Cross-Layer Optimization for OFDM Wireless Networks—Part I: Theoretical Framework

Guocong Song, *Student Member, IEEE* and Ye (Geoffrey) Li, *Senior Member, IEEE*

***Abstract*—In this paper, we provide a theoretical framework for cross-layer optimization for orthogonal frequency division multiplexing (OFDM) wireless networks. The utility is used in our study to build a bridge between the physical layer and the media access control (MAC) layer and to balance the efficiency and fairness of wireless resource allocation. We formulate the cross-layer optimization problem as one that maximizes the average utility of all active users subject to certain conditions, which are determined by adaptive resource allocation schemes. We present necessary and sufficient conditions for utility-based optimal subcarrier assignment and power allocation and discuss the convergence properties of optimization. Numerical results demonstrate a significant performance gain for the cross-layer optimization and the gain increases with the number of active users in the networks.**

***Index Terms*—Cross-layer optimization, efficiency and fairness, orthogonal frequency division multiplexing (OFDM) network, utility function.**

I. Introduction

FAIRNESS and efficiency are two crucial issues in resource allocation for wireless networks. Traditionally, spectral efficiency is evaluated in terms of the aggregate throughput, which is sometimes unfair to those users far away from a base-station or with bad channel conditions. On the other hand, absolute fairness may lead to low bandwidth efficiency. Therefore, an effective tradeoff between efficiency and fairness is desired in wireless resource allocation.

The issues on efficient and fair resource allocation have been well studied in economics, where utility functions are used to quantify the benefit of usage of certain resources. Similarly, utility theory can be used in communication networks to evaluate the degree to which a network satisfies service requirements of users' applications, rather than in terms of system-centric quantities like throughput, outage probability, packet drop rate, power, etc. [1]. In wireline networks, utility and pricing mechanisms have been used for flow control [2], [3], congestion control [4], and routing [5]. In wireless networks, the pricing of uplink power control in code division multiple access (CDMA) has been investigated in [6]–[8]. Utility-based power allocation on CDMA downlinks for voice and data applications has been proposed in [9]–[11]. Utility also offers a tangible metric for network provisioning when application performance is the key concern. In this paper, we use utility not only to tradeoff the fairness and efficiency of resource allocation but also to build a bridge between the physical and the media access control (MAC) layers to achieve cross-layer optimization.

Orthogonal frequency division multiplexing (OFDM) divides an entire channel into many orthogonal narrowband subchannels (subcarriers) to deal with frequency-selective fading and support a high data rate. Furthermore, in an OFDM wireless network, different subcarriers can be allocated to different users to provide a flexible multiuser access scheme [12] and to exploit multiuser diversity. Therefore, we focus on utility-based cross-layer optimization for OFDM wireless networks.

There is plenty of room to exploit the high degree of flexibility of radio resource management in the context of OFDM. Since channel frequency responses are different at different frequencies or for different users, data rate adaptation over each subcarrier, dynamic subcarrier assignment (DSA), and adaptive power allocation (APA) can significantly improve the performance of OFDM systems/networks. Using data rate adaptation [13], [14], the transmitter can send higher transmission rates over the subcarriers with better conditions so as to improve throughput and simultaneously to ensure an acceptable bit-error rate (BER) on each subcarrier. Despite the use of data rate adaptation, deep fading on some subcarriers still leads to low channel capacity.

On the other hand, channel characteristics for different users are almost mutually independent in multiuser environments; the subcarriers experiencing deep fading for one user may not be in a deep fade for other users; therefore, each subcarrier could be in a good condition for some users in a multiuser OFDM wireless network. By dynamically assigning subcarriers, the network can benefit from this multiuser diversity [15]. Resource allocation issues and the achievable regions for multiple access and broadcast channel have been investigated in [16] and [17], respectively, which have proved that the largest data rate region is achieved when the same frequency range is shared with overlap by multiple users in broadcast channels. However, when optimal power allocation is used, from [18], there is only a small range of frequency with overlapping power sharing. Thus, optimal power allocation with dynamic subcarrier (nonoverlap) assignment [19]–[22] can achieve a data transmission rate close to the channel capacity boundary. In [19], the authors have investigated optimal resource allocation in multiuser OFDM systems to minimize the total transmission power while satisfying a minimum rate for each user. In [16], [17], [21], and [22], the authors have discussed fairness using linear priority factors; how to assign those factors, however, is still unknown.

Manuscript received April 10, 2003; revised November 14, 2003; accepted January 28, 2004. The editor coordinating the review of this paper and approving it for publication is F.-C. Zheng. This work was supported by the National Science Foundation under Grant CCR-0121565, by Nortel Networks, and by the U.S. Army Research Laboratory under the Collaborative Technology Alliance Program, Cooperative Agreement DAAD19-01–20–0011.

The authors are with the School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, GA 30332-0250 USA (e-mail: songgc@ece.gatech.edu, liye@ece.gatech.edu).

Digital Object Identifier 10.1109/TWC.2004.843065

1536-1276/$20.00 © 2005 IEEE

To take efficiency and fairness into account, we have investigated resource allocation in OFDM networks based on jointly optimizing the physical layer and MAC layer in [23]–[25].

In this paper, we use utility functions to balance efficiency and fairness and to perform cross-layer optimization for the downlink of OFDM wireless networks. By means of convex analysis, we investigate the properties of the optimal subcarrier allocation and power allocation associated with utility-based optimization. Furthermore, we demonstrate that the utility-based resource allocation naturally balances efficiency and fairness. In summary, we provide a theoretical framework for efficient and fair resource allocation in multiuser frequency-selective fading environments. The rest of this paper is organized as follows. In Section II, we describe the channel model, the general properties of utility functions, the model of subcarrier rate adaptation, and formulate the cross-layer optimization problems. In Sections III and IV, we investigate the optimal subcarrier assignment and power allocation, respectively. Next, in Section V, we prove the convexity of the achievable rate region with continuous frequency assignment. In Section VI, we discuss the efficiency and fairness issues. Finally, we demonstrate the performance improvement of the cross-layer optimization through numerical results in Section VII.

II. Problem Formulation

In this section, we briefly describe the channel model, utility functions, adaptive modulation, and frequency power allocation, and formulate the utility-based cross-layer optimization problems.

