

# Joint Resource Allocation and Base-Station Assignment for the Downlink in CDMA Networks

Jang Won Lee, Ravi R. Mazumdar, and Ness B. Shroff

School of Electrical and Computer Engineering

Purdue University

West Lafayette, IN 47907, USA

{lee46, mazum, shroff}@ecn.purdue.edu

## Abstract

In this paper, we consider the joint resource allocation and base-station assignment problem for the downlink in CDMA networks with heterogeneous data services. We first study a power and rate control problem that attempts to maximize the expected throughput of the system. This problem is inherently difficult because it is in fact a non-convex optimization problem. To solve this problem, we propose a distributed algorithm based on dynamic pricing. This algorithm provides a power and rate allocation that is Pareto-optimal and asymptotically optimal. We study the properties of the proposed power and rate allocation and the global optimal power and rate allocation and show that they have similar properties. We also characterize the optimality conditions of pure TDMA and CDMA type of transmissions, which depend not only on the fixed system parameters but also on the channel state and the maximum data rate of each mobile. Finally, using the outcome of the proposed power and rate allocation algorithm, we propose a pricing based base-station assignment algorithm resulting in joint resource allocation and base-station assignment. In this algorithm, a base-station is assigned to each mobile taking into account the congestion level of the base-station as well as the transmission environment of the mobile. We show that it provides higher performance than the SIR based base-station assignment algorithm in which only the transmission environment of the mobile is considered.

## I. INTRODUCTION

The increasing demand for high data rate services in wireless networks and scarcity of radio resources necessitate efficient usage of radio resources. Some of the main difficulties in using radio resources efficiently are the time and location dependent channel characteristics and the demand to accommodate mobiles

This research has been supported in part by NSF grants ANI-0073359, ANI-9805441, and ANI-0207728.

with diverse service requirements. Hence, for the efficient utilization of radio resources, it is necessary to adopt a resource allocation scheme according to the channel status and the service requirements of mobiles. Moreover, if the joint allocation of two or more resources is considered, the performance gain could be significantly increased.

Recently, there have been a number of papers that have studied joint resource allocation problems [1], [2], [3], [4], [5], [6], [7]. Oh and Wasserman [1] consider a joint power and spreading gain allocation problem for the uplink in Code Division Multiple Access (CDMA) networks. They consider a single cell system and formulate an optimization problem for a single class of mobiles. Their objective is to maximize the total system throughput of mobiles with a constraint on the maximum transmission power of each mobile. In this work, they do not impose any constraint on the spreading gain (i.e., data rate). They show that for the optimal solution, mobiles are selected for transmission according to the channel state and if a mobile is selected, it transmits at its maximum transmission power. They generalize their algorithm to multi-cellular networks in [2]. Bedekar *et al.* [3] and Berggren *et al.* [4] consider joint power and rate allocation for the downlink of the CDMA system with a constraint on the total transmission power for the base-station. However, their models do not consider that each mobile could have a maximum data rate constraint. They show that, for the downlink of the system without the maximum data rate constraint, selecting only one mobile at a time and allocating the total transmission power to that mobile (i.e., the pure Time Division Multiple Access (TDMA) type of strategy) is an optimal strategy.

Joint power allocation and base-station assignment problems for the uplink have been considered by Hanly [5] and Yates and Huang [6]. In these papers, the minimum transmission power satisfying the Signal to Interference Ratio (SIR) threshold of each mobile is obtained by joint power allocation and base-station assignment. At each iteration of the algorithm, each mobile is assigned to the base-station that provides the maximum SIR. Saraydar *et al.* [7] also consider a joint power allocation and base-station assignment problem for the uplink case. They model the problem as an  $N$ -person non-cooperative power allocation game. Each mobile selects the optimal power level and the base-station that maximizes its net utility, (i.e., utility minus cost) without considering other mobiles. The performance of their algorithm depends on the choice of the price. However, they do not provide a strategy of how to determine the optimal price.

In this paper, we study a joint resource allocation and base-station assignment problem for CDMA networks. We focus on the downlink problem in this paper, since the downlink is typically considered a bottleneck due to the asymmetric bandwidth demand between the downlink and the uplink for data services [8],

[9], [10], [11]. We first consider a joint power and rate allocation problem by focusing on one cell of the system. Power is the most important resource in CDMA networks. By increasing the transmission power level of a specific mobile, the performance of the mobile can be improved. But this increase in transmission power also deteriorates the performance of the other mobiles, since it produces more interference to other mobiles. Hence, the power allocation among mobiles is an important resource allocation problem in CDMA networks. Further, there is a trade off between the data rate and  $E_b/I_0$  (bit energy to interference density ratio) that is related to the transmission success probability. At a given transmission power level, if the data rate is increased,  $E_b/I_0$  is decreased and vice versa. Therefore, by jointly optimizing the allocation of power and data rate, the throughput of the system can be improved significantly.

In this paper, the goal of the joint power and rate allocation is to maximize the total expected system throughput. We allow for constraints on the total transmission power at the base-station and the maximum data rate for each mobile. It turns out that this problem is a non-convex programming problem. In general, obtaining a global optimal solution of a non-convex programming problem is not easy and requires a complex algorithm. Hence, in this paper, we propose a simple distributed power and rate allocation algorithm that provides a Pareto-optimal power and rate allocation as well as an asymptotic optimal power and rate allocation. We also study the properties of power and rate allocation in the downlink of the system in the presence of the constraint on the maximum data rate and heterogeneous mobiles. As mentioned above, it was proven that, for the downlink in the idealized system in which there is no constraint on the maximum data rate for each mobile, the optimal strategy is to transmit only one mobile at a time. In addition, it can be easily shown that if all mobiles are homogeneous, transmitting a mobile in the best transmission environment is an optimal strategy. However, in practice, due to either the physical limit of the hardware or the application limit, each mobile cannot receive more than some maximum data rate and there exists a constraint on the maximum data rate. Further, it is possible that each mobile has a different modulation scheme and a different coding level that depend on the channel state and the application [12]. In this case, each mobile has a different function  $f_i$ , since it depends on the modulation and the coding schemes. Because of these two generalizations (that are practically important), a strategy that transmits only one mobile in the best transmission environment may not be optimal. In this paper, we characterize the optimality conditions for pure CDMA type of transmission in which the base-station transmits to all mobiles simultaneously and pure TDMA type of transmission in which the base-station transmits to only one mobile at a time, when there exists a constraint on the maximum data rate for each mobile. We also study the properties of our

mobile selection strategy, when there are heterogeneous mobiles in the system.

We next consider a multi-cellular system. Here we propose a pricing based base-station assignment algorithm. In the traditional SIR based base-station assignment, a mobile is assigned the base-station that can transmit to it at the highest SIR. Hence, in this scheme, only the transmission environment (the channel state and the interference level) between a base-station and a mobile is considered in determining which base-station should communicate to a mobile. However, when a cell is congested, by reassigning some of the mobiles in the cell to other cells that are less congested, we can expect to improve the system performance by balancing the load of cells. Therefore, when a base-station is assigned to a mobile, by taking into account the congestion level of the cell as well as the transmission environment of the mobile, we can significantly improve the performance of the system. However, obtaining optimal assignment of base-stations may need complicated algorithm to be implemented in practice, since it may require global coordinations among base-stations. Hence, we propose a heuristic and simple algorithm in this paper. To measure the congestion level in the cells and the transmission environment of mobiles, we will use the outcomes of the proposed power and rate allocation algorithm in this paper. Our base-station assignment algorithm has two distinguishing features compared with base-station assignment algorithms proposed in [5], [6], [7]. First, it is simple and needs less signaling costs than the other algorithms. In our algorithm, each mobile needs only base-station-specific information that can be broadcasted from the base-station, while in other algorithms, each mobile needs mobile-specific information that requires for the base-station to maintain signaling channels with not only mobiles in its own cell but also mobiles in other cells. Second, our algorithm is based on the downlink performance, while others are based on the uplink performance. As mentioned before, in wireless networks, the downlink is the bottleneck link. Hence, base-station assignment based on downlink performance may be a more appropriate approach. In many problems in wireless networks, the algorithm for the uplink can be used for the uplink with simple modifications. However, as shown in [13], in the joint power allocation and base-station assignment problems, the uplink problem and the downlink problem have different properties. This implies that the algorithm for the uplink cannot be easily modified for the use in the downlink, which necessitates the study of the downlink problem independently.

The rest of the paper is organized as follows. In Section II, we describe the system model. We present the joint power and rate allocation algorithm in Section III and study the properties of power and rate allocation in Section IV. In Section V, the base-station assignment algorithm is presented. We provide numerical results using computer simulation in Section VI and conclude in Section VII.

## II. SYSTEM MODEL

We consider the downlink in a CDMA network consisting of  $B$  base-stations (cells) and  $M$  mobiles. The system is assumed to be time-slotted. At each time slot (this could be a generic interval of time), the power and data rate allocation algorithm and the base-station assignment algorithm are executed. A time slot in our system is an arbitrary interval of time and could consist of one packet or several packets. Each base-station has its maximum power limit  $P_T$  and each mobile communicates with one base-station. For a given mobile  $i$ ,  $R_i^{max}$  is the maximum data rate at which it can receive and  $f_i$  is the probability of packet transmission success, which is a function of  $\gamma_i \triangleq E_b/I_0$  (bit energy to interference density ratio) for mobile  $i$  and depends on the modulation and coding level. We can express  $\gamma_i$  for mobile  $i$  that communicates with base-station  $b$  as:

$$\begin{aligned} \gamma_i(R_i, \bar{P}(b)) &= \frac{W}{R_i} \frac{G_i(b)P_i}{\theta G_i(b) \sum_{m \in M(b)} P_m - \theta G_i(b)P_i + I_i(b)} \\ &= \frac{W}{R_i \theta} \frac{P_i}{\sum_{m \in M(b)} P_m - \theta P_i + \frac{I_i(b)}{G_i(b)}} \\ &= \frac{W}{R_i \theta} \frac{P_i}{\sum_{m \in M(b)} P_m - \theta P_i + A_i(b)}, \end{aligned}$$

where

$W$  : Chip rate.

$\theta$  : Orthogonality factor.

$P_i$  : Allocated power for mobile  $i$ .

$R_i$  : Data rate for mobile  $i$ .

$\bar{P}(b)$  : Power allocation vector for mobiles that communicate with base-station  $b$ .

$G_i(b)$  : Path gain from base-station  $b$  to mobile  $i$ .

