quad ,$$

As we will see in this section, under realistic scenarios this problem has a distributed solution. This is extremely desirable for the practical design of DO systems. The rest of this section provides this distributed solution to Problem (P).

We now make the following assumption:

Assumption 2: Each user i's maximum rate is such that $R_b\alpha_i^{\max} < \frac{W}{\gamma}$.

This assumption is necessary to guarantee that no user is allowed to monopolize the total bandwidth, i.e. each user's transmission rate is such that, even at its maximum, it still leaves half of the total effective bandwidth unused, i.e. $r_i \leq \frac{K}{2\gamma(K+1)}$. It is easy to show that, under Assumption 2, Problem (P) reduces to optimizing a strictly concave function subject to a set of linear constraints. In other words, Problem (P) can be solved using the dual method.

Consider the Lagrangian

$$\begin{aligned}
&\mathcal{L}(\underline{r}, \underline{\mu}) \\
&= \sum_{i=1}^{N} U(r_i) - \sum_{l=1}^{L} \mu_l (\sum_{i=1}^{N} \frac{g_{il}}{g_{ib(i)}} r_i - \frac{K}{\gamma(1+K)}) \\
&= \sum_{i=1}^{N} \left(U(r_i) - r_i (\sum_{l=1}^{L} \frac{g_{il}}{g_{ib(i)}} \mu_l) \right) + \frac{K}{\gamma(1+K)} \sum_{l=1}^{L} \mu_l
\end{aligned}$$

where $U(r_i) = U_i^2(\frac{Wr_i}{1-\gamma r_i})$.

Since, under Assumption Assumption 2, $U(r)$ is a monotone non-decreasing and concave function of r, there is no duality gap [23]. Hence, the above optimization problem can be addressed by solving the dual problem:

$$\min_{\underline{\mu} \geq 0} \sum_{i=1}^{N} \phi_i(p_i) + \frac{K}{\gamma(1+K)} \sum_{l=1}^{L} \mu_l \tag{15}$$

where $p_i = \sum_{l=1}^{L} \frac{g_{il}}{g_{ib(i)}} \mu_l$ and

$$\phi_i(p_i) = \max_{r_i} (U(r_i) - r_i p_i). \tag{16}$$

Equations (15) and (16) are powerful tools in proposing distributed algorithms in the studied system. Convex duality implies that at the optimum $\underline{\mu}^*$ (which may not be unique), the optimum r^* (maximizing the individual utility minus the cost in Eqn (16)) is exactly the solution to the primal problem. In other words, provided the equilibrium prices μ^* can be made to align with the Lagrange multipliers, the individual optima, computed in a decentralized fashion by sources, will align with global optima of Eqn (14). Furthermore, we will see how a simple gradient projection method can provide an effective method to compute the optimum $\underline{\mu}^*$.

A natural way to interpret the above Lagrangian multipliers is to introduce a price per unit rate, p_i. Note that these prices are not dollar value prices, but rather regulating/coordinating signals which are produced by each base station to indicate the level of interference at each sector. In this way, each mobile uses the indication of high levels of interference to back off its rate appropriately. When we calculate the Lagrangian multipliers using gradient projection, the system forms a distributed feedback loop, as shown in Figure 1.

In the rest of this section, we use this structure to propose a distributed implementation for rate control. Note that we will need certain assumptions, such as symmetry between forward and reverse links, negligible thermal noise, etc. These assumptions are needed to extend the above convergence and optimality to the proposed rate control scheme. It is worth noting that these assumptions are not necessary for the operation of the system, but only to guarantee the optimality and convergence of the distributed scheme. In Section VI, we address the impact of violating these assumptions in the overall system performance.