A. Multiuser Frequency-Selective Fading Channels

We consider a network with one transmitter (base-station) and M receivers (users), as shown in Fig. 1(a). The complex baseband representation of the impulse response of a wireless channel for user i can be described by

$$h_i(t,\tau) = \sum_k \gamma_{k,i}(t)\delta(\tau - \tau_{k,i})$$

where $\tau_{k,i}$ is the delay of the kth path and $\gamma_{k,i}(t)$ is the corresponding complex amplitude. The $\gamma_{k,i}(t)$'s are assumed to be wide-sense stationary, narrowband, complex Gaussian processes, which are independent for different paths or users. The frequency response of the channel impulse response can be expressed as

$$H_i(f,t) = \int_{-\infty}^{+\infty} h_i(t,\tau)e^{-j2\pi f\tau}\, d\tau = \sum_k \gamma_{k,i}(t)e^{-j2\pi f\tau_{k,i}}. \tag{1}$$

It is assumed that the channel fading rate is slow enough so that the frequency response does not change during an OFDM block. If only instantaneous channel conditions are considered, the channel frequency response corresponding to user i is denoted by $H_i(f)$. Consequently, the M-user frequency-selective broadcast fading channel can be represented as in Fig. 1(b). The quality of each user's channel can be indicated by the signal-to-noise ratio (SNR) function $\rho_i(f)$ when the transmission power density $p(f) = 1$, which is defined as

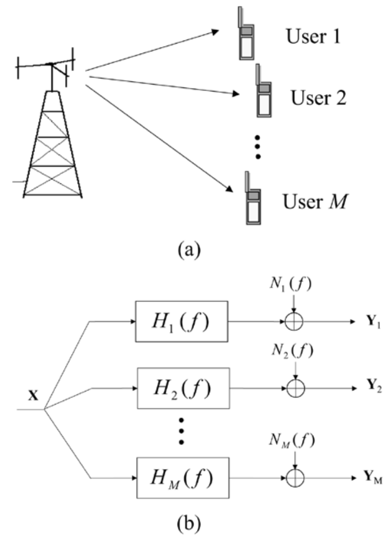

Fig. 1. (a) Downlink of multiuser system. (b) Channel model.

$$\rho_i(f) = \frac{\mid H_i(f) \mid^2}{N_i(f)}$$

where $N_i(f)$ is the noise power density function of user i.

There are many ways to obtain the channel state information at the base station. In a frequency division duplex (FDD) system, using pilot symbols that are inserted in the downlink with a certain time-frequency pattern, the mobile terminals can effectively estimate the channel parameters $H_i(f)$'s and $\rho_i(f)$'s [26] and feed back the $\rho_i(f)$'s to the base station. In a time division duplex (TDD) system, since the symmetry of the channel characteristics for the downlink and uplink, the base station can obtain the channel state information by directly measuring the uplink channels. Therefore, the base station knows the channel state information for the utility-based cross-layer optimization.

B. Rate Adaptation and Power Allocation

Using adaptive modulation [13], [14], the transmitter can send higher data rates over the subcarriers with better channel conditions to improve throughput and simultaneously ensure an acceptable BER in all subcarriers.

Let $c_i(f)$ denote the achievable throughput of user i at frequency f for a given BER and a transmission power density $p(f)$. When continuous rate adaptation is used, $c_i(f)$ can be expressed as [27]

$$\begin{aligned} c_i(f) &= \log_2\left(1 + \frac{\beta p(f) \mid H_i(f) \mid^2}{N_i(f)}\right) \left(\frac{\frac{\text{bits}}{\text{sec}}}{\text{Hz}}\right) \\ &= \log_2(1 + \beta p(f)\rho_i(f)) \end{aligned} \tag{2}$$

where β is determined by

$$\beta = \frac{1.5}{-\ln(5\text{BER})}.$$

Since it indicates the difference between the SNR needed to achieve a certain data transmission rate for a practical system and the theoretical limit, respectively, β is usually called the SNR gap [27].

C. Utility Functions

As indicated earlier, a utility function is used for the cross-layer optimization and balancing the efficiency and fairness. The utility function maps the network resources a user utilizes into a real number. In almost all wireless applications, a reliable data transmission rate is the most important factor to determine the satisfaction of users. Therefore, the utility function $U(r)$ should be a nondecreasing function of the data rate r. In particular, when $U(r) = r$, the utility is just the throughput, which is the objective of most traditional network optimizations. Therefore, our work can be regarded as a general extension of traditional network optimizations.

Utility functions serve as an optimization objective for the adaptive physical and MAC layer techniques. Consequently, it can be used to optimize radio resource allocation for different applications and to build a bridge among the physical, MAC, and higher layers.

When a utility function is used to capture the user's feeling, such as the level of satisfaction for assigned certain resources, it cannot be obtained only through theoretical derivation. In this case, it can be estimated from subjective surveys. For best effort traffic [28], a utility function can be described by

$$U(r) = 0.16 + 0.8\ln(r - 0.3) \tag{3}$$

where r is in unit of k b/s. To prevent assigning too much resource to the user with good channel conditions, the slope of the utility curves decreases with an increase in the data rate. We will discuss more on the issue of fairness and efficiency in Section VI.

D. Formulation of Utility-Based Cross-Layer Optimization

To obtain the performance bound of the cross-layer optimization, we assume that there is an infinite number of orthogonal subcarriers in all frequency resources, or the bandwidth of each orthogonal subcarrier is infinitesimal, which can be regarded as an extreme situation of OFDM. In a practical OFDM system, the minimum granularity of resource allocation is one subcarrier. The OFDM system in which $\triangle f \rightarrow 0$ provides an infinitesimal granularity of resource allocation, thereby presenting the performance upper bound.

Consider a single cell consisting of M users. The entire frequency band $[0, B]$ is divided into M nonoverlapping frequency sets, each for one user. Define D_i as the frequency set assigned to user i. Then

$$\bigcup_{i=1}^{M} D_i \subseteq [0, B] \tag{4}$$

$$D_i \bigcap D_j = \varnothing, \quad i \neq j \tag{5}$$

where $\varnothing$ denotes an empty set. The transmission throughput of user i can be expressed as

$$\begin{aligned} r_i &= \int_{D_i} c_i(f) df \\ &= \int_{D_i} \log_2[1 + \beta p(f)\rho_i(f)] df. \end{aligned} \tag{6}$$

Besides dynamically assigning the frequency sets, the transmission power density at different frequencies can also be adjusted to improve the network performance, with a total transmission power constraint by

$$\frac{1}{B}\int_0^B p(f)\, df \leq 1. \tag{7}$$

Let the utility function of user i be $U_i(\cdot)$. If user i has a data rate r_i, the user's utility is $U_i(r_i)$. The utility-based cross-layer optimization is to assign wireless resources (including frequency band and power density) to maximize the average utility of the network, which can be expressed as

$$\frac{1}{M}\sum_{i=1}^{M} U_i(r_i). \tag{8}$$

In Sections III and IV, we will discuss dynamic subcarrier assignment and power allocation, respectively.

III. DSA

In this section, we investigate DSA to improve the performance of an OFDM-based network when the transmission power is uniformly distributed over the entire available frequency band, that is, $p(f) = 1$, then the achievable throughput at frequency f, $c_i(f)$, can be expressed as

$$c_i(f) = \log_2(1 + \beta\rho_i(f)).$$

Thus, the DSA problem is to maximize

$$\frac{1}{M}\sum_{i=1}^{M} U_i(r_i) = \frac{1}{M}\sum_{i=1}^{M} U_i\left(\int_{D_i} c_i(f)\, df\right) \tag{9}$$

subject to

$$\bigcup_{i=1}^{M} D_i \subseteq [0, B] \tag{10}$$

$$D_i \bigcap D_j = \varnothing, \quad i \neq j \text{ and } i, j = 1, 2, \ldots, M. \tag{11}$$

We first present the results for a network with two users and then extend to general networks.