$I_i(b)$  : Background noise and intercell interference to mobile  $i$  that communicates with base-station  $b$ .

$M(b)$  : Set of mobiles that communicate with base-station  $b$ .

$A_i(b)$  :  $I_i(b)/G_i(b)$ .

We assume that  $f_i$  has the following properties.

### Assumptions:

- (a)  $f_i$  is an increasing function of  $\gamma_i$ .
- (b)  $f_i$  is twice continuously differentiable.

- (c)  $f_i(0) = 0$ .
- (d)  $\frac{\partial^2 f_i(\gamma_i)}{\partial \gamma_i^2} \left( \frac{W}{R_i} + \theta \gamma_i \right) + 2\theta \frac{\partial f_i(\gamma_i)}{\partial \gamma_i} = 0$  has at most one solution for  $\gamma_i > 0$ .
- (e) If  $\frac{\partial^2 f_i(\gamma_i)}{\partial \gamma_i^2} \left( \frac{W}{R_i} + \theta \gamma_i \right) + 2\theta \frac{\partial f_i(\gamma_i)}{\partial \gamma_i} = 0$  has one solution at  $\gamma_i^o > 0$ ,  $\frac{\partial^2 f_i(\gamma_i)}{\partial \gamma_i^2} \left( \frac{W}{R_i} + \theta \gamma_i \right) + 2\theta \frac{\partial f_i(\gamma_i)}{\partial \gamma_i} > 0$  for  $\gamma_i < \gamma_i^o$  and  $\frac{\partial^2 f_i(\gamma_i)}{\partial \gamma_i^2} \left( \frac{W}{R_i} + \theta \gamma_i \right) + 2\theta \frac{\partial f_i(\gamma_i)}{\partial \gamma_i} < 0$  for  $\gamma_i > \gamma_i^o$ .

*Remark 1: By the assumptions above, if  $\sum_{i \in M(b)} P_i$  is constant,  $\frac{\partial^2 f_i(\gamma_i)}{\partial P_i^2}$  has at most one inflection point. In addition, if it has an inflection point  $P_i^o$ , then  $\frac{\partial^2 f_i(\gamma_i)}{\partial P_i^2} > 0$  for  $P_i < P_i^o$  and  $\frac{\partial^2 f_i(\gamma_i)}{\partial P_i^2} < 0$  for  $P_i > P_i^o$ . We will show that  $\sum_{i \in M(b)} P_i = P_T$  later. Therefore,  $f_i$  can be one of three types: a sigmoidal-like function of its own power allocation<sup>1</sup>, a concave function of its own power allocation, or a convex function of its own power allocation. In most cases, the probability of packet transmission success can be characterized by one of these three types of functions. In Fig. 1, we provide some examples for Binary Phase-Shift Keying (BPSK), Differential Phase-Shift Keying (DPSK), and Frequency-Shift Keying (FSK) modulation schemes. In this figure, we assume a packet consists of 800 bits without channel coding and set  $P_T = 10$ ,  $\theta = 1$ ,  $\frac{W}{R_i} = 16$ , and  $A_i(b) = 0.7407$ .*

### III. POWER AND RATE ALLOCATION

In this section, we study the power and data rate allocation problem focusing on one cell of the system. For notational convenience, we will omit the parameter for the base-station in variables. We assume that mobiles from 1 to  $M$  communicate with their target base-station. First, we define  $U_i$ , the utility function of

<sup>1</sup>A sigmoidal-like function means a function  $f_i(x)$  that has one inflection point,  $x^o$  and  $\frac{d^2 f_i(x)}{dx^2} > 0$  for  $x < x^o$  and  $\frac{d^2 f_i(x)}{dx^2} < 0$  for  $x > x^o$ , i.e.  $f_i(x)$  is a convex function for  $x < x^o$  and a concave function for  $x > x^o$ .

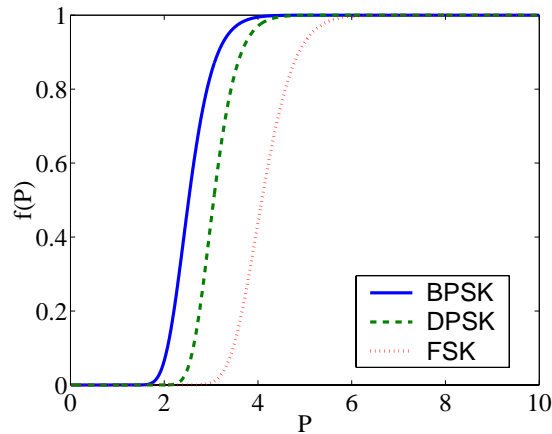


Fig. 1. Probabilities of packet transmission success for BPSK, DPSK, and FSK modulation schemes.

mobile  $i$  as

$$U_i(R_i, \bar{P}) = R_i f_i(\gamma_i(R_i, \bar{P})), \quad i = 1, 2, \dots, M, \quad (1)$$

which is the expected throughput of mobile  $i$ . Then, the optimization problem for power and rate allocation considered in this paper is given by

$$\begin{aligned} \text{(A)} \quad & \max \sum_{i=1}^M U_i(R_i, \bar{P}) \\ & \text{subject to } \sum_{i=1}^M P_i \leq P_T, \\ & 0 \leq P_i \leq P_T, \quad i = 1, 2, \dots, M, \\ & 0 \leq R_i \leq R_i^{\max}, \quad i = 1, 2, \dots, M. \end{aligned}$$

Therefore, our goal is to *maximize the total expected system throughput* (the sum of the expected throughput of all mobiles) with constraints on the total transmission power at the base-station and the maximum data rate for each mobile.

To solve problem (A), we first calculate the optimal data rate given a power allocation and, next, obtain the power allocation. It can easily be shown that, to maximize the total system throughput, the base-station must transmit at its maximum power limit,  $P_T$ . This implies that  $\sum_{i=1}^M P_i = P_T$  and

$$\begin{aligned} \gamma_i(R_i, \bar{P}) &= \frac{W}{R_i} \frac{P_i}{\theta \sum_{m=1}^M P_m - \theta P_i + A_i} \\ &= \frac{W}{R_i} \frac{P_i}{\theta P_T - \theta P_i + A_i} \\ &\triangleq \gamma_i(R_i, P_i), \quad i = 1, 2, \dots, M. \end{aligned}$$

Note that  $\gamma_i(R_i, P_i)$  does not depend on the power allocation for other mobiles and we can rewrite the utility function for mobile  $i$  defined in Equation (1) as

$$U_i(R_i, P_i) = R_i f_i(\gamma_i(R_i, P_i)), \quad (2)$$

which does not depend on the power allocation for other mobiles. Therefore, each mobile  $i$  can determine  $R_i^*(P_i)$ , the optimal data rate for a given power allocation  $P_i$  without considering the other mobiles. The optimal data rate for mobile  $i$  for a given power allocation is obtained by the following proposition.

*Proposition 1: For the power allocation  $P_i$ ,  $R_i^*(P_i)$ , the optimal rate for mobile  $i$ , is given by*

$$R_i^*(P_i) = \begin{cases} \frac{WP_i}{\gamma_i^*(\theta P_T - \theta P_i + A_i)}, & \text{if } P_i \leq \frac{R_i^{max} \gamma_i^*(\theta P_T + A_i)}{W + \theta R_i^{max} \gamma_i^*}, \\ R_i^{max}, & \text{otherwise,} \end{cases}$$

where  $\gamma_i^* = \arg \max_{\gamma \geq 1} \{\frac{1}{\gamma} f_i(\gamma)\}$ .

*Proof:* See Appendix A. ■

From Proposition 1 and Equation (2), the utility function of mobile  $i$  at the optimal data rate allocation for a given power allocation,  $P_i$ ,  $U_i(R_i^*(P_i), P_i)$  is expressed as

$$\begin{aligned} U_i^{R_i^*}(P_i) &\triangleq U_i(R_i^*(P_i), P_i) \\ &= \begin{cases} \frac{W}{\gamma_i^*} \frac{P_i}{\theta P_T - \theta P_i + A_i} f_i(\gamma_i^*), & \text{if } P_i \leq \frac{R_i^{max} \gamma_i^*(\theta P_T + A_i)}{W + \theta R_i^{max} \gamma_i^*}, \\ R_i^{max} f_i(\gamma_i(R_i^{max}, P_i)), & \text{otherwise,} \end{cases} \end{aligned} \quad (3)$$

and problem (A) is equivalent to the following optimization problem.<sup>2</sup>

$$\begin{aligned} \text{(B)} \quad &\max \sum_{i=1}^M U_i^{R_i^*}(P_i) \\ &\text{subject to} \quad \sum_{i=1}^M P_i \leq P_T, \\ &0 \leq P_i \leq P_T, \quad i = 1, 2, \dots, M. \end{aligned}$$

In problem (B), the problem is reduced to a power allocation problem from the original power and rate allocation problem. After obtaining the appropriate power allocation by solving problem (B), the optimal transmission data rate can be obtained via Proposition 1.

*Remark 2: In Equation (3),  $U_i^{R_i^*}(P_i)$  is a convex function for  $P_i \leq \frac{R_i^{max} \gamma_i^*(\theta P_T + A_i)}{W + \theta R_i^{max} \gamma_i^*}$  and the shape of  $U_i^{R_i^*}(P_i)$  follows the shape of  $f_i$  for  $\frac{R_i^{max} \gamma_i^*(\theta P_T + A_i)}{W + \theta R_i^{max} \gamma_i^*} < P_i \leq P_T$ . Therefore,  $U_i^{R_i^*}(P_i)$  is a convex function, or a sigmoidal-like function of  $P_i$ , since by assumption,  $f_i$  is a convex function, a concave function, or a sigmoidal-like function of  $P_i$ .*

Since  $U_i^{R_i^*}$  could be a sigmoidal-like function, which is the more interesting case, in general, it is not easy to obtain an optimal solution for problem (B). However, it turns out that the structure of problem (B) is similar to that of the problem studied in [14]. In [14], a power allocation algorithm that provides a power allocation that is Pareto-optimal as well as a good approximation of the global optimal power allocation

<sup>2</sup>A mobile with the fixed data rate  $R^f$  can be easily accommodated in this problem, if we assume that  $R_i^*(P_i) = R^f$  for  $0 \leq P_i \leq P_T$ .

for sigmoidal-like, convex, and concave utility functions. Further, it is proved that the power allocation is asymptotically optimal in [15]. We can hence exploit the power allocation strategy in [14] to obtain a Pareto-optimal and asymptotically optimal power allocation and obtain the optimal data rate for the power allocation using Proposition 1. In the following, we briefly describe the power allocation algorithm. We refer readers to [14], [15] for more details.