First assume that each base makes an announcement of an optimum multiplier μ_l. Under such a scenario, each mobile is only required to vary its rate according to the following expression to maximize its own "profit."

$$r_i = \arg\max_r (U(r) - r \sum_{l=1}^{L} \frac{g_{il}}{g_{ib(i)}} \mu_l). \tag{17}$$

Mobile Algorithm: At each computation epoch, each mobile has to compute its (seemingly selfish) optimal rate by solving Eqn (17). The challenge seems to be that each mobile needs to compute its weighted price $\sum_{l=1}^{L} \frac{g_{il}}{g_{ib(i)}} \mu_l$. We next show that through the introduction of pricing pilot signals, $\sum_{l=1}^{L} \frac{g_{il}}{g_{ib(i)}} \mu_l$ can be easily available to the mobiles without an assumption on the mobiles' knowledge of the channel gain values g_{il}, $g_{ib(i)}$.

In CDMA each base station transmits a pilot signal (PS) with a fixed transmission power (P_t^P) over the forward link channel. Assuming symmetric forward and reverse link gains (a reasonably common assumption), this pilot signal is usually used by mobiles to perform channel estimation and power control [15]. Similarly, we propose that a *Pricing Pilot Signal (PPS)* is implemented as follows. Each PPS is transmitted on the forward link channel synchronously. The transmitted power of PPS for base l is μ_l times the transmission power of the primary pilot signal, P_t^P. This is similar to the reverse activity bit (RAB) in CDMA2000 [15]. Assuming synchronized PPS transmissions from all bases, negligible thermal noise, and symmetry between forward and reverse links, for each mobile i we have

$$\sum_{l=1}^{L} \frac{g_{il}}{g_{ib(i)}} \mu_l = \frac{E_R^{PPS}}{E^P(b(i))},$$

where $E_R^{PPS} \propto P_t^P \sum_{l=1}^{L} g_{il}\mu_l$ denotes the total PPS energy measured by mobile i, and $E^P(b(i)) \propto P_t^P g_{ib(i)}$ is the PS energy measured by mobile i from its tracking sector $b(i)$.

In summary, at every decision epoch, each mobile computes $r_i^* = \arg\max \left(U(r) - r\frac{E_R^{PPS}}{E^P(b(i))} \right)$. Mobile i then transmits its information with rate $\alpha_i^* = \frac{Wr_i^*}{1-\gamma r_i^*}$.

Our next challenge is to compute the multipliers (optimal prices) in a distributed manner at each base station over a large network. The simplest algorithm that guarantees these equilibrium prices μ^* is based on the gradient projection [5]:

$$\begin{aligned}
\dot{\mu_l} &= \begin{cases} \beta(\sum_{i=1}^{N} \frac{g_{il}}{g_{ib(i)}} r_i - \frac{K}{\gamma(1+K)}) & \text{if } \mu_l(t) > 0 \\ \beta[\sum_{i=1}^{N} \frac{g_{il}}{g_{ib(i)}} r_i - \frac{K}{\gamma(1+K)}]^+ & \text{if } \mu_l(t) = 0 \end{cases} \\
&= \begin{cases} \frac{\beta}{\gamma}(\sum_{k=1}^{L} \frac{Y_{kl}}{1+Z_k} - \frac{K}{\gamma(1+K)}) & \text{if } \mu_l(t) > 0 \\ \frac{\beta}{\gamma}[\sum_{k=1}^{L} \frac{Y_{kl}}{1+Z_k} - \frac{K}{\gamma(1+K)}]^+ & \text{if } \mu_l(t) = 0 \end{cases}
\end{aligned}$$

where $Y_{kl} = \sum_{i \in M_k} g_{il} P_i$ is the interference received from sector k at the base station l, and the second equality is due to Eqn (6). Note that if k and l are far from each other $Y_{kl} \approx 0$.