A. Network With Two Users

Assume a network with only two users sharing the bandwidth $[0,\ B]$. Define

$$\bar{D}_1(\alpha) = \left\{ f \in [0, B] : \frac{c_2(f)}{c_1(f)} = \frac{\log_2(1 + \beta\rho_2(f))}{\log_2(1 + \beta\rho_1(f))} \leq \alpha \right\} \tag{12}$$

and

$$D_1(\alpha) = \left\{ f \in [0, B] : \frac{c_2(f)}{c_1(f)} = \frac{\log_2(1+\beta\rho_2(f))}{\log_2(1+\beta\rho_1(f))} < \alpha \right\}. \tag{13}$$

Similarly, we can define $\bar{D}_2(\alpha)$ and $D_2(\alpha)$ as the regions where $c_2(f)/c_1(f) >= \alpha$ and $c_2(f)/c_1(f) > \alpha$, respectively. It can be easily seen that

$$\bar{D}_2(\alpha) \cup D_1(\alpha) = \bar{D}_1(\alpha) \cup D_2(\alpha) = [0, B]$$

and

$$\bar{D}_2(\alpha) \cap D_1(\alpha) = \bar{D}_1(\alpha) \cap D_2(\alpha) = \varnothing.$$

The following theorem is proved in Appendix A and it determines the optimal subcarrier assignment for the cross-layer optimization.

Theorem 1: For a network with two users, if the subcarrier assignment $\{D_1^*, D_2^*\}$ is optimal, then D_1^* and D_2^* satisfy

$$D_1(\alpha^*) \subseteq D_1^* \subseteq \bar{D}_1(\alpha^*), \quad D_2^* = [0, B] - D_1^*$$

$$\alpha^* = \frac{U_1'(r_1^*)}{U_2'(r_2^*)}$$

and

$$r_i^* = \int_{D_i^*} c_i(f)\, df = \int_{D_i^*} \log_2(1+\beta\rho_i(f))\, df \quad \text{for } i = 1, 2$$

where

$$U_i'(r) = \frac{dU_i(r)}{dr}.$$

Fig. 2 demonstrates the difference between the utility-based optimization and the traditional throughput-based optimization. For the traditional optimization, $U_i(x) = x$; therefore, the threshold, $\alpha^* = U_1'(r_1)/U_2'(r_2)$ is always 1. Consequently, a subcarrier or frequency is allocated to the user with the larger channel gain, as in Fig. 2(b). To balance the efficiency and fairness, an increasing utility curve with a decreasing slope is usually used. In this case, the threshold α^* depends on how much resource each user has already occupied. Since the channel corresponding to user 2 is not as good as that of user 1 in Fig. 2(a), user 2 gets more frequency resource in the utility-based optimization than in the throughput-based optimization, as in Fig. 2(c).

It should be noted that the optimal subcarrier assignment is not unique as we can see in a network with flat fading channels. However, α^*, r_1^*, and r_2^* are unique.

B. Network With Multiple Users

The results for a two-user network can be extended to the general case of more than two users, which is summarized in the following theorem.

Theorem 2: For a network with M users, if the subcarrier assignment D_i^*'s for $i = 1, 2, \ldots, M$ maximizes the average utility, then for any $f \in D_i^*$, we have

$$U_j'(r_j^*)c_j(f) \le U_i'(r_i^*)c_i(f), \text{ for any } j \ne i \tag{14}$$

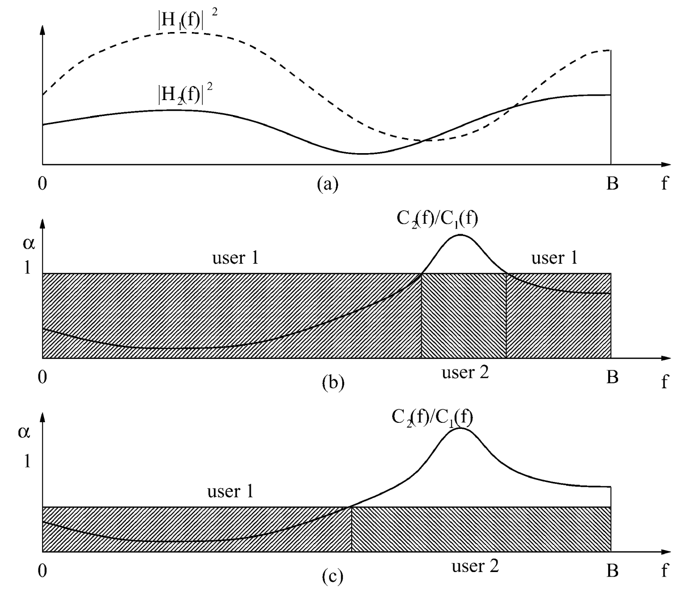

Fig. 2. Optimal subcarrier assignment for a two-user network. (a) Frequency responses for two users. (b) Subcarrier assignment resulting from throughput-based optimization. (c) Subcarrier assignment resulting from utility-based optimization.

and

$$r_i^* = \int_{D_i^*} c_i(f)df.$$

The proof of this theorem is very similar to that of Theorem 1 and is omitted here.

IV. APA

In the previous section, we discussed using DSA to maximize the network performance, where the power allocation is assumed to be fixed. In this section, we first investigate APA with fixed subcarrier assignment and then study joint DSA and APA. Since the achievable throughput is a function of the power allocation, it becomes

$$c_i(f) = \log_2(1+\beta p(f)\rho_i(f)).$$

A. APA With Fixed Subcarrier Assignment

When a subcarrier assignment is fixed, the APA optimization can be formulated as: given a fixed subcarrier assignment, D_i's for $i = 1, 2, \ldots, M$, assign the power density, $p(f)$, to maximize

$$\frac{1}{M}\sum_{i=1}^{M} U_i(r_i) = \frac{1}{M}\sum_{i=1}^{M} U_i\left(\int_{D_i} \log_2[1+\beta p(f)\rho_i(f)]\, df\right) \tag{15}$$

subject to

$$\frac{1}{B}\int_0^B p(f)\, df \le 1$$

and

$$p(f) \ge 0. \tag{16}$$

To achieve its optimality, a utility-based multilevel water-filling is needed, which is stated in the following theorem.

Theorem 3: For a given fixed subcarrier assignment, D_i's for all i, the optimal power allocation $p^*(f)$ satisfies

$$p^*(f) = \left[\frac{U_i'(r_i^*)}{\lambda} - \frac{1}{\beta\rho_i(f)}\right]^+ \quad \lambda > 0, \quad f \in D_i \tag{17}$$

where λ is a constant for the normalization of the optimal power density

$$[x]^+ = \begin{cases} x & x \geq 0 \\ 0 & x < 0 \end{cases}$$

and λ as well as the r_i^*'s satisfy

$$\frac{1}{B}\int_0^B p^*(f)df = 1, \qquad \text{and } r_i^* = \int_{D_i} \log_2[1 + \beta p^*(f)\rho_i(f)]df$$

where the r_i^*'s and $p^*(f)$ are the optimal values of the rates and the power density, respectively.