The power allocation algorithm is a dynamic pricing based algorithm. The base-station broadcasts  $\lambda$ , the price per unit power to all mobiles that communicate with it. Based on  $\lambda$ , each mobile  $i$  requests a power level  $P_i(\lambda)$  that maximizes its net utility, i.e.,

$$P_i(\lambda) = \arg \max_{0 \leq P \leq P_T} \{U_i^{R^*}(P) - \lambda P\}. \quad (4)$$

Based on the amount of power requested by all mobiles, the base-station updates  $\lambda$  to maximize the total system utility. The algorithm consists of two stages, the mobile selection stage and the power allocation stage. At the mobile selection stage, mobiles are selected by the mobile selection algorithm:

#### Mobile Selection Algorithm

- (a) If  $P_1(\lambda_1^{max}) = P_T$ , mobile 1 is selected.
- (b) If  $\sum_{j=1}^{k-1} P_j(\lambda_{k-1}^{max}) < P_T$  and  $\sum_{j=1}^{k-1} P_j(\lambda_k^{max}) \geq P_T$ , mobiles 1, 2,  $\dots$ ,  $k-1$  are selected.
- (c) If  $\sum_{j=1}^{k-1} P_j(\lambda_k^{max}) < P_T$  and  $\sum_{j=1}^k P_j(\lambda_k^{max}) > P_T$ , mobiles 1, 2,  $\dots$ ,  $k-1$  are selected.
- (d) If  $\sum_{j=1}^M P_j(\lambda_M^{max}) \leq P_T$ , mobiles 1, 2,  $\dots$ ,  $M$  are selected.

where  $\lambda_i^{max}$  is the maximum willingness to pay per unit power of mobile  $i$  and is defined as

$$\lambda_i^{max} = \arg \min_{0 \leq \lambda \leq \infty} \{ \max_{0 \leq P \leq P_T} \{U_i^{R^*}(P) - \lambda P\} = 0 \},$$

and we assume that  $\lambda_1^{max} > \lambda_2^{max} > \dots > \lambda_M^{max}$ . We call  $\lambda_i^{max}$  the maximum willingness to pay per unit power of mobile  $i$ , since  $P_i(\lambda) > 0$  for  $\lambda \leq \lambda_i^{max}$  and  $P_i(\lambda) = 0$  for  $\lambda > \lambda_i^{max}$ . Therefore, at the mobile selection stage, mobiles are selected in a decreasing order of  $\lambda_i^{max}$  as many as possible satisfying the power constraint. Each mobile  $i$  can calculate  $\lambda_i^{max}$  as:

$$\lambda_i^{max} = \begin{cases} \frac{\partial U_i^{R^*}(P)}{\partial P} \Big|_{P=P^*}, & \text{if } 0 < P_i^o < P_T \text{ and } P^* \text{ exists,} \\ \frac{U_i^{R^*}(P_T)}{P_T}, & \text{otherwise,} \end{cases} \quad (5)$$

where  $P_i^o$  is an inflection point of  $U_i^{R^*}(P)$  and  $P^*$  is a solution of the following equation.

$$U_i^{R^*}(P) - P \frac{\partial U_i^{R^*}(P)}{\partial P} = 0, \quad P_i^o \leq P \leq P_T.$$

After selecting mobiles, the problem is reduced to a convex programming problem. Thus, at the power allocation stage, power is allocated optimally to the selected mobiles using a simple algorithm. In [14], the power allocation algorithm was implemented using a bisection algorithm. It was shown that if a condition in either (a), (b), or (d) of the mobile selection algorithm is satisfied, the power allocation is a *global optimal power allocation*. But, if the condition in (c) of the mobile selection algorithm is satisfied, the power allocation might be different from the global optimal power allocation.

#### IV. STUDY OF THE PROPOSED POWER AND RATE ALLOCATION ALGORITHM

##### A. Mobile selectivity

We study the properties of the mobile selectivity in this subsection. Recall that  $A_i$  is defined by  $I_i/G_i$ , where  $I_i$  is background noise and intercell interference to mobile  $i$  and  $G_i$  is path gain from the base-station to mobile  $i$ ,  $f_i$  is the probability of packet transmission success for mobile  $i$ , and  $R_i^{max}$  is the maximum data rate for mobile  $i$ . Thus,  $A_i$  represents the degree of “goodness” of the transmission environment from the base-station to mobile  $i$ . Smaller  $A_i$  implies that a better transmission environment.

The next proposition shows that if mobile  $i$  is more efficient than mobile  $j$ , then  $\lambda_i^{max} \geq \lambda_j^{max}$  and mobile  $i$  has a greater chance to be selected by the algorithm than mobile  $j$ . This is the case because in the mobile selection algorithm, mobiles are selected in a decreasing order of  $\lambda_i^{max}$ . As pointed out in the previous section, the power allocation is a global optimal power allocation for the selected mobiles. Therefore, with this efficient mobile selection combined with the globally optimal power allocation for the selected mobiles, our algorithm can achieve high efficiency of the system. We first define the efficiency of mobiles.

*Definition 1: Mobile  $i$  is said to be more efficient than mobile  $j$  if  $U_i^{R_i^*}(P) \geq U_j^{R_j^*}(P)$  for  $0 \leq P \leq P_T$ .*

*Proposition 2: If mobile  $i$  is more efficient than mobile  $j$ , then  $\lambda_i^{max} \geq \lambda_j^{max}$ .*

*Proof:* See Appendix B. ■

*Corollary 1: Suppose  $f_i(\gamma) = f_j(\gamma)$  and  $A_i = A_j$ . Then, if  $R_i^{max} > R_j^{max}$ ,  $\lambda_i^{max} \geq \lambda_j^{max}$ .*

*Proof:* See Appendix C. ■

Corollary 1 implies that if other conditions are same, the mobile with a higher maximum data rate has a higher priority to be selected than the mobile with a lower maximum data rate. Corollaries 2 and 3 can be proved in a similar way to Corollary 1.

*Corollary 2: Suppose  $R_i^{max} = R_j^{max}$  and  $A_i = A_j$ . Then, if  $f_i(\gamma) \geq f_j(\gamma)$  for all  $\gamma$ ,  $\lambda_i^{max} \geq \lambda_j^{max}$ .*

In Corollary 2,  $f_i(\gamma) \geq f_j(\gamma)$  implies that mobile  $i$  needs a lower value of  $E_b/I_0$  to achieve the same transmission success probability as mobile  $j$ , i.e., mobile  $i$  has a more efficient transmission scheme than mobile  $j$ . Hence, Corollary 2 implies that if other conditions are same, the mobile with the more efficient transmission scheme has a higher priority to be selected than the mobile with the less efficient transmission scheme.

*Corollary 3: Suppose  $f_i(\gamma) = f_j(\gamma)$  and  $R_i^{max} = R_j^{max}$ . Then, if  $A_i < A_j$ ,  $\lambda_i^{max} \geq \lambda_j^{max}$ .*

In Corollary 3,  $A_i < A_j$  implies that mobile  $i$  is in a better transmission environment than mobile  $j$ . Hence, if other conditions are same, a mobile in a better transmission environment has a higher priority to be selected than a mobile in a worse transmission environment. This also implies that if all mobiles are homogeneous, then mobiles are selected in an increasing order of  $A_i$ , i.e., mobiles are selected according to their transmission environment.

It can be easily shown that the results in this subsection can be applied to the global optimal power and rate allocation. Therefore, in the system with heterogeneous mobiles, the mobile selectivity depends not only on the transmission environment of mobiles but also on the maximum data rate and the transmission scheme of mobiles. This implies that, in general situations in which several conditions are mixed, it is not easy to determine which mobiles must be selected for the global optimal solution. However, our mobile selection that selects mobiles in a decreasing order of  $\lambda_i^{max}$  provides a simple and unified strategy of mobile selection, while providing a good approximation of the global optimal solution.

### B. Properties of the proposed power and rate allocation

In this subsection, we study properties of the proposed power and rate allocation. The next proposition gives us the optimal transmission data rate of the selected mobiles.

*Proposition 3: Suppose that mobile  $i$  is selected by the mobile selection algorithm, then  $R_i^*$ , the transmission data rate for mobile  $i$  is given by*

$$R_i^* = \begin{cases} \frac{P_T W}{\gamma_i^* A_i} < R_i^{max}, & \text{if } A_i > \frac{P_T W}{R_i^{max} \gamma_i^*}, \\ R_i^{max}, & \text{otherwise.} \end{cases}$$

*Proof:* See Appendix D. ■

Proposition 3 implies that each mobile can compute its transmission data rate when it is selected by the proposed power and rate allocation algorithm before the power allocation procedure, since  $R_i^*$  does not depend on the individual power level  $P_i$ . This is true even though the optimal data rate of a mobile is a function of the level of the power allocation, as shown in Proposition 1. From the proof of Proposition 3, if  $R_i^* < R_i^{max}$ ,  $U_i^{R_i^*}(P)$  is a convex function and mobile  $i$  is allocated the total transmission power  $P_T$  at price  $\lambda^* = \lambda_i^{max}$ . This implies that if a mobile is selected for transmission, it is always allocated the data rate either  $R_i^* = R_i^{max}$  that is a constant or  $R_i^* < R_i^{max}$  that requires the total transmission power  $P_T$ . Thus, we can think that  $R_i^*$  does not depend on the individual power level  $P_i$ . Using this property, mobile  $i$  can calculate its transmission data rate  $R_i^*$  after measuring its transmission environment  $A_i$  and it can act like a mobile with a fixed data rate  $R_i^*$  during the power allocation procedure. We also have the next corollary that indicates that the joint power and data rate allocation for a mobile using the proposed algorithm is one of boundary points of the individual constraint set of the mobile.

*Corollary 4: If mobile  $i$  is selected by the mobile selection algorithm and  $R_i^* < R_i^{max}$ , only mobile  $i$  is selected and the total power is allocated to it.*

Corollary 4 implies that at most one mobile can be allocated a data rate that is smaller than its maximum data rate. Further, if there exists such a mobile, other mobiles cannot be selected. Therefore, if we assume that mobiles are selected sequentially (in a decreasing order of the maximum willingness to pay) in the proposed scheme, a new mobile can be selected, only when both the mobiles that are already selected and the new mobile can be allocated their maximum data rates.