Base Algorithm 1: Based on the above, we propose the following algorithm:

$$\dot{\mu_l} = \begin{cases} \frac{\beta}{\gamma}(\sum_{i=1}^{L} \frac{\hat{Y_{kl}}}{1+\hat{Z_{kl}}} - \frac{K}{\gamma(1+K)}) & \text{if } \mu_l(t) > 0 \\ \frac{\beta}{\gamma}[\sum_{k=1}^{L} \frac{\hat{Y_{kl}}}{1+\hat{Z_{kl}}} - \frac{K}{\gamma(1+K)}]^{+} & \text{if } \mu_l(t) = 0 \end{cases}$$

where $\hat{Y_{kl}}$ and $\hat{Z_{kl}}$ are estimates of Y_{kl} and Z_k by base l. In other words, the gradient of the Lagrangian depends on each base station's estimate of the load of its neighboring cells. The better these estimates are, the closer the solution is to the optimal rate assignment.

When estimation error in $\hat{Y_{kl}}$ and $\hat{Z_{kl}}$ are negligible, the dynamic interaction between the mobile and base algorithms, in effect, simulates a gradient projection-based computation of the solution to the dual problem (15). These algorithms are similar in nature to those studied in [24], [4], and constitute a distributed computation system. Similarly, Theorems 2 and 3 in [4] can be used to establish the convergence of the distributed computation. This implies that the proposed (decentralized) mobile and base algorithms converge to a proportional fair rate assignment. In real systems, the performance of these algorithms depends on the ability of base l to obtain good estimates for Y_{kl} and Z_k. This requires that each base 1) periodically receives information from its neighboring bases regarding their ROT, and 2) identifies sector source of the uplink interference it experiences. These are significant drawbacks for the implementation of this algorithm in real systems. Base Algorithm 2, proposed below, is a heuristic attempt to mitigate these drawbacks.

Base Algorithm 2: Our second algorithm uses a gradient projection method, but uses an overestimation on the value of $\hat{Z_{kl}}$, in the sense that we use $\hat{Z_{kl}} = K \; \forall l, k$.

$$\dot{\mu_l} = \begin{cases} \beta(Z_l - K) & \text{if } \mu_l(t) > 0 \\ \beta[Z_l - K]^{+} & \text{if } \mu_l(t) = 0 \end{cases}$$

This algorithm does not require any estimate of the loading of or the interference caused by the neighboring cells.

We do not have any analytical results on the convergence/equilibrium-point when Base Algorithm 2 is used. Figure 2 shows such an equilibrium point. This equilibrium in general, violates the linearized constraints LC1, but remains limited to Δ_Z.

It is important to note that, while the implementation details differ, the distributed feedback structure described in this section is identical to the structure described by the CDMA2000 standards [15]. Thus, the algorithms described here represent a comprehensive, optimal rate assignment scheme that fits within the existing structure of CDMA2000.

VI. Numerical Examples and Simulations

In order to examine the behavior of the distributed algorithms described in the previous section, simulations were run using a layout of four base stations (each 2500m apart) and 20 randomly positioned mobiles. The simulations use a cost-231 propagation model at 1.9 GHz between the mobiles and bases [15]. The values for γ and K are 4 dB and 6 dB. The chip bandwidth W is 1.2 MHz, and the pilot rate R_b is 4.8 Kbps. For simulations involving mobility, a subset of five mobiles were allowed to move using a model of constant speed (120 km/hr) and random directions. When shadowing is modeled, we use log-normal shadowing with a standard deviation of $\sigma = 8$dB, and the thermal noise level is -179 dBm/Hz. Soft-handoff is not simulated. The *Base Algorithm* is run every 20 ms, and the mobiles respond to new prices by running the *Mobile Algorithm* to generate new rates every 20 ms. For simulations involving power control, a closed-loop power control mechanism is run every 5 ms using ΔP=.25 dB.