It should be indicated that Theorem 3 only gives a necessary condition for the globally optimal power allocation. The proof of this theorem is similar to the water-filling theorem [29], which is summarized in Appendix B.

Similar to the classical water-filling [29], the optimal power allocation cannot be directly calculated from (17), and iterative algorithms are needed to obtain the optimal one satisfying the power constraint.

There are two major differences between the classical water-filling and the one in Theorem 3. First, the water-level for each user is proportional to its current marginal utility value $U_i'(r_i)$. In other words, the power allocation is also related to the utility functions. Since the data rates of users are unlikely equal, it is from (17) that the water levels, $U_i'(r_i^*)/\lambda$'s, are different for different users. Second, the power constraint is the total transmission power rather than the power of an individual user. As shown in Fig. 3, the utility-based multilevel water-filling (17) can be regarded as an extension of the fixed-priority multilevel water-filling in [30].

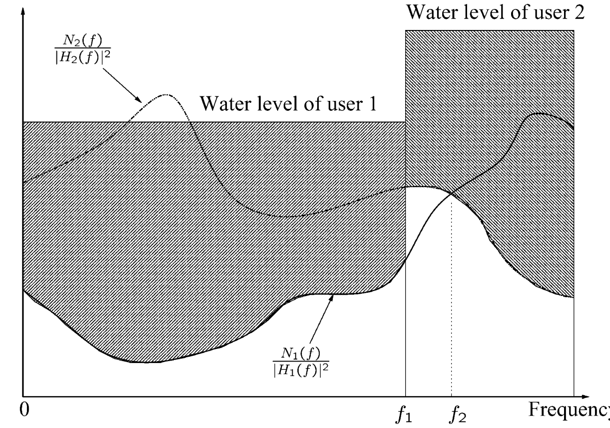

Fig. 3. Multilevel water-filling for adaptive power allocation in a two-user network.

B. Joint DSA and APA

The DSA and APA can be used simultaneously for the cross-layer optimization. The joint DSA and APA optimization can be formulated as follows: adjust the D_i's and $p(f)$ to maximize

$$\frac{1}{M}\sum_{i=1}^{M} U_i(r_i) = \frac{1}{M}\sum_{i=1}^{M} U_i\left(\int_{D_i} \log_2[1 + \beta p(f)\rho_i(f)]\, df\right) \tag{18}$$

subject to

$$\bigcup_{i=1}^{M} D_i \subseteq [0, B], \tag{19}$$

$$D_i \bigcap D_j = \varnothing, \quad i \neq j \text{ and } i, j = 1, 2, \ldots, M, \tag{20}$$

and

$$\frac{1}{B}\int_0^B p(f)df \leq 1$$

and

$$p(f) \geq 0. \tag{21}$$

Obviously, there are two necessary conditions for the global optimum for the joint DSA and APA.

1) Fixing the optimal subcarrier assignment, any change of the power allocation does not increase the total utility.
2) Fixing the optimal power allocation, any change of the subcarrier assignment does not increase the total utility.

Therefore, an optimal frequency assignment D_i^*'s for all i and power allocation $p^*(f)$ must satisfy the conditions in both Theorems 3 and 4. Consequently, we have the following theorem.

Theorem 4: Let the D_i^*'s for $i = 1, 2, \ldots, M$ and $p^*(f)$ be the optimal subcarrier assignment and power allocation, respectively. Then, they satisfy the conditions in (22), located at the bottom of the page, where the r_i^*'s and λ are constrained by

$$\frac{1}{B}\int_0^B p^*(f)df = 1, \qquad \text{and } r_i^* = \int_{D_i^*} \log_2(1 + \beta p^*(f)\rho_i(f))df.$$

$$\begin{cases} U_j'\left(r_j^*\right)\log_2(1 + \beta p^*(f)\rho_j(f)) \leq U_i'(r_i^*)\log_2(1 + \beta p^*(f)\rho_i(f)), & f \in D_i^*, \\ p^*(f) = \left[\frac{U_i'(r_i^*)}{\lambda} - \frac{1}{\beta\rho_i(f)}\right]^+, & \lambda > 0\ f \in D_i^*, \end{cases} \tag{22}$$

When the utility function is just the throughput $U_i(r_i) = r_i$, the optimal subcarrier assignment is independent of the optimal power allocation. In this case, the optimal subcarrier assignment and power allocation have the following closed forms:

$$\begin{cases} D_i^* = \{f \in [0:B] : \rho_i(f) = \max_m \rho_m(f)\} \\ p^*(f) = \left[\frac{1}{\lambda} - \frac{1}{\beta \max_m \rho_m(f)}\right]^+ \\ \frac{1}{B}\int_0^B p^*(f)df = 1 \end{cases}$$

which is identical to the result in [31]. It illustrates that frequency division multiple access (FDMA)-type systems can achieve the Shannon capacity when they are optimized for the sum of throughputs.

V. Properties of Cross-Layer Optimization

In this section, we will prove the convexity of the achievable data rate region and show that, if the utility function is concave, then a local maximum is also a global maximum. Therefore, the necessary conditions in Theorems 2–4 are also sufficient ones.

A. Convexity of Instantaneous Data Rate Region

A data rate vector $\mathbf{r}$ is defined as

$$\mathbf{r} = (r_1, r_2, \ldots, r_M)^T \in \mathbb{R}_+^M$$

where M is the number of users. The instantaneous data rate region $\mathcal{C}_\pi$ is a set that consists of the total achievable data rate vectors under the constraint of a resource allocation policy π, such as DSA, APA, or joint DSA and APA. The instantaneous data rate region is obviously determined by the channel conditions and the resource allocation constraints. It is intuitive that more adaptive resource allocation techniques will result in a larger feasible region.

The objective function is

$$U(\mathbf{r}) = \frac{1}{M}\sum_{i=1}^{M} U_i(r_i).$$

Thus, the optimization problem can be regarded as

$$\max_{\mathbf{r}\in\mathcal{C}_\pi} U(\mathbf{r}).$$

Therefore, if $\mathcal{C}_\pi$ is convex, the optimization problem will become tractable. The convexity of the instantaneous data rate region with frequency assignment and power allocation can be described by the following theorem, which is proved in Appendix C.

Theorem 5: For an OFDM-based network with infinitesimal subcarrier space and with DSA, APA, or joint DSA and APA, the achievable data rate region is convex.

With this theorem, we can obtain the following property of the cross-layer optimization.

Lemma 1: Let the *boundary of the instantaneous data rate region* be a subset of the data rate region with the following property: no component of any data rate vector can be increased while the other data rate components remain fixed. The data rate with respect to the maximum of the average utility must be on the boundary of the data rate region if each utility function is strictly increasing.