A similar result can be shown for the global optimal power and rate allocation.

*Proposition 4: If  $\bar{P}^* = (P_1^*, \dots, P_M^*)$  is a global power allocation, then at most one mobile achieves utility value in the convex region of the utility function among the selected mobiles for transmission.*

*Proof:* See Appendix E. ■

Proposition 4 implies that at the global optimal power and rate allocation, at most one mobile can be allocated a data rate that is smaller than its maximum data rate (similar to the case in our scheme), since at the concave region of the utility function, a mobile is allocated its maximum data rate. If some mobile, that is allocated a smaller data rate than its maximum data rate, is selected for transmission along with other mobiles, we can think that this mobile utilizes power most inefficiently among the selected mobiles. This has to be true, otherwise, by taking some power from the other mobiles and allocating it to the mobile, we can increase its

data rate up to the maximum data rate and improve the overall system performance. Hence, if we assume that mobiles are selected for transmission sequentially (in a decreasing order of the efficiency of power utilization) at the global optimal power and rate allocation, a new mobile can be selected, only when mobiles that are already selected can be allocated their maximum data rate as in our scheme. But in the global optimal scheme, a new mobile can be allocated a data rate that is smaller than its maximum data rate, which is not allowed in our scheme.

Therefore, we can infer that if we have a situation when a mobile is in the convex region at the global optimal power and rate allocation, the proposed algorithm does not allocate power to the mobile but distributes power to be allocated to the mobile to the other selected mobiles optimally. As pointed out before, in general, the mobile in the convex region at the global optimal power and rate allocation utilizes power less efficiently than the other selected mobiles and the utility value in the convex region is relatively small. But finding the mobile in the convex region at the global optimal power and rate allocation is not easy. *Hence, by ignoring the mobile that is relatively inefficient, we can obtain a good approximation of the global power and rate allocation with the simple proposed algorithm.*

### C. Study of the optimality conditions for TDMA and CDMA type of transmissions

In this subsection, we study the optimality conditions of pure TDMA type of transmission that transmits to only one mobile at a time and pure CDMA type of transmission that transmits to all mobiles simultaneously. The next proposition gives us the optimal conditions for the TDMA type of transmission and the CDMA type of transmission.

*Proposition 5:* (a) *If  $P_1(\lambda_2^{max}) \geq P_T$ , where  $\lambda_1^{max} > \lambda_2^{max} > \dots > \lambda_M^{max}$ , then selecting only mobile 1 and allocating the total transmission power to it is an optimal strategy.*

(b) *If  $\sum_{i=1}^M P_i(\lambda_M^{max}) \leq P_T$ , where  $\lambda_1^{max} > \lambda_2^{max} > \dots > \lambda_M^{max}$ , then selecting all mobiles and transmitting to all mobiles simultaneously is an optimal strategy.*

*Proof:* See Appendix F. ■

Note that the optimality conditions in Proposition 5 depend not only on system parameters that are constant, but also on the transmission environment of each mobile (time varying and location dependent) and the maximum data rate of each mobile (also time varying from the point of view of the system). Moreover, selecting a subset of mobiles at a time and transmitting only to them can be an optimal strategy depending

on the system status. This implies that a static strategy for transmission could be inefficient in some cases and to obtain a high efficiency of the system, we need a strategy that can be adapted to dynamic characteristics of the system, such as the scheme proposed in this paper.

From Proposition 5 (a), we can also infer that if  $R_i^{max} > \frac{P_T W}{A_i \gamma_i^*}$  for all mobiles (i.e., if mobiles have a high enough maximum data rate or the transmission environment of mobiles is bad because of high interference or high channel loss), then pure TDMA type of transmission is optimal. In this case,  $R_i^* < R_i^{max}$  for all mobiles by Proposition 3 and  $U_i^{R_i^*}(P)$  is a convex function for each mobile  $i$  that has  $P_i(\lambda_i^{max}) = P_T$ . Therefore, by Proposition 5 (a), selecting only one mobile and allocating the total transmission power to it is an optimal strategy. This also implies that if there is no constraint on the maximum data rate for mobiles, then TDMA type of transmission is always optimal regardless of transmission environment of mobiles as shown in [3], [4].

## V. BASE-STATION ASSIGNMENT

In the previous sections, we considered a power and rate allocation problem by focusing on one cell of the system. In this section, we consider a multi-cellular system and it brings us two additional problems that were not considered in the previous sections. One is the total transmission power allocation problem for each base-station and the other is the base-station assignment problem for each mobile.

To maximize the improvement of the system performance, the power allocation for base-stations must be done considering the status of each cell and the channel state of each mobile. But, in practice, this would require a very complex power allocation algorithm, since it needs the cooperation among base-stations (at least among adjacent base-stations). In addition, each base-station may require information not only from mobiles in its cell, but also from mobiles in other cells, resulting in a significant amount of signaling cost. For instance, Oh *et al.* extend their power control and spreading gain allocation algorithm for a single cell system [1] to a multi-cellular system [2]. To perform the algorithm, each base-station needs to know the detailed information of the status of the mobiles in its own cell and adjacent cells. However, even with this significant amount of information, the algorithm is not guaranteed to converge to the optimal allocation and requires much longer convergence time than the algorithm for the single cell system.

Therefore, in this paper, we adopt a strategy that each base-station tries to maximize its total system utility without considering the status of other cells. Using this strategy, each base-station performs the power allocation algorithm in the previous section by independently transmitting at its maximum power

level, which makes the algorithm simple. This is typically called a non-cooperative situation and results in a Nash equilibrium operating point, which might be *inefficient*, i.e., there might exist another total power allocation with which each mobile can obtain higher utility than the Nash equilibrium power allocation. If the load of each cell is unbalanced, the inefficiency might be large. For example, if one cell is heavily loaded and an adjacent cell is lightly loaded, by decreasing the total transmission power of the base-station in the lightly loaded cell, interference for mobiles in the heavily loaded cell can be decreased. This results in a decrease of the utility of the lightly loaded cell and an increase in the utility of the heavily loaded cell. In such a case, generally, the increase can be larger than the decrease in utility and hence the total system utility could be higher than the total system utility at the Nash equilibrium. To cope with this situation, we propose a base-station assignment algorithm. By reassigning some mobiles in the heavily loaded cells to lightly loaded cells, we can expect to improve the total system utility by balancing the load among the base-stations. We next describe the proposed base-station assignment algorithm.

Each base-station performs the power and data rate allocation algorithm in the previous section independently. For the base-station assignment, we use the results of the power and data rate allocation algorithm, which is a pricing based algorithm. Thus, we call our base-station assignment the *pricing based base-station assignment*. We first define some variables.

- $S^c(b)$ : Set of mobiles that are not selected for transmission by base-station  $b$  among mobiles that communicate with base-station  $b$ .
- $\lambda_i^{max}(b)$ : Maximum willingness to pay per unit power of mobile  $i$  for base-station  $b$ .
- $\lambda^*(b)$ : Equilibrium price per unit power for base-station  $b$ .
- $w(b) = \begin{cases} \max_{i \in S^c(b)} \{\lambda_i^{max}(b)\}, & \text{if } S^c(b) \neq \phi, \\ 0, & \text{otherwise.} \end{cases}$

Also, recall that  $A_i(b)$  indicates the transmission environment between base station  $b$  and mobile  $i$ . Smaller  $A_i(b)$  indicates a better transmission environment.

Assume that mobile  $i$  is currently communicating to base-station  $b$ . If mobile  $i$  is selected for transmission by base-station  $b$ , it continues to communicate to base-station  $b$  in the next time slot. But, if mobile  $i$  is not selected for transmission by base-station  $b$ , at the next time slot, it selects and connects to a base-station  $d$  that satisfies

$$d = \arg \max_{k \in g(i)} \{\lambda_i^{max}(k) - \lambda^*(k)\}, \quad (6)$$

where

$$g(i) = \{k \in \mathbf{B} | \lambda_i^{max}(k) \geq w(k)\} \quad (7)$$

and  $\mathbf{B}$  is the set of base-stations. Note that  $\mathbf{B}$  may not be the set of all base-stations. It could be a subset of all base-stations. For example,  $\mathbf{B}$  could be the set of adjacent base-stations of base-station  $b$  or the set of base-stations from which mobile  $i$  receives pilot signal with enough strength. If  $w(b) > \lambda_i^{max}(b)$ , mobile  $i$  cannot be selected by base-station  $b$  and in Equation (7), we exclude base-station  $b \in \mathbf{B}$  from  $g(i)$  whose  $w(b)$  is larger than  $\lambda_i(b)$ . Therefore,  $g(i)$  can be thought as the set of candidate base-stations by which mobile  $i$  can be selected. By Equation (6), the mobile selects a base-station that has the maximum value of the maximum willingness to pay for the base-station minus the equilibrium price of the base-station among the candidate base-stations.

We can explain the intuition of the proposed algorithm using two interpretations. First, from the point of view of pricing, we can think that the mobile selects the base-station which gives it the highest profit per unit power. Since  $\lambda_i^{max}(b)$  can be interpreted as the maximum value for unit power of mobile  $i$  at base-station  $b$  and  $\lambda^*(b)$  can be interpreted as the current price for unit power at base-station  $b$ ,  $\lambda_i^{max}(b) - \lambda^*(b)$  can be interpreted as the profit per unit power that mobile  $i$  can obtain if it is selected by base-station  $b$ . Another interpretation is that the mobile selects the base-station with which it has the relatively best transmission environment considering the congestion level of the base-station. The next proposition tells us that  $\lambda_i^{max}(b)$  is an indicator of the transmission environment between base-station  $b$  and mobile  $i$ .

*Proposition 6: If  $\lambda_i^{max}(n) < \lambda_i^{max}(m)$ , then  $A_i(n) > A_i(m)$ .*

*Proof:* See Appendix G. ■

By the definition of  $A_i(b)$ ,  $\lambda_i^{max}(n) < \lambda_i^{max}(m)$  implies that the transmission environment between base-station  $m$  and mobile  $i$  is better than the transmission environment between base-station  $n$  and mobile  $i$ . Thus,  $\lambda_i^{max}(b)$  can be interpreted as the “goodness” of the transmission environment of mobile  $i$  when it is connected to base-station  $b$ . Also, we can interpret  $\lambda^*(b)$  as being the congestion level at base-station  $b$  from the following propositions.