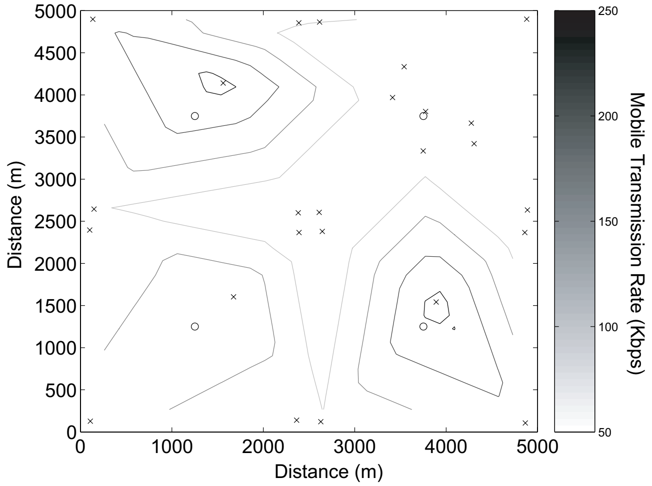

Fig. 4. Contour plot of equilibrium rate assignments.

The performance of the distributed algorithms will obviously be impacted by the particular choice of parameter values described above. Discussion of this falls beyond the scope of this paper, but can be found in [25]. For purposes of this paper, we have chosen to simulate our algorithms using realistic parameter values as described in [26].

Figure 4 shows a contour plot illustrating the equilibrium point of the distributed rate allocations for various mobiles. To illustrate the dependency of rate on the sum of the gain ratios for neighboring cells, we have ignored log-normal shadowing. We note that this shows that a proportional fair rate assignment might require unbalanced and unequal rate assignments. This result confirms that an equal rate assignment performs sub-optimally with respect to the sector capacity because it does not distinguish the mobiles in the inner part (close to only one base) from those at the boundary whose interference affects more than one base.

So far, we have shown that the proposed distributed algorithm converges to an equilibrium when the network topology and gains are static. However, dynamic variations can cause significant performance degradation and may need to be addressed. Through extensive simulations and observation, shown in Figures 5, 6, and 7, we have classified these dynamic variations into three types: fast and random variations, sudden but quasi-static topological changes, and persistent dynamic variations. Note that in Figures 5 and 6 we examine the effect of time-varying parameters one at a time, while Figure 7 illustrates the performance of the algorithms in a practical setting incorporating realistic dynamics.

Remark: Figure 5 shows the transmission rate for an individual user in the presence of dynamic variations. This transmission rate does not necessarily represent the mobile's

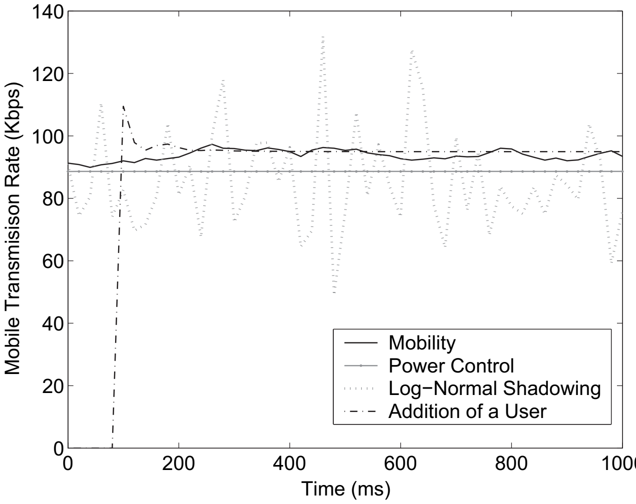

Fig. 5. Rate assignment for an individual user in the presence of dynamic variations.