Proof: Suppose that the maximum can be achieved by a data rate vector $\mathbf{r}$, which is not on the boundary of the data rate region. There must exist a vector $\mathbf{r}^*$ such that $\mathbf{r} \leqslant \mathbf{r}^*$[1] with $r_i < r_i^*$ for some i, then $U(\mathbf{r}) < U(\mathbf{r}^*)$. The contradiction shows Lemma 1. ∎

Lemma 1 implies that using a strictly increasing utility function intends to assign all resources including all power and bandwidth to users.

B. Global Optimum

For general differentiable utility functions, the conditions (14), (17), and (22) are sufficient and necessary for *locally* optimal solutions of respective optimization problems; hence, they are only necessary for the *global* optimality. With concave utility functions, however, the global optimality of the cross-layer optimization can be described by the following theorem.

Theorem 6: If all $U_i(r_i)$'s are concave functions, then a local maximum of $U(\mathbf{r})$ is also a global maximum, and conditions (14), (17), and (22) are not only necessary but also sufficient, respectively.

Proof: The proof simply uses the following two consequences in convex analysis [32].

1) If all $U_i(r_i)$ are concave functions, then the objective function $U(\mathbf{r}) = (1/M)\sum_{i=1}^{M} U_i(r_i)$ is also a concave function.
2) If $\mathcal{C}_\pi \in \mathbb{R}^n$ is a convex set and $U : \mathcal{C}_\pi \mapsto \mathbb{R}$ is a concave function, then a local maximum of U is also a global maximum.

∎

The sufficiency of conditions (14), (17), and (22) for a global optimum is indispensable for algorithm design. If, in addition, the $U_i(r_i)$'s are all strictly concave, there is a unique global maximum solution to the optimization problems. Note that the unique global maximum implies that there is only one optimal data rate vector. However, there may be different frequency and power allocation schemes corresponding to the optimal data rate vector as we can see from a network with flat fading channels for all users.

The relation between the feasible data rate region and concave utility functions is shown in Fig. 4. Heuristically, Lemma 1 shows that the rate vector corresponding to the maximum is located on the boundary of the achievable rate region. Therefore, the optimal rate vector should be a point of tangency between the region boundary and an average utility contour.

VI. Efficiency and Fairness

Both efficiency and fairness issues are very important for resource allocation in wireless networks. An allocation scheme is said to be *efficient* if there is no other scheme that would simultaneously benefit someone and harm nobody in terms of their utilities. Therefore, the utility-based optimization is obviously efficient. Note that it differs from the spectral efficiency that is

[1] $\mathbf{r} \leqslant \mathbf{r}^*$ means $r_i \leq r_i^*$, for all i.

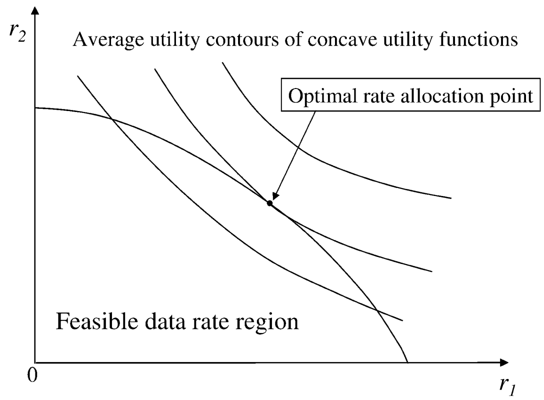

Fig. 4. Feasible data rate region and optimal rate allocation.

measured in terms of the total throughput over the bandwidth. Clearly, the maximum spectral efficiency is achieved by using a utility function $U_i(r_i) = r_i$ for all i.

With the channel knowledge for each user at the base station, the DSA scheme tends to assign subcarriers to users with a better SNR on the corresponding subcarriers, thereby having high spectral efficiency. It is obvious from (14) that the utility-based DSA penalizes the users with poor channel conditions.

When $U_i(r_i) = r_i$, $U_i'(r_i) = 1$. In this case, each subcarrier is assigned to the user with the best channel conditions among all users; therefore, the system can obtain the largest multiuser diversity with respect to spectral efficiency. Although the multiuser diversity is similar to the traditional selection diversity, its diversity gain results from the number of users, rather than from the number of antennas. It can be proved in Appendix D that, in the case of Rayleigh fading, the diversity gain grows approximately as $\ln(M)$, where M is the number of users in the system.

Fairness requires a fair share of bandwidth among competing users and protection from aggressive connections. Two representative types of fairness are proportional fairness [2] and max–min fairness [33]. Proportional fairness provides each connection a priority inversely proportional to its data rate. A vector of rates $\mathbf{r} \in \mathcal{C}$ is said to be proportionally fair if for any other feasible rate vector $\mathbf{r}' \in \mathcal{C}$, the aggregate of proportional changes is zero or negative

$$\sum_{i=1}^{M} \frac{r_i' - r_i}{r_i} \leq 0. \tag{23}$$

For a concave utility function $U(\mathbf{r})$ and a convex set $\mathcal{C}_\pi$, $\mathbf{r}$ is optimal if and only if

$$\nabla U(\mathbf{r})^T(\mathbf{r}' - \mathbf{r}) \leq 0 \text{ for all } \mathbf{r}' \in \mathcal{C}_\pi \tag{24}$$

where $\nabla U(\mathbf{r}) = [U_1'(r_1), U_2'(r_2), \ldots, U_M'(r_M)]^T$. When the logarithmic utility function $U(r) = \ln(r)$ is used, (24) is identical to (23). Therefore, the logarithmic utility function is associated with the proportional fairness for the utility-based optimization.

A data rate vector $\mathbf{r}$ is max–min fair if for each $m \in 1, 2, \ldots, M$, r_m cannot be increased without decreasing r_i form some i for which $r_i < r_m$. Obviously, max–min fairness has a strict fairness criterion since lower rates can get an absolute priority.

Consider a family of utility functions that is expressed as

$$U(r) = -\frac{r^{-\alpha}}{\alpha}, \quad \alpha > 0. \tag{25}$$

Obviously, the parameter α determines the degree of fairness. As α increases, the fairness of the corresponding utility function becomes stricter and stricter. When $\alpha \to \infty$, it turns out to be the max–min fairness.

It can be also seen from (14) that increasing utility functions encourage the users having good channel conditions, and decreasing marginal utility functions assign a high priority to the users with a low data rate. Therefore, utility-based resource allocation can guarantee both efficiency and fairness.

VII. Numerical Results

In this section, we present numerical results to illustrate the performance of the cross-layer optimization for OFDM wireless networks. To obtain the numerical results, the channel is assumed to be characterized by bad urban (BU) delay profile [34] and to suffer from shadowing with 8.0 dB standard deviation. The acceptable BER is chosen to be 10^{-6} for rate adaptation.