*Proposition 7: Suppose mobile  $i$  is selected by base-station  $n$  and mobile  $j$  is selected by base-station  $m$ . Further suppose that  $f_i = f_j$ ,  $R_i^{max} = R_j^{max}$ , and  $A_i(n) = A_j(m)$ . Then, if  $\lambda_i^*(n) < \lambda_j^*(m)$ ,  $P_i(\lambda^*(n)) \geq P_j(\lambda^*(m))$ .*

*Proof:* See Appendix H. ■

What Proposition 7 tells us is that if two mobiles are in the same condition and they are selected by different base-stations, a mobile in the base-station with a lower equilibrium price is allocated more power than a mobile in the base-station with a higher equilibrium price. In other words, the base-station with a higher equilibrium price has more demand for power than the base-station with a lower equilibrium price. Thus, we can interpret  $\lambda^*(b)$  as being the congestion level of base-station  $b$ . Hence,  $\lambda_i^{max}(b) - \lambda^*(b)$  in Equation (6) can be interpreted as the relative “goodness” of the transmission environment between mobile  $i$  and base-station  $b$ , taking into account the congestion level of the cell. From this interpretation, Equation (6) implies that mobile  $i$  is reassigned to base-station  $d$  which has the relatively best transmission environment considering the congestion level of the cell among the candidate base-stations from which it can be selected.

To perform the base-station assignment algorithm in Equations (6) and (7), mobile  $i$  must calculate  $\lambda_i^{max}(b)$  and know  $\lambda^*(b)$  and  $w(b)$  for  $b \in \mathbf{B}$ . Mobile  $i$  can calculate  $\lambda_i^{max}(b)$ , if it knows  $A_i(b)$  and  $P_T$  and  $A_i(b)$  can be measured at mobile  $i$ . Hence, the information that mobile  $i$  needs to perform the algorithm are  $P_T$ ,  $\lambda^*(b)$  and  $w(b)$  for  $b \in \mathbf{B}$  and these parameters must be transmitted from base-station  $b$ . But this is base-station-specific information not mobile-specific information and this can be broadcasted by each base-station. In contrast, for the algorithms in [5], [6], [7], each mobile needs to know some mobile-specific information that must be transmitted from the base-stations that are candidates for the assignment, requiring each mobile to maintain its own signaling channels with several base-stations. Therefore, our algorithm needs less signaling costs than the algorithms in [5], [6], [7].

## VI. NUMERICAL RESULTS

In this section, we provide numerical results of our power and rate allocation algorithm and base-station assignment algorithm by computer simulations. We also compare the performance of our algorithms with the performance of Qualcomm’s HDR [8], [9], which is proposed for downlink transmission for high data rate services. In HDR, each base-station transmits a pilot signal. Each mobile measures its received SIR from the pilot signals of several base-stations and is connected to the base-station with the highest SIR. At each base-station, for downlink transmission, the base-station transmits at the maximum power, if there is at least one mobile to transmit to, without considering the status of other cells. The downlink transmissions are time multiplexed and at each time slot, each base-station transmits to only one mobile. Therefore, ignoring fairness, the optimal strategy that maximizes the throughput in HDR is to transmit to a mobile that can achieve the highest expected throughput. We adopt this strategy to obtain the performance of HDR, since in this paper, we do not consider fairness.

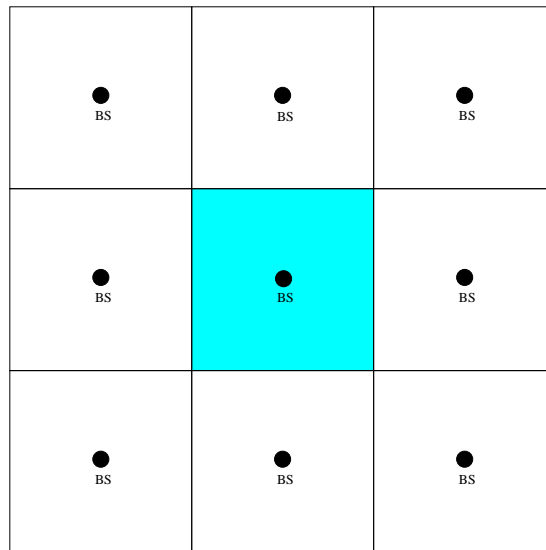


Fig. 2. Cellular network model.

TABLE I

PARAMETERS FOR THE SYSTEM.

Maximum power ( $P_T$ )	10
Chip rate ( $W$ )	100000
Distance loss exponent ( $\alpha$ )	4
Variance of log-normal distribution ( $\sigma^2$ )	8
Length of the side of the cell	1000

For the simulation, we model the cellular network with 9 square cells as shown in Fig. 2. We assume that the base-station is located at the center of each cell. We model the path gain from a base-station  $i$  to a mobile  $j$ ,  $G_{i,j}$  as

$$G_{i,j} = \frac{K_{i,j}}{d_{i,j}^\alpha}, \quad (8)$$

where  $d_{i,j}$  is the distance from the base-station  $i$  to mobile  $j$ ,  $\alpha$  is a distance loss exponent and  $K_{i,j}$  is the log-normally distributed random variable with mean 0 and variance  $\sigma^2$  (dB), which represents shadowing [16]. The parameters for the system are summarized in Table I. For the simulation, we use a sigmoid function to represent  $f_i(\gamma)$ . The sigmoid function is expressed as

$$f_i(\gamma) = c_i \left\{ \frac{1}{1 + e^{-a_i(\gamma - h_i)}} - d_i \right\}, \quad (9)$$

where we set  $c_i = \frac{1 + e^{a_i h_i}}{e^{a_i h_i}}$  and  $d_i = \frac{1}{1 + e^{a_i h_i}}$  for the normalization.

TABLE II

THE PROBABILITY OF PACKET TRANSMISSION SUCCESS ( $a = 3, h = 3.5$ ).

$\gamma(dB)$	5	6	7	8
$f(\gamma)$	0.2662	0.8089	0.9894	0.9998

We first present the performance of the power and rate allocation algorithm without considering base-station assignment. We focus on the cell at the center of the system and assume that the base-stations in other cells transmit at the maximum power limit. In Tables III - V, we assume that there are two classes of mobiles and set the orthogonality factor,  $\theta = 1$ . In Tables III - IV, we have 10 mobiles that are located independently according to a uniform distribution and mobiles in each class are generated with probability 0.5.

In Table III, each class is assumed to have the same  $f_i(\gamma) = f(\gamma)$  but different  $R_i^{max}$ . For  $f(\gamma)$ , we set  $a = 3$  and  $b = 3.5$  and the values of it for several  $\gamma$ 's are provided in Table II. We provide simulation results by setting  $R_2^{max} = 6250$  ( $N_2^{min} = 16$ ) and varying  $R_1^{max}$  ( $N_1^{min}$ ). First, the selection ratio of mobiles of each class is provided. The results in Table III indicate that the class with the higher  $R_i^{max}$  has a higher selection ratio of mobiles than the class with the lower  $R_i^{max}$ . This implies that the proposed power and rate allocation gives a higher priority to be selected to mobiles with the higher maximum data rate, as proved in Corollary 1. We also provide the system utilities obtained by the proposed power and rate allocation and HDR, and the ratio of their utilities. When  $R_1^{max} = 25000$ , these two algorithms give almost the same utility. This is because if there exist mobiles with high maximum data rate, the optimal strategy is to select one mobile and

TABLE III

COMPARISON OF PERFORMANCES OF OUR POWER AND RATE ALLOCATION AND HDR.

$R_1^{max}$ ( $N_1^{min}$ )	1562.5 (64)	3125 (32)	6250 (16)	12500 (8)	25000 (4)
Selection ratio of class 1	0.50052	0.427702	0.388386	0.347643	0.197597
Selection ratio of class 2	0.567581	0.509915	0.391784	0.121819	0.0197844
Utility (Our algorithm)	21324.7	22298.9	24085.7	25379.3	25307.1
Utility (HDR)	6243.77	6247.18	6250	12485.8	24919
Utility (Our algorithm)/Utility (HDR)	3.41536	3.56944	3.85372	2.03265	1.01557

TABLE IV

COMPARISON OF PERFORMANCES OF OUR POWER AND RATE ALLOCATION AND HDR.

$h_1$	2.5	3.0	3.5	4.0	4.5
Selection ratio of class 1	0.565704	0.471902	0.391296	0.306037	0.230125
Selection ratio of class 2	0.288467	0.320855	0.388662	0.448649	0.484367
Utility (Our algorithm)	26222.5	24570.6	24077.7	23196.7	22031.6
Utility (HDR)	6250	6250	6250	6250	6250
Utility (Our algorithm)/Utility (HDR)	4.1956	3.93129	3.85243	3.71148	3.52505

allocate the total transmission power to that mobile. However, as  $R_1^{max}$  decreases, the utility achieved by our power and rate allocation becomes much higher than that of HDR. The reason for this is that if  $R_1^{max}$  is not high, the optimal strategy is to select several mobiles and transmit to them simultaneously. From the results, we know that HDR is optimized only for high data rate services and it could be very inefficient, if there exist heterogeneous services, while our algorithm can be used in any cases with high efficiency.

In Table IV, numerical results are provided by varying  $h_1$ . We set  $R_1^{max} = R_2^{max} = 6250$  ( $N_1^{min} = N_2^{min} = 16$ ),  $a_1 = a_2 = 3$ , and  $h_2 = 3.5$ . As presented in this table, the class with the lower value of  $h_i$  has a higher selection ratio than the class with the higher value of  $h_i$ . By Equation (9), the mobile with the lower  $h_i$  needs the lower SIR to achieve the same probability of packet transmission success than the mobile with the higher  $h_i$ . Hence, the former has the more efficient transmission scheme and a higher priority to be selected than the latter, as proved in Corollary 2. We also compare the performance of the proposed power and rate allocation and HDR. From the results, we know that as  $h_1$  decreases the difference between the two algorithms gets larger. Hence, the results indicates that as mobiles get more efficient, the base-station can select more mobiles and transmit them simultaneously to improve the system throughput.

In table V, we assume that each class has the same maximum data rate,  $R_i^{max} = 6250$ , and the same function for the probability of packet transmission success,  $f_i = f$ . We set  $a = 3$  and  $b = 3.5$ . But we divide the cell into two regions: an inner region that is a square at the center of the cell whose length of the side is a half of the length of the side of the cell and an outer region that is the remaining region of the cell. We say that mobiles in the inner region belong to class 1 and the mobiles in the outer region belong to class 2. Thus, a mobile in class 1 is closer to the base-station than a mobile in class 2 and, in general, the former is in a better transmission environment than the latter. As we studied in Corollary 3, the comparison of selection

TABLE V

COMPARISON OF PERFORMANCES OF OUR POWER AND RATE ALLOCATION AND HDR.