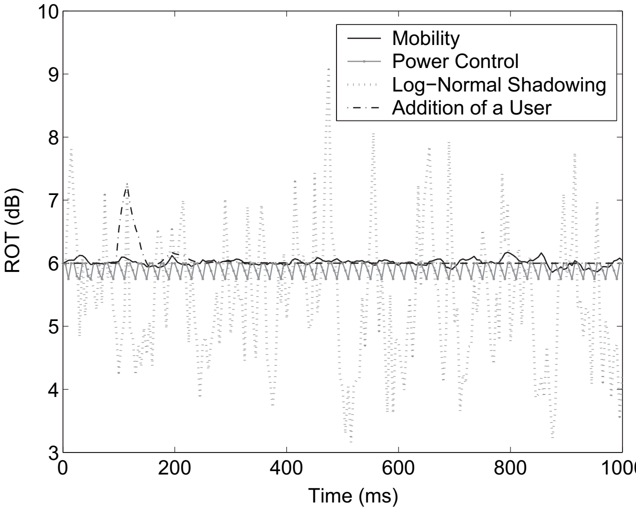

Fig. 6. ROT Levels in the presence of dynamic variations ($ROT > 6dB$ yields unacceptable BER.

throughput, as during the periods in which $ROT > K$ dB, the mobile encounters higher (even unacceptable) BER.

Fast and Random Variations

Fast and random variations captures the use of imperfect power control and mobility. These dynamics tend to introduce a level of uncertainty around the equilibrium point, but only cause minor performance degradation. As seen in Figures 5 and 6, the system performance remains close to the predicted equilibrium point even when considering a realistic level of imperfect power control ($\Delta P = .25dB$) or mobile speeds on the order of 120 km/hr.

Persistent Dynamic Variations

Persistent dynamic variations in network setting occur continually over time, and include shadowing. Unlike fast and random variations, however, persistent dynamic variations do cause significant performance degradation and require some level of compensation. In other words, the time constants of significant change in network setting and the distributed algorithm are comparable. An important example is shadowing where the time scales of change for the gains $g_{il}(t)$ are

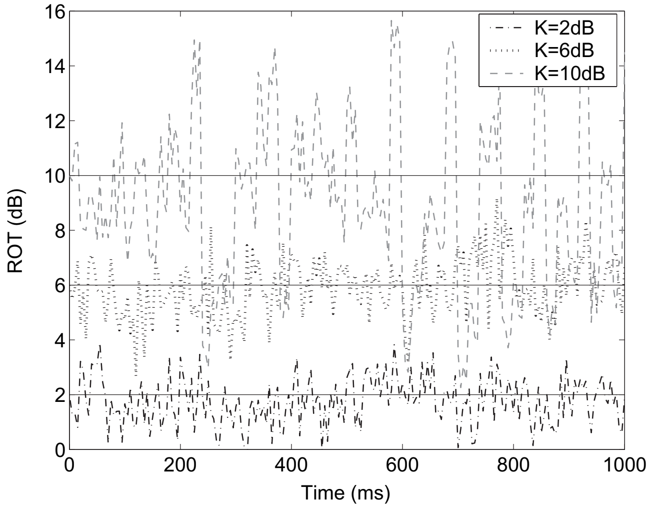

Fig. 7. ROT levels in the presence of dynamic variations for various levels of K - lower values of K increase robustness.

at the same order as the time constants of the distributed algorithm. In Figures 5 and 6, we see the effect of shadowing on the algorithm performance. When considering these types of dynamics, the transient mode is the dominant mode. Figure 7 shows the performance of the algorithms under realistic conditions for several levels of K. We observe that when the value of K is lowered, the algorithm tends to be more resistant to these dynamic variations. Intuitively, K seems to be damping the transient behavior, as we had discussed in Section III.

Sudden but Quasi-Static Topological Changes

Most topological changes can be classified as quasi-static: that is, they cause sudden and significant changes but tend to occur sparsely over time. Examples of such changes include the addition and removal of users from the system. Depending on the behavior of upper layer protocols, (particularly the transport layer) changes in user demand may (or may not) be classified as quasi-static. The latter is the case when the parameters α_i^{max} (introduced in Section IV) change in a quasi-static manner. These types of changes, while occurring infrequently, tend to cause a significant spike in ROT at the affected base, violating the constraints C1 and C2 before the distributed algorithm converges to a new equilibrium.