It is impossible to simulate the case of infinite subcarriers. For the BU delay profile, its corresponding coherent bandwidth is about 80 kHz; thus, the bandwidth of each subcarrier should be less than 80 kHz [34]. When the subcarrier bandwidth is small enough to let each subchannel experience flat fading, the optimization performance approaches closely to that with infinite subcarriers.

In simulation, we assume that the bandwidth of each subcarrier is 10 kHz, and the utility function in (3) is used. To compare various results, we fix the average bandwidth per user to 80 kHz, which implies that the overall bandwidth increases linearly with the number of users.

Fig. 5 shows the results for different resource allocation schemes. From the figure, we can see that the DSA, APA, and joint DSA and APA effectively improve the network performance compared to the fixed subcarrier assessment (FSA), and the performance gain increases with the number of users. The gain of joint DSA and APA is about 2.5 dB for a two-user network and as large as 5 dB for a 16-user network. We have proved, in Appendix D, that the diversity gain increases logarithmically with the number of users when the utility function is just the throughput $(U_i(r_i) = r_i)$. Unlike the asymptotic analysis of multiuser diversity gain in [15], our analysis is valid for an arbitrary user number.

It should be noted from the figure that the performance improvement of APA is limited for continuous rate adaptation. However, it can be shown in [35] that there is a huge performance gain in the realistic situation where discrete order of modulation is used.

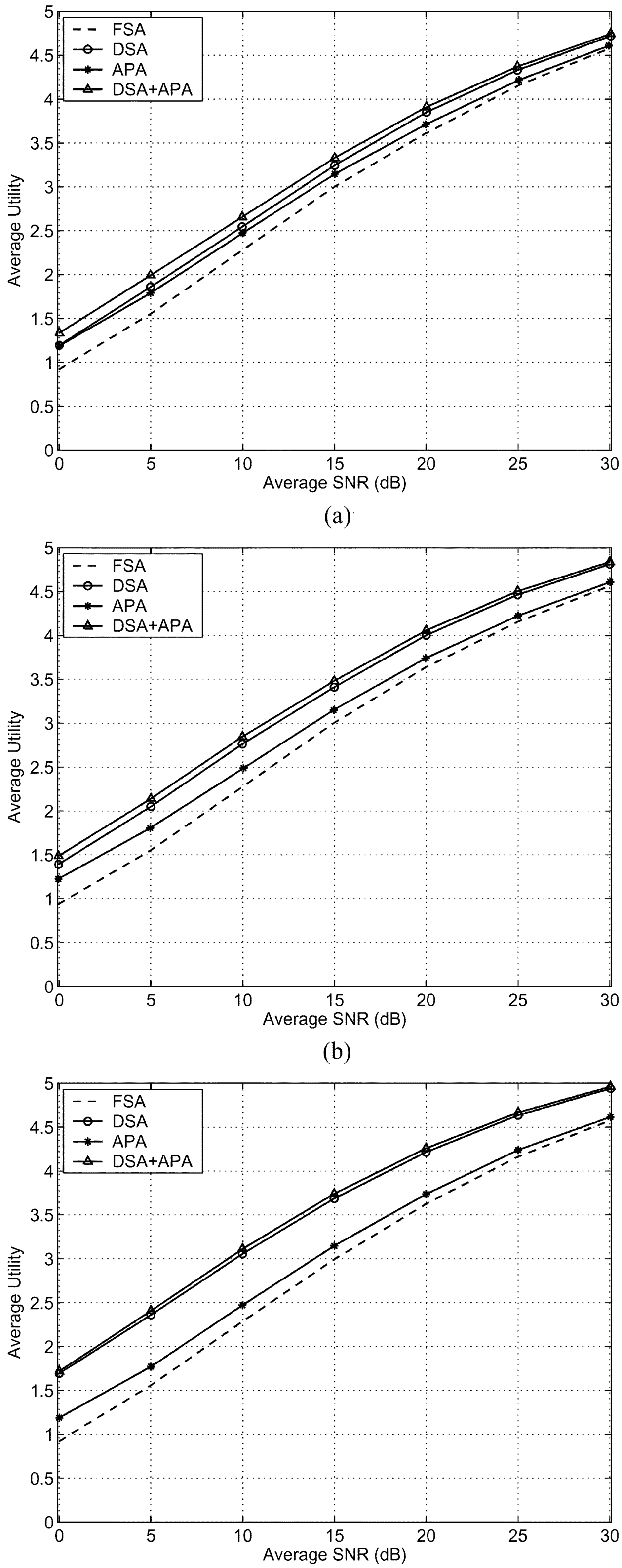

Fig. 5. Average user utility versus SNR for the OFDM wireless network with FSA, DSA, APA, and joint DSA and APA. (a) Two users. (b) Four users. (c) Sixteen users.

VIII. Conclusion and Current Research

In this paper, we have presented utility-based cross-layer optimization for OFDM-based wireless networks. The utility is used here to build a bridge between the physical and MAC layers and to balance the efficiency and fairness of resource allocation. In particular, we have investigated the necessary and sufficient conditions for finding an optimum for the DSA, APA, and joint DSA and APA schemes. Through numerical results, we have also demonstrated the significant performance gain of cross-layer optimization. Our research here provides a theoretical background for developing practical algorithms for the cross-layer optimization in future wireless networks.

In our research, we have assumed that the OFDM signal is composed of an infinite number of subcarriers, and that the modulation order can be continuously changed based on the channel conditions. In [35], we will investigate the cross-layer optimization under more realistic conditions and develop low-complexity approaches for the OFDM wireless network to exploit multiuser diversity.

Appendix A
Proof of Theorem 1

Proof: If the $\bar{D}_1^*$'s are optimal, then any change of allocation will not increase the average utility. Let $(f-(1/2)\triangle f,\ f+(1/2)\triangle f) \in D_1^*$. If $(f-(1/2)\triangle f,\ f+(1/2)\triangle f)$ is assigned to the other user, then the data rate of user 1 will be decreased by $\triangle r_1 = c_1(f)\triangle f$, while the data rate of user 2 will be increased by $\triangle r_2 = c_2(f)\triangle f$. However, the new average utility will be equal to or less than the optimal one, that is

$$U_1(r_1^* - \triangle r_1) + U_2(r_2^* + \triangle r_2) \leq U_1(r_1^*) + U_2(r_2^*)$$

which is equivalent to

$$U_2(r_2^* + \triangle r_2) - U_2(r_2^*) \leq U_1(r_1^*) - U_1(r_1^* - \triangle r_1).$$

Dividing both sides by $\triangle f$, we have

$$\frac{U_2(r_2^* + \triangle r_2) - U_2(r_2^*)}{\triangle f} \leq \frac{U_1(r_1^*) - U_1(r_1^* - \triangle r_1)}{\triangle f}.$$

Since $\triangle r_1 = c_1(f)\triangle f$ and $\triangle r_2 = c_2(f)\triangle f$, we have

$$\frac{U_2(r_2^* + \triangle r_2) - U_j(r_2^*)}{\triangle r_2} c_2(f) \leq \frac{U_1(r_1^*) - U_1(r_1^* - \triangle r_1)}{\triangle r_1} c_1(f).$$