Ratio of class 1	0.2	0.4	0.6	0.8
Selection ratio of class 1	0.981354	0.849091	0.653282	0.4994
Selection ratio of class 2	0.02528	0.00391738	0.000125534	0
Utility (Our algorithm)	13328.8	21279.4	24450.7	24877.3
Utility (HDR)	6141.44	6242.9	6250	6250
Utility (Our algorithm)/Utility (HDR)	2.17031	3.40857	3.91211	3.98037

TABLE VI

COMPARISON OF PERFORMANCES OF OUR POWER AND RATE ALLOCATION AND HDR.

$\theta$	0.2	0.4	0.6	0.8	1.0
Utility (Our algorithm)	92914.3	63127.7	49559.4	44446.7	25000
Utility (HDR)	25000	25000	25000	25000	25000
Utility (Our algorithm)/Utility (HDR)	3.71657	2.52511	1.98238	1.77787	1

ratio indicates that mobiles in a better transmission environment have a higher priority to be selected. The comparison of the performance of our scheme and that of HDR shows that as the number of mobiles in a better transmission environment increases, our scheme more outperforms a scheme like HDR. This is because mobiles in a better transmission environment need less power to achieve the desired performance and, thus, as the number of mobiles in a better transmission environment increases, the base-station can transmit to more mobiles simultaneously.

In table VI, we assume that a single class of mobiles are in the cell. We set  $R^{max} = 25000$ ,  $a = 3$ , and  $b = 3.5$ . We provide the result varying  $\theta$ , the orthogonality factor of the system. As we have a smaller orthogonality factor, mobiles have less intracell interference and the base-station can transmit to more mobiles simultaneously. This is indicated by the result that our scheme more outperforms HDR, as the orthogonality factor gets smaller.

We now present the performance of the proposed base-station assignment algorithm. Recall that in the proposed base-station assignment algorithm, power and rate are allocated using our power and rate allocation in the previous section. We call this *Algorithm A*. We also compare it with two other algorithms. One is called

*Algorithm B*, in which the power and rate are allocated using the proposed power and rate allocation, but each mobile is assigned to the base-station with the highest SIR, as in HDR. In this case, mobile  $i$  is assigned to base-station  $d$  such that

$$d = \arg \min_{b \in \mathbf{B}} \{A_i(b)\}. \quad (10)$$

The comparison of these two algorithms shows us the performance gain of the pricing based base-station assignment algorithm over the SIR based base-station assignment algorithm. The other algorithm that we compare with is the one in HDR, in which each mobile is assigned to the base-station using Equation (10), and each base-station transmits with full power to only one mobile that can receive the highest data rate. We assume that there are two classes of mobiles and set  $a_1 = a_2 = 3$ ,  $h_1 = h_2 = 3.5$ ,  $R_1^{max} = 12500$  ( $N_1^{min} = 8$ ), and  $R_2^{max} = 1562.5$  ( $N_2^{min} = 64$ ). Mobiles in class 1 are generated with probability 0.2 and mobiles in class 2 with probability 0.8. For the simulation, we model hot spot situation as follows. First,  $M_u$  mobiles are located in each cell with uniform distribution. We call these mobiles *balanced mobiles*. Thus, there exist a total of  $9 \times M_u$  *balanced mobiles* in the system. *Balanced mobiles* are generated in the system one by one in each cell in a sequential order, while trying to preserve a balanced load for each cell. Next,  $M_h$  mobiles are located one by one in the center cell with uniform distribution. We call these mobiles *hot spot mobiles*. Therefore, a total of  $9 \times M_u + M_h$  mobiles are in the system with  $M_u + M_h$  mobiles in the center cell and  $M_u$  mobiles in each of other cells. In addition, we assume that  $\sigma^2 = 0$  in Equation (8), which implies that at the SIR based base-station assignment, each mobile is assigned the closest base-station from it.

In Figs. 3–5, we compare the system utility achieved by each algorithm by varying  $M_u$ . In the way we generate mobiles, if the number of mobiles is less than or equal to 9, each base-station has at most one mobile, and, thus, all three algorithms provide the same results. For *balanced mobiles* (i.e., up to  $9 \times M_u$  mobiles) *Algorithm A* and *Algorithm B* achieve almost the same utility. But, during the hot spot period, *Algorithm A* outperforms *Algorithm B*. This implies that the pricing based base-station assignment outperforms the SIR based base-station assignment at the hot spot situation by reassigning mobiles connected to the more congested base-station to less congested base-stations. As shown in the single cell results, *Algorithm B* outperforms HDR, and, thus, so does *Algorithm A*. In Fig. 6, we provide the ratio of the system utility achieved by *Algorithm A* to that achieved by *Algorithm B* for  $M_u = 1, 2, 3$ , respectively. The only difference between *Algorithm A* and *Algorithm B* is the base-station assignment algorithm. Hence, the difference in the performance of these two algorithms provides the effect of the base-station algorithm for

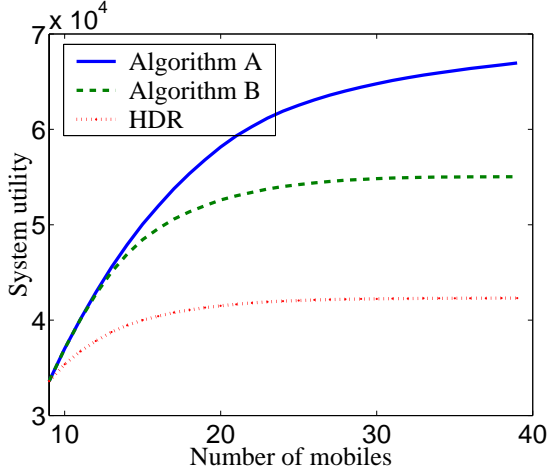
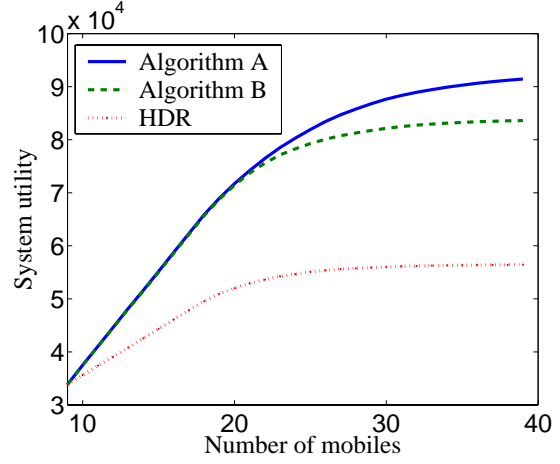
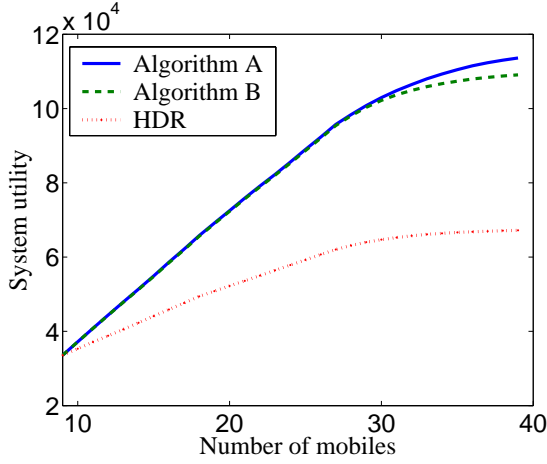
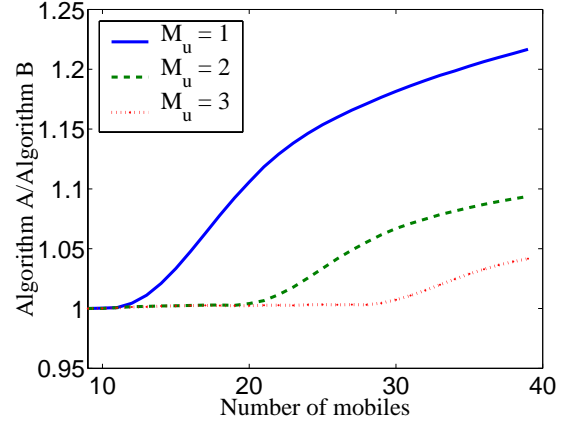
Fig. 3. System utility ( $M_u = 1$ ).Fig. 4. System utility ( $M_u = 2$ ).Fig. 5. System utility ( $M_u = 3$ ).

Fig. 6. The ratio of utility of Algorithm A to that of Algorithm B.

the system performance. Since, as the “unbalancedness” of the system increases (i.e., as  $M_u$  decreases), more mobiles in the more congested cell can be reassigned to less congested cells, the performance gain of *Algorithm A* over *Algorithm B* increases with the “unbalancedness” of the system. The results also implies that balancing the load of each cell by connecting mobiles to the base-station considering the load of each cell and the channel state is more beneficial to increase the system utility than connecting mobiles to the base-station with the best channel state.

We now compare the increment of the system utility during the hot spot period in Figs. 7–9. In these figures, the increment in system utility is the increment in the system utility from the system utility achieved when only *balanced mobiles* exist, (i.e., the system utility achieved by all mobiles minus the system utility when only  $9 \times M_u$  *balanced mobiles* exist). The number of mobiles during the hot spot period,  $M_h$ , means that the number of mobiles added at the center cell after  $9 \times M_u$  *balanced mobiles* are located. Hence

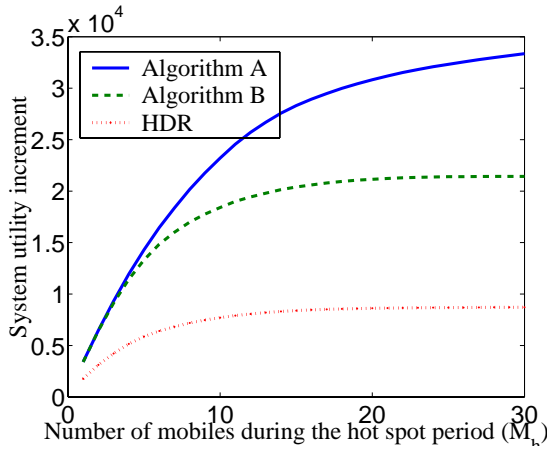


Fig. 7. System utility increment during the hot spot period ( $M_u = 1$ ).