In Figures 5 and 6, we see the effect of adding a new user at time $t = 100ms$. The spike in ROT that accompanies the addition of a new user violates system constraints, causing unacceptably high BER. In addition, we see that the new user sees not only a high initial transmit rate, but a large back-off in rate and a fairly slow convergence time. It is a topic of current research to develop modified base and mobile algorithms that will mitigate the undesirable effects introduced by these dynamics [25].

VII. Conclusion

In this paper we address the issue of rate allocation in a wideband CDMA system serving users with varying rates. We first provided various definitions of feasible rate vectors. We then formulated the rate allocation problem as a global optimization problem, the solution of which depends on the mobile layouts and tracking rules. The objective function of such an optimization problem determines the trade-off between total throughput and fairness. Furthermore, we have provided pairs of distributed mobile/base algorithms whose equilibrium point coincides with the proportional fair solution to the original global optimization problem. In other words, we propose these algorithms as desirable solutions to the problem of decentralized and distributed rate assignment in CDMA systems in scenarios where changes in the characteristics of the network (layout, tracking assignment, and utilities) are slow. Finally, we demonstrate the performance of these algorithms via numerical examples.

Appendix I
Proof of Theorems 1 and 2

We use Condition C1 to express the vector of powers $\underline{\mathbf{P}}$ in terms of effective rates $r_i,\ i = 1, 2, \ldots, N$. From Condition C1, for any mobile i and its assigned base l (i.e. $b(i) = l$), we have

$$P_i g_{il}(s_i + \gamma) = \gamma(N_0 W + \sum_{j=1}^{N} P_j g_{jl}) = \gamma N_0 W(1 + Z_l).$$

Notice that the right hand side of the above equation is the same for all mobiles tracked by base l, i.e. for $\forall i \in M_l$, we have

$$P_i g_{il} = \gamma N_0 W r_i (1 + Z_l). \tag{18}$$

On the other hand, we can rewrite Eqn (3) as

$$\begin{aligned} Z_l &= \frac{\sum_{j=1}^{N} P_j g_{jl}}{N_0 W} = \frac{1}{N_0 W} \sum_{k=1}^{L} \sum_{j \in M_k} P_j g_{jl} \qquad (19) \\ &= \frac{1}{N_0 W} \sum_{k=1}^{L} \sum_{j \in M_k} P_j g_{jk} \frac{g_{jl}}{g_{jk}} \end{aligned}$$

Now using Eqn (18) we have

$$\begin{aligned} Z_l &= \gamma \sum_{k=1}^{L} \left((1 + Z_k) \sum_{j \in M_k} \frac{r_j g_{jl}}{g_{jk}} \right) \qquad (20) \\ &= \gamma \sum_{k=1}^{L} \left((1 + Z_k) \sum_{j=1}^{N} \frac{\Psi_{jk} r_j g_{jl}}{g_{jk}} \right) \\ &= \gamma \sum_{j=1}^{N} G_{lj} \left(\sum_{k=1}^{L} R_{jk} (1 + Z_k) \right) \end{aligned}$$

We can write Eqn (20) in the matrix form:

$$\underline{Z} = \gamma G R (\underline{\mathbf{1}}_L + \underline{Z}) \tag{21}$$

Now under Assumption 1 and using Proposition 1, we solve the above equation:

$$\underline{Z} = \gamma (I - \gamma G R)^{-1} G R \underline{\mathbf{1}}_L = (I - \gamma G R)^{-1} \underline{\mathbf{1}}_L - \underline{\mathbf{1}}_L \tag{22}$$

where the second equality follows from $(I - A)^{-1}(I - A) = I$.