When $\triangle f \to 0$, $\triangle r_1 \to 0$ and $\triangle r_2 \to 0$. Consequently

$$\lim_{\triangle r_2 \to 0} \frac{U_2(r_2^* + \triangle r_2) - U_j(r_2^*)}{\triangle r_2} c_2(f) \leq \lim_{\triangle r_1 \to 0} \frac{U_1(r_1^*) - U_1(r_1^* - \triangle r_1)}{\triangle r_1} c_1(f)$$

or

$$U_2'(r_2^*) c_2(f) \leq U_1'(r_1^*) c_1(f)\ f \in \bar{D}_1^* \tag{A.1}$$

which implies, for any $f \in \bar{D}_1^*$

$$\frac{c_2(f)}{c_1(f)} \leq \frac{U_1'(r_1^*)}{U_2'(r_2^*)} (= \alpha^*)$$

that is, $f \in \bar{D}_1(\alpha^*)$ and $D_1^* \subseteq \bar{D}_1(\alpha^*)$.

Similarly, we can prove that

$$D_2^* \subseteq \bar{D}_2(\alpha^*).$$

Therefore,

$$\begin{aligned} D_1(\alpha^*) &= [0,\ B] - \bar{D}_2(\alpha^*) \\ &\subseteq [0,\ B] - D_2^* \\ &= D_1^*. \end{aligned}$$

∎

Appendix B
Proof of Theorem 3

Proof: For a fixed subcarrier assignment D_i for all i, we define $p_i(f)$ for $i = 1,\ 2, \ldots,\ M$ as

$$p_i(f) = \begin{cases} p(f), & f \in D_i \\ 0, & \text{otherwise.} \end{cases}$$

Using the Lagrangian method, the above optimization problem with the power constraint becomes to maximize

$$\frac{1}{M}\sum_{i=1}^{M} U_i\left(\int_{D_i} \log_2\left[1+\beta p(f)\rho_i(f)\, df\right]\right) - \lambda'\left[\frac{1}{B}\int_0^B p(f)\, df - 1\right]$$

or

$$\frac{1}{M}\sum_{i=1}^{M} \left\{U_i\left(\int_{D_i} \log_2\left[1+\beta p_i(f)\rho_i(f)\, df\right]\right) - \lambda'\left[\frac{1}{B}\int_{D_i} p_i(f)\, df - 1\right]\right\}$$

where $\lambda' \geq 0$.

With the Karush-Kuhn-Tucker (KKT) conditions [32], we have

$$\frac{1}{M}U_i'(r_i^*)\frac{\partial}{\partial p_i(f)}\log_2\left\{1+\beta p_i(f)\rho_i(f)\right\} - \frac{\lambda'}{B}\frac{\partial}{\partial p_i(f)}p_i(f)\bigg|_{p_i(f)=p_i^*(f)} = 0, \quad \text{for all } i, \tag{B.1}$$

$$\lambda' \geq 0 \tag{B.2}$$

$$\lambda'\left[\sum_{i=1}^{M}\frac{1}{B}\int_{D_i} p_i(f)\, df - 1\right] = 0. \tag{B.3}$$

Equation (B.1) is equivalent to

$$U_i'(r_i^*)\frac{\beta\rho_i(f)}{1+\beta\rho_i(f)p_i^*(f)} - \lambda'\frac{M}{\log_2(e)B} = 0, \quad \text{for all } i.$$

Let $\lambda = \lambda'(M/\log_2(e)B)$. Then, the optimal power allocation for a fixed subcarrier assignment satisfies:

$$\begin{cases} p_i^*(f) = \left[\frac{U_i'(r_i^*)}{\lambda} - \frac{1}{\beta\rho_i(f)}\right]^+ & f \in D_i \\ \sum_{i=1}^{M}\frac{1}{B}\int_{D_i} p_i^*(f)\, df = 1 \end{cases}$$

or

$$\begin{cases} p^*(f) = \left[\frac{U_i'(r_i^*)}{\lambda} - \frac{1}{\beta\rho_i(f)}\right]^+ & f \in D_i \\ \frac{1}{B}\int_0^B p^*(f)\, df = 1. \end{cases}$$

Appendix C
Proof of Theorem 5

Proof: Assume that the system has joint DSA and APA. Then, $\forall \mathbf{r}^{(1)}, \mathbf{r}^{(2)} \in \mathcal{C}_{DSA+APA}, \alpha \in [0,1]$, we need to show that $\alpha\mathbf{r}^{(1)}+(1-\alpha)\mathbf{r}^{(2)} \in \mathcal{C}_{DSA+APA}$. $\mathbf{r}^{(1)} = [r_1^{(1)}, r_2^{(1)}, \ldots, r_M^{(1)}]^T$ is achieved with $D_m^{(1)}$ and $p^{(1)}(f)$, $\mathbf{r}^{(2)} = [r_1^{(2)}, r_2^{(2)}, \ldots, r_M^{(2)}]^T$ is achieved with $D_m^{(2)}$ and $p^{(2)}(f)$, where for $m \in \{1, 2, \ldots, M\}$. Of course, $D_m^{(1)}$ and $D_m^{(2)}$ satisfy (4) and (5); $p^{(1)}(f)$ and $p^{(2)}(f)$ yield (7). We represent those two power allocations as $\mathbf{p}^{(1)}$ and $\mathbf{p}^{(2)}$, respectively.

We define the measure of a frequency set as follows. When the frequency set $D = \bigcup_i [a_i, b_i]$, $b_i \leq a_{i+1}$, the measure μ is given by $\mu(D) = \sum_i (b_i - a_i)$. For user m, we have

$$\begin{aligned} r_m^{(1)} &= \int_{D_m^{(1)}} c_m^{\mathbf{p}^{(1)}}(f)\, d\mu \\ r_m^{(2)} &= \int_{D_m^{(2)}} c_m^{\mathbf{p}^{(2)}}(f)\, d\mu \end{aligned}$$

where $c_m^{\mathbf{p}}(f)$ denotes the achievable throughput of user m at frequency f with power allocation $\mathbf{p}$.

We divide $[0, B]$ into a family of sets F_n's so that

$$\bigcup_n F_n = [0, B], \quad F_i \bigcap F_j = \varnothing \quad i \neq j \tag{C.1}$$

$$\max_{f \in F_n}\left\{c_m^{\mathbf{p}^{(1)}}(f)\right\} - \min_{f \in F_n}\left\{c_m^{\mathbf{p}^{(1)}}(f)\right\} \to 0 \quad \text{for all } m, n \tag{C.2}$$

$$\max_{f \in F_n}\left\{c_m^{\mathbf{p}^{(2)}}(f)\right\} - \min_{f \in F_n}\left\{c_m^{\mathbf{p}^{(2)}}(f)\right\} \to 0 \quad \text{for all } m, n. \tag{C.3}$$

Equations (C.2) and (C.3) imply

$$\begin{aligned} &\max_{f \in F_n} p^{(1)}(f) - \min_{f \in F_n} p^{(1)}(f) \to 0 \quad \text{for all } n \\ &\max_{f \in F_n} p^{(2)}(f) - \min_{f \in F_n} p^{(2)}(f) \to 0 \quad \text{for all } n. \end{aligned}$$

Each F_n is divided into two subsets F_n^{α} and $F_n^{(1-\alpha)}$ that satisfy

$$F_n^{\alpha} \bigcup F_n^{(1-\alpha)} = F_n, \quad F_n^{\alpha} \bigcap F_n^{(1-\alpha)} = \varnothing \tag{C.4}$$

and $\mu(F_n^{\alpha}) = \alpha\mu(F_n)$.