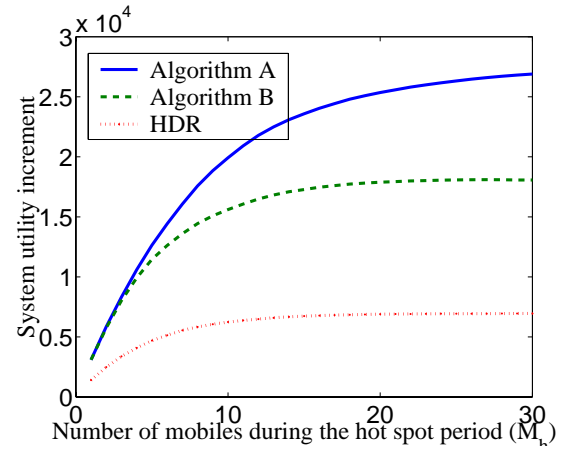


Fig. 8. System utility increment during the hot spot period ( $M_u = 2$ ).

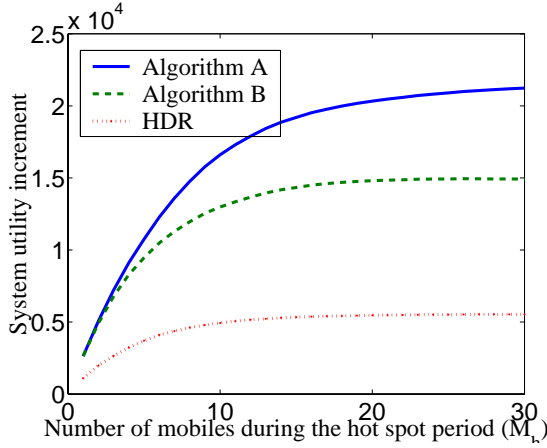


Fig. 9. System utility increment during the hot spot period ( $M_u = 3$ ).

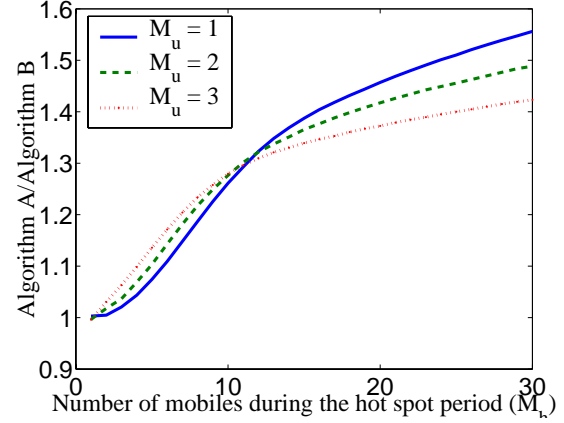


Fig. 10. The ratio of utility increment of Algorithm A to that of Algorithm B during the hot spot period.

when the number of mobiles during the hot spot period is  $M_h$ , the total number of mobiles in the system is  $9 \times M_u + M_h$  and the number of mobiles in the center cell is  $M_u + M_h$ . As shown the figures, *Algorithm A* gives the highest performance while HDR gives the lowest performance.

In Fig. 10, the ratio of the increment in the utility of *Algorithm A* to that of *Algorithm B* is provided varying  $M_u$  to investigate the effect of the base-station assignment algorithm. When  $M_h$  is small, the performance gain of *Algorithm A* over *Algorithm B* is increasing, as  $M_u$  is increasing. In such a case, the hot spot cell with a smaller  $M_u$  is less congested and, thus, it has smaller performance gain. However, when the hot spot cell is congested (i.e.,  $M_h$  is large), the performance gain depends on the congestion level of adjacent cells. As adjacent cells are less congested (i.e., smaller  $M_u$ ), the number of mobiles that can be reassigned to adjacent cells increase and, thus, the performance gain is increasing. Thus, from the results, we know that the performance gain of the proposed base-station assignment algorithm over the SIR based base-station

assignment depends not only on the congestion level of the cell itself but also on the difference of the congestion levels among cells.

## VII. CONCLUSION

In this paper, we study the joint resource allocation and base-station assignment problem for the downlink in CDMA networks with heterogeneous data services. The resource allocation part of the problem that we consider is a joint power and rate allocation problem intended to maximize the total expected throughput with constraints on the total transmission power for the base-station and the maximum data rate for each mobile. We propose an asymptotically optimal power and rate allocation algorithm which is based on a utility and pricing based algorithm. For the downlink of an idealized system with homogeneous mobiles without a constraint on the maximum data rate, it is well known that transmitting only one mobile in the best transmission environment at a time is an optimal strategy. However, we show that in the presence of heterogeneous mobiles and a constraint on the maximum data rate (both of are more practically relevant), this may not be true. We show that the proposed algorithm in this paper achieves high system throughput in a various different situations. For the base-station assignment problem, we propose a pricing based base-station assignment algorithm that uses information from the proposed power and rate allocation algorithm, which results in joint resource allocation and base-station assignment. The proposed algorithm assigns a base-station to a mobile taking into account the congestion level of the cell as well as the transmission environment between the base-station and the mobile. The numerical results show that the proposed algorithm outperforms the SIR based base-station assignment algorithm that considers only the transmission environment between a base-station and a mobile. This implies that load balancing among cells as well as the transmission environment between the base-station and the mobile is quite important to improve the system performance.

## APPENDIX

### A. Proof of Proposition 1

First, suppose that  $R_i^{max} = \infty$ , i.e.,  $N_i^{min} \triangleq W/R_i^{max} = 0$ . Then, by Proposition 1 in [1],

$$R_i^*(P_i) = \frac{WP_i}{\gamma_i^*(\theta P_T - \theta P_i + A_i)},$$

where  $\gamma_i^* = \arg \max_{\gamma \geq 1} \{\frac{1}{\gamma} f_i(\gamma)\}$ . In this paper, we have a constraint on the maximum data rate,  $R_i^{max}$ .

Therefore, considering the constraint

$$R_i^*(P_i) = \begin{cases} \frac{WP_i}{\gamma_i^*(\theta P_T - \theta P_i + A_i)}, & \text{if } \frac{WP_i}{\gamma_i^*(\theta P_T - \theta P_i + A_i)} \leq R_i^{max}, \\ R_i^{max}, & \text{otherwise,} \end{cases}$$

and the equation above is equivalent to

$$R_i^*(P_i) = \begin{cases} \frac{WP_i}{\gamma_i^*(\theta P_T - \theta P_i + A_i)}, & \text{if } P_i \leq \frac{R_i^{max} \gamma_i^*(\theta P_T + A_i)}{W + \theta R_i^{max} \gamma_i^*}, \\ R_i^{max}, & \text{otherwise.} \end{cases}$$

### B. Proof of Proposition 2

Let us define

$$\begin{aligned} w_i(\lambda) &= \max_{0 \leq P \leq P_T} \{U_i^{R_i^*}(P) - \lambda P\}, \\ w_j(\lambda) &= \max_{0 \leq P \leq P_T} \{U_j^{R_j^*}(P) - \lambda P\}. \end{aligned}$$

Then,  $w_i(\lambda) \geq w_j(\lambda)$  by the definition of the mobile efficiency and they are non-increasing functions of  $\lambda$ .

Therefore,

$$w_i(\lambda_j^{max}) \geq w_j(\lambda_j^{max}) = 0$$

and this implies that

$$\lambda_i^{max} \geq \lambda_j^{max},$$

since we must have  $w_i(\lambda_i^{max}) = 0$ .

### C. Proof of Corollary 1

Since  $A_i = A_j$ ,  $\gamma_i(R, P) = \gamma_j(R, P)$  and

$$\begin{aligned} U_i^{R_i^*}(P) &= \max_{0 \leq R \leq R_i^{max}} \{R f_i(\gamma_i(R, P))\} \\ &\geq \max_{0 \leq R \leq R_j^{max}} \{R f_i(\gamma_i(R, P))\} \\ &= \max_{0 \leq R \leq R_j^{max}} \{R f_j(\gamma_j(R, P))\} \\ &= U_j^{R_j^*}(P), \quad 0 \leq P \leq P_T. \end{aligned}$$

Therefore, mobile  $i$  is more efficient than mobile  $j$  and by Proposition 2, the proof is completed.

#### D. Proof of Proposition 3

Since we assume that mobile  $i$  is selected,  $P_i(\lambda^*) > 0$ , where  $\lambda^*$  is an equilibrium price. Suppose  $A_i > \frac{P_T W}{R_i^{max} \gamma_i^*}$ , then

$$P_T < \frac{R_i^{max} \gamma_i^* (\theta P_T + A_i)}{W + \theta R_i^{max} \gamma_i^*} \quad (11)$$

and by Equation (3),  $U_i^{R_i^*}(P_i)$  is a convex function for  $0 \leq P_i \leq P_T$ . By Equation (5), mobile  $i$  has  $\lambda_i^{max} = \frac{U_i^{R_i^*}(P_T)}{P_T}$  and it has  $P_i(\lambda) = 0$  for  $\lambda > \lambda_i^{max}$  and  $P_i(\lambda) = P_T$  for  $\lambda \leq \lambda_i^{max}$ . Since  $P_i(\lambda^*) > 0$ , mobile  $i$  is allocated power level  $P_T$ . Hence, by Proposition 1 and Equation (11),  $R_i^* = \frac{P_T W}{\gamma_i^* A_i}$  and  $R_i^* < R_i^{max}$ . Now, suppose  $A_i \leq \frac{P_T W}{R_i^{max} \gamma_i^*}$ , then

$$P_T \geq \frac{R_i^{max} \gamma_i^* (\theta P_T + A_i)}{W + \theta R_i^{max} \gamma_i^*}. \quad (12)$$

In this case,  $U_i^{R_i^*}(P_i)$  is a convex function or a sigmoidal-link function for  $0 \leq P_i \leq P_T$ . If  $U_i^{R_i^*}(P_i)$  is a convex function, mobile  $i$  is allocated power level  $P_T$ . Hence, by Proposition 1 and Equation (12),  $R_i^* = R_i^{max}$ . If  $U_i^{R_i^*}(P_i)$  is a sigmoidal-like function, mobile  $i$  is allocated power level  $P_i(\lambda^*)$  that forces  $U_i^{R_i^*}(P_i(\lambda^*))$  to be in the concave region, since  $P_i(\lambda) > 0$  that forces  $U_i^{R_i^*}(P_i(\lambda))$  to be in the convex region does not satisfy the second order necessary condition for the solution of Equation (4),  $\frac{\partial^2 U_i^{R_i^*}(P)}{\partial P^2} < 0$ . This implies that  $P_i(\lambda^*) > \frac{R_i^{max} \gamma_i^* (\theta P_T + A_i)}{W + \theta R_i^{max} \gamma_i^*}$ , since  $U_i^{R_i^*}(P_i)$  is a convex function for  $P_i \leq \frac{R_i^{max} \gamma_i^* (\theta P_T + A_i)}{W + \theta R_i^{max} \gamma_i^*}$  by Equation (3). Therefore, by Proposition 1,  $R_i^* = R_i^{max}$ .