On the other hand, from Eqn (18) we have

$$\underline{P} = \gamma N_0 W R (\underline{\mathbf{1}}_L + \underline{Z}) = \gamma N_0 W R (I - \gamma G R)^{-1} \underline{\mathbf{1}}_L \tag{23}$$

Hence, satisfying Condition AC1 is sufficient and necessary to guarantee Condition C2. This completes the proof of Theorem 1.

Similarly, Eqn (22) implies that satisfying Condition AC2 is sufficient and necessary to guarantee Condition C4. This completes the proof of Theorem 2.

■

Appendix II
Proof of Theorem 3

We show that LC1 is a sufficient condition for AC2. To do so, we notice that LC1 can be written as:

$$\|\gamma \mathbf{GR}\|_\infty \le \frac{K}{1 + K}. \tag{24}$$

Using the norm inequality $\|(I - A)^{-1}\| \le \frac{1}{1 - \|A\|}$ and LC1, we have

$$\|(I - \gamma \mathbf{GR})^{-1}\|_\infty \le \frac{1}{1 - \|\gamma \mathbf{GR}\|_\infty} \le \frac{1}{1 - \frac{K}{1+K}} = 1 + K \tag{25}$$

Put vector $\mathbf{V} = (I - \gamma \mathbf{GR})^{-1}$. From the definition of the infinity norm and Eqn (25), we have $\max_i V_i \le 1 + K$, or $\mathbf{V} \le (1 + K)\underline{1}_L$. This is nothing but AC2, and we are done.

■

Acknowledgment

The authors would also like to thank Dr. C. Lott and Dr. D. Zhang at Qualcomm Corporate R & D for their thoughtful suggestions.

References

[1] F. Kelly, "Mathematical modelling of the Internet," in *Mathematics Unlimited - 2001 and Beyond*, B. Engquist and W. Schmid, Eds. Springer-Verlaq, 2001, pp. 685–702.

[2] S. Shenker, "Fundamental design issues for the future Internet," *IEEE J. Sel. Areas Commun.*, vol. 13, pp. 1176–1188, 1995.

[3] F. Kelly, "Charging and rate control for elastic traffic," *European Trans. Telecommun.*, vol. 8, pp. 33–37, 1997.

[4] S. H. Low and D. E. Lapsley, "Optimization flow control, I: basic algorithm and convergence," *IEEE/ACM Trans. Networking*, vol. 7, no. 6, pp. 861–874, Dec. 1999.

[5] S. H. Low, F. Paganini, and J. Doyle, "Internet congestion control," *IEEE Control Systems Mag.*, Feb 2002.

[6] J. Mo and J. Walrand, "Fair end-to-end window-based congestion control," *IEEE/ACM Trans. Networking*, vol. 8, no. 5, pp. 555–567, Oct. 2000.

[7] N. Bambos, S. Chen, and G. Pottie, "Channel access algorithms with active link protection for wireless communication networks with power control," *IEEE/ACM Trans. Networking*, vol. 8, no. 5, pp. 583–597, Oct. 2000.

[8] G. Foschini and Z. Miljanic, "A simple distributed autonomous power control algorithm and its convergence," *IEEE Trans. Veh. Technol.*, vol. 42, no. 4, pp. 641–646, Nov. 1993.

[9] C. Huang and R. D. Yates, "Rate of convergence for minimum power assignment algorithms in cellular radio systems," *ACM Wireless Networks*, vol. 4, no. 3, pp. 223–231, Apr. 1998.

[10] D. Mitra and J. Morrison, "A distributed power control algorithm for bursty transmissions on cellular spread spectrum wireless networks," in *Wireless Information Networks*, J. Holtzman, Ed. Kluwer, 1996.

[11] R. Leelahakriengkrai and R. Agrawal, "Scheduling in multimedia CDMA wireless networks," *IEEE Trans. Veh. Technol.*, vol. 52, no. 1, Jan. 2003.

[12] S. Chakravarty, R. Pankaj, and E. Esteves, "An algorithm for reverse traffic channel rate contol for CDMA2000 high rate packet data systems," in *Proc. IEEE Globecom 2001*.

[13] E. Esteves, "On the reverse link performance of CDMA2000 high rate packet data systems," in *Proc. IEEE International Communication Conference 2002*.
[14] S. Oh, D. Zhang, and K. Wasserman, "Optimal resource allocation in multiservice CDMA networks," *IEEE Trans. Wireless Commun.*, vol. 2, no. 4, pp. 811–821, July 2003.
[15] 3rd Generation Partnership Project 2 (3GPP2), "cdma2000 high rate packet data air interface specification," technical report C.S0024-A v2.0, July 2005.
[16] S. Chen, N. Bambos, and G. Pottie, "Admission control schemes for wireless communication networks with adjustable tranmitter powers," in *Proc. IEEE INFOCOM 94*.
[17] R. Horn and C. Johnson, *Matrix Analysis*. Cambridge University Press, 1985.
[18] A. Elwalid and D. Mitra, "Effective bandwidth of general Markovian traffic sources and admission control of high speed networks," *IEEE/ACM Trans. Networking*, vol. 1, pp. 329–343, June 1993.
[19] J. S. Evans and D. Everitt, "Effective bandwidth-based admission control for multiservice CDMA cellular networks," *IEEE Trans. Veh. Technol.*, vol. 48, no. 1, pp. 36–46, Jan. 1999.
[20] F. Kelly, "Notes on effective bandwidths," in *Stochastic Networks: Theory and Applications*, F. Kelly, S. Zachary, and I. Ziedins, Eds. Oxford University Press, 1996, pp. 141–168.
[21] D. Zhang, S. Oh, and N. Sindhushayana, "Optimal resource allocation for data service in cdma reverse link," in *Proc. Wireless Communications and Networking Conference 2004*, vol. 3, pp. 1601–1606.
[22] B. Radunovic and J. LeBoudec, "Rate performance objectives of multihop wireless networks," in *Proc. INFOCOM 04*.
[23] D. Luenberger, *Linear and Nonlinear Programming*. Kluwer Academic Publishers, 2003.
[24] D. Bertsekas and J. Tsitsiklis, *Parallel and Distributed Computation*. Prentice-Hall Inc., 1989.
[25] J. Price and T. Javidi, "On dual methods for adaptive resource allocation in wireless networks: a taxonomy of practical challenges in CDMA," in *Resource Allocation in Next Generation Wireless Networks*, Y. Pan and W. Li, Eds. Nova Science Publishers, Inc., 2006.
[26] P. Black and Q. Wu, "Link budget of CDMA2000 1xev-do wireless Internet access system," in *Proc. 13th IEEE International Symposium on Personal, Indoor and Mobile Radio Communications 2002*, vol. 4, pp. 1847–1852.

Jennifer Price (S'02) received her B.S. and M.S. degrees from the University of Oklahoma and the University of Washington in 2002 and 2004, respectively. Currently, she is pursuing her Ph.D. in Communications Theory and Systems at the University of California, San Diego.

Her research interests include communication networks, optimization, distributed resource allocation, network pricing and economics, and game theory.

Tara Javidi (S'96-M'02) studied electrical engineering at Sharif University of Technology, Tehran, Iran from 1992 to 1996. She received the MS degrees in electrical engineering (systems), and in applied mathematics (stochastics) from the University of Michigan, Ann Arbor, in 1998 and 1999, respectively. She received her Ph.D. in electrical engineering and computer science from the University of Michigan, Ann Arbor, in 2002.

From 2002 to 2004, she was an assistant professor at the Electrical Engineering Department, University of Washington, Seattle. She joined University of California, San Diego, in 2005, where she is currently an assistant professor of electrical and computer engineering. She was a Barbour Scholar during 1999-2000 academic year and received an NSF CAREER Award in 2004.

Her research interests are in communication networks, stochastic resource allocation, stochastic control theory, and wireless communications.