#### E. Proof of Proposition 4

First, in the next lemma, we show that at the global optimal power and rate allocation, the marginal utilities of the selected mobiles are same.

*Lemma 1: If  $\bar{P}^* = (P_1^*, \dots, P_M^*)$  is a global power allocation and  $K$  is the set of the selected mobiles for transmission, then  $\frac{\partial U_i^{R_i^*}(P_i)}{\partial P_i} \Big|_{P_i=P_i^*} = \frac{\partial U_j^{R_j^*}(P_j)}{\partial P_j} \Big|_{P_j=P_j^*}$  for all  $i, j \in K$ .*

*Proof:* We will prove this by using contradiction. Suppose that  $\bar{P}^* = (P_1^*, \dots, P_M^*)$  is a global optimal power allocation,  $P_i^* > 0$ ,  $P_j^* > 0$ , and  $\frac{\partial U_i^{R_i^*}(P_i)}{\partial P_i} \Big|_{P_i=P_i^*} > \frac{\partial U_j^{R_j^*}(P_j)}{\partial P_j} \Big|_{P_j=P_j^*}$ . Thus, we can find  $\delta P$  such that  $\frac{\partial U_i^{R_i^*}(P_i)}{\partial P_i} > \frac{\partial U_j^{R_j^*}(P_j)}{\partial P_j}$  for all  $P_i \in [P_i^*, P_i^* + \delta P]$  and  $P_j \in [P_j^* - \delta P, P_j^*]$ . Hence

$$\int_{P_i^*}^{P_i^* + \delta P} \frac{\partial U_i^{R_i^*}(P_i)}{\partial P_i} dp > \int_{P_j^* - \delta P}^{P_j^*} \frac{\partial U_j^{R_j^*}(P_j)}{\partial P_j} dp.$$

This implies that

$$U_i^{R_i^*}(P_i^* + \delta P) - U_i^{R_i^*}(P_i^*) > U_j^{R_j^*}(P_j^*) - U_j^{R_j^*}(P_j^* - \delta P),$$

$$U_i^{R_i^*}(P_i^* + \delta P) + U_j^{R_j^*}(P_j^* - \delta P) > U_i^{R_i^*}(P_i^*) + U_j^{R_j^*}(P_j^*),$$

and utilities for all other mobiles are unchanged, which is contradiction.  $\blacksquare$

Now, we will prove Proposition 4. We will prove this by using contradiction. Suppose that  $\bar{P}^* = (P_1^*, \dots, P_M^*)$  is a global optimal power allocation,  $P_i^* > 0$ ,  $P_j^* > 0$ , and  $U_i^{R_i^*}(P_i^*)$  and  $U_j^{R_j^*}(P_j^*)$  are in the convex region. Since  $P_i^*$  and  $P_j^*$  are optimal allocation, by Lemma 1

$$\left. \frac{\partial U_i^{R_i^*}(P_i)}{\partial P_i} \right|_{P_i=P_i^*} = \left. \frac{\partial U_j^{R_j^*}(P_j)}{\partial P_j} \right|_{P_j=P_j^*}$$

and we can find  $\delta P$  such that  $U_i^{R_i^*}(P_i)|_{P_i=P_i^*+\delta P}$  and  $U_j^{R_j^*}(P_j)|_{P_j=P_j^*-\delta P}$  are still in the convex region. Hence

$$\int_{P_i^*}^{P_i^*+\delta P} \frac{\partial U_i^{R_i^*}(P_i)}{\partial P_i} dp > \int_{P_j^*-\delta P}^{P_j^*} \frac{\partial U_j^{R_j^*}(P_j)}{\partial P_j} dp$$

since  $\frac{\partial U_i^{R_i^*}(P_i)}{\partial P_i}$  and  $\frac{\partial U_j^{R_j^*}(P_j)}{\partial P_j}$  are increasing functions of  $P_i$  and  $P_j$ , respectively. This implies that

$$U_i^{R_i^*}(P_i^* + \delta P) - U_i^{R_i^*}(P_i^*) > U_j^{R_j^*}(P_j^*) - U_j^{R_j^*}(P_j^* - \delta P),$$

$$U_i^{R_i^*}(P_i^* + \delta P) + U_j^{R_j^*}(P_j^* - \delta P) > U_i^{R_i^*}(P_i^*) + U_j^{R_j^*}(P_j^*),$$

and utilities for all other mobiles are unchanged, which is contradiction.

#### F. Proof of Proposition 5

We first prove Proposition 5(a). If  $P_1(\lambda_2^{max}) \geq P_T$ , condition (a) in the mobile selection algorithm is satisfied or condition (b) in the mobile selection algorithm is satisfied with  $k - 1 = 1$ . Thus, only mobile 1 is selected and allocated the total transmission power to it and this is a global optimal solution.

Now, we prove Proposition 5(b). If  $\sum_{i=1}^M P_i(\lambda_M^{max}) \leq P_T$ , then condition (d) in the mobile selection algorithm is satisfied. In this case, all mobiles are selected and power can be allocated to all mobiles and it is a global optimal solution.

### G. Proof of Proposition 6

This is equivalent to the situation when there exist two mobiles, mobile  $n$  and mobile  $m$ , in the same cell. Mobile  $n$  has  $f_i$ ,  $R_i^{max}$ ,  $\lambda_i^{max}(n)$ , and  $A_i(n)$ . Mobile  $m$  has  $f_i$ ,  $R_i^{max}$ ,  $\lambda_i^{max}(m)$ , and  $A_i(m)$ . By Corollary 3, if  $\lambda_i^{max}(n) < \lambda_i^{max}(m)$ , then  $A_i(n) \geq A_i(m)$ . However,  $A_i(n) \neq A_i(m)$ , since if  $A_i(n) = A_i(m)$ ,  $\lambda_i^{max}(n) = \lambda_i^{max}(m)$ , which is contradiction to the assumption. Therefore,  $A_i(n) > A_i(m)$ .

### H. Proof of Proposition 7

Since mobiles  $i$  and  $j$  are in the same condition,  $P_i(\lambda) = P_j(\lambda)$ . From Equation (4),  $P_i(\lambda)$  is non-increasing as  $\lambda$  is increasing. Therefore,  $P_i(\lambda^*(n)) \geq P_j(\lambda^*(m))$ , since  $\lambda_i^*(n) < \lambda_j^*(m)$ .

## REFERENCES

- [1] S.-J. Oh and K. M. Wasserman, "Optimality of greedy power control and variable spreading gain in multi-class CDMA mobile networks," in *ACM Mobicom'99*, 1999, pp. 102–112.
- [2] S.-J. Oh, T. L. Olsen, and K. M. Wasserman, "Distributed power control and spreading gain allocation in CDMA data networks," in *IEEE Infocom'00*, 2000, pp. 379–385.
- [3] A. Bedekar, S. Borst, K. Ramanan, P. Whiting, and E. Yeh, "Downlink scheduling in CDMA data networks," in *IEEE Globecom'99*, 1999, pp. 2653–2657.
- [4] F. Berggren, S.-L. Kim, R. Jäntti, and J. Zander, "Joint power control and intracell scheduling of DS-CDMA nonreal time data," *IEEE Journal on Selected Areas in Communications*, vol. 19, pp. 1860–1870, 2001.
- [5] S. V. Hanly, "An algorithm for combined cell-site selection and power control to maximize cellular spread spectrum capacity," *IEEE Journal on Selected Areas in Communications*, vol. 13, pp. 1332–1340, 1995.
- [6] R. D. Yates and C.-Y. Huang, "Integrated power control and base station assignment," *IEEE Transactions on Vehicular Technology*, vol. 44, pp. 638–644, 1995.
- [7] C. U. Saraydar, N. B. Mandayam, and D. J. Goodman, "Power control in a multicell CDMA data system using pricing," in *IEEE VTC'00*, 2000, pp. 484–491.
- [8] P. Bender, P. Black, M. Grob, R. Padovani, N. Sindhushayana, and A. Viterbi, "CDMA/HDR: A bandwidth-efficient high-speed wireless data service for nomadic users," *IEEE Communications Magazine*, vol. 38, pp. 70–77, July 2000.

- [9] Y. Jou, “Developments in Third Generation (3G) CDMA Technology,” in *IEEE ISSSTA'00*, 2000, pp. 460–464.
- [10] S. Parkvall, E. Dahlman, P. Frenger, P. Beming, and M. Persson, “The evolution of WCDMA towards higher speed downlink packet data access,” in *IEEE VTC'01 Spring*, 2001, pp. 2287–2291.
- [11] M. Frodigh, S. Parkvall, C. Roobol, P. Johansson, and P. Larsson, “Future-generation wireless networks,” *IEEE Personal Communications*, vol. 8, pp. 10–17, Oct. 2001.
- [12] S. Nanda, K. Balachandran, and S. Kumar, “Adaptation techniques in wireless packet data services,” *IEEE Communications Magazine*, vol. 38, pp. 54–64, Jan. 2000.
- [13] R. D. Yates, “A framework for uplink power control in cellular radio systems,” *IEEE Journal on Selected Areas in Communications*, vol. 13, pp. 1341–1347, 1995.
- [14] J. W. Lee, R. R. Mazumdar, and N. B. Shroff, “Downlink power allocation for multi-class CDMA wireless networks,” in *IEEE Infocom'02*, 2002, pp. 1480–1489.
- [15] J. W. Lee, R. R. Mazumdar, and N. B. Shroff, “Downlink power allocation for multi-class wireless systems,” submitted for publication, 2002.  
(<http://expert.cc.purdue.edu/~lee46/Documents/pa.pdf>).
- [16] G. Stuber, *Principles of Mobile Communication*, Kluwer Academic Publishers, 1996.