If $F_n \in D_m$, we use $D_{m,n}$ to denote F_n. Thus,

$$r_m^{(1)} = \sum_n c_m^{\mathbf{P}^{(1)}}(n)\mu(D_{m,n}^{(1)})$$
$$r_m^{(2)} = \sum_n c_m^{\mathbf{P}^{(2)}}(n)\mu(D_{m,n}^{(2)}).$$

In the same way, using $D_{m,n}^{\alpha}$ to denote $F_n^{\alpha} \in D_m$, we have

$$\int_{D_m^{(1),\alpha}} c_m^{\mathbf{P}^{(1)}}(f)\, d\mu = \sum_n c_m^{\mathbf{P}^{(1)}}(n)\mu(D_{m,n}^{(1),\alpha}) = \alpha r_m^{(1)}$$
$$\int_{D_m^{(2),(1-\alpha)}} c_m^{\mathbf{P}^{(2)}}(f) d\mu = \sum_n c_m^{\mathbf{P}^{(2)}}(n)\mu(D_{m,n}^{(2),(1-\alpha)}) = (1-\alpha) r_m^{(2)}$$
$$\text{where} \quad D_m^{(1),\alpha} = \bigcup_n D_{m,n}^{(1),\alpha}$$
$$D_m^{(2),(1-\alpha)} = \bigcup_n D_{m,n}^{(2),(1-\alpha)}.$$

Therefore, with the new frequency assignment $D_m = D_m^{(1),\alpha} \cup D_m^{(2),(1-\alpha)}$ and the new power allocation

$$p(f) = \begin{cases} p^{(1)}(f), & f \in D_m^{(1),\alpha} \\ p^{(2)}(f), & f \in D_m^{(2),(1-\alpha)} \end{cases}$$

the new data rate for user m is

$$r_m = \int_{D_m^{(1),\alpha}} c_m^{\mathbf{P}^{(1)}}(f)\, d\mu + \int_{D_m^{(2),(1-\alpha)}} c_m^{\mathbf{P}^{(2)}}(f)\, d\mu$$
$$= \alpha r_m^{(1)} + (1-\alpha) r_m^{(2)}.$$

Furthermore, due to (C.1) and (C.4), the D_m's satisfy (4) and (5). In addition

$$\begin{aligned}\frac{1}{B}\int_0^B p(f)\, d\mu &= \frac{1}{B}\sum_m \int_{D_m} p(f)\, d\mu \\ &= \frac{1}{B}\sum_m \int_{D_m^{(1),\alpha}} p^{(1)}(f)\, d\mu \\ &\quad + \frac{1}{B}\sum_m \int_{D_m^{(2),(1-\alpha)}} p^{(2)}(f)\, d\mu \\ &= \frac{\alpha}{B}\sum_m \int_{D_m^{(1)}} p^{(1)}(f) d\mu \\ &\quad + \frac{1-\alpha}{B}\sum_m \int_{D_m^{(2)}} p^{(2)}(f) d\mu \\ &= \alpha \frac{1}{B}\int_0^B p^{(1)}(f) d\mu \\ &\quad + (1-\alpha)\frac{1}{B}\int_0^B p^{(2)}(f)\, d\mu \\ &\le 1.\end{aligned}$$

Therefore, there are feasible frequency assignment and power allocation schemes such that $\alpha \mathbf{r}^{(1)} + (1-\alpha)\mathbf{r}^{(2)} \in \mathcal{C}$.

Let $p^{(1)}(f) = p^{(2)}(f)$ in the previous proof. Then, we have that the achievable data rate region is convex when only DSA is used. Let $D_m^{(1)} = D_m^{(2)}$ for all m in the previous proof. Similarly, we have that the achievable data rate region is also convex when only APA is used. ■

Appendix D
Maximum Order of Multiuser Diversity

We consider the case where $U_i(r_i) = r_i$, for all i, and where there is no APA The optimal frequency assignment should be

$$D_i^* = \left\{ f \in [0:B] : \rho_i(f) = \max_{m \in \{1,2,\ldots,M\}} \rho_m(f) \right\}. \quad \text{(D.1)}$$

After assigning the frequency, the channel SNR at frequency f, $\mathrm{SNR}(f)$, is expressed as

$$\mathrm{SNR}(f) = \max_{m \in \{1,2,\ldots,M\}} \rho_m(f). \quad \text{(D.2)}$$

It is assumed that fading at frequency f has a Rayleigh distribution, which is identical and independent of other users. Therefore, the PDF of $\rho_m(f)$ is

$$p_{\rho_m(f)}(\gamma) = \frac{e^{-\gamma/\Gamma}}{\Gamma}$$

where $\Gamma = \overline{\rho}_m(f)$ is the mean channel SNR at frequency f. Similar to selection diversity [36], the PDF of $\mathrm{SNR}(f)$ can be obtained as

$$p_{\mathrm{SNR}(f)}(\gamma) = \frac{M}{\Gamma}(1 - e^{-\gamma/\Gamma})^{M-1} e^{-\gamma/\Gamma}. \quad \text{(D.3)}$$

The average $\mathrm{SNR}(f)$, $\overline{\mathrm{SNR}}(f)$, can be calculated by

$$\overline{\mathrm{SNR}}(f) = \int_0^\infty \gamma \cdot p_{\mathrm{SNR}(f)}(\gamma)\, d\gamma = \Gamma \sum_{m=1}^{M} \frac{1}{m}.$$

The diversity gain, thus, is

$$\frac{\overline{\mathrm{SNR}}(f)}{\Gamma} = \sum_{m=1}^{M} \frac{1}{m}. \quad \text{(D.4)}$$

Since

$$\frac{1}{2(M+1)} < \sum_{m=1}^{M} \frac{1}{m} - (\ln M + E) < \frac{1}{2M}$$

where E is the Euler constant $(0.57721566\cdots)$, $\ln M + E$ is a tight lower bound on the diversity gain when M is large. Consequently, the multiuser diversity gain increases approximately with $\ln M$.

Acknowledgment

The authors would like to thank Prof. L. J. Cimini, Jr. for his insightful comments.

References

[1] S. Shenker, "Fundamental design issues for the future internet," *IEEE J. Sel. Areas Commun.*, vol. 13, no. 7, pp. 1176–1188, Sep. 1995.

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

[3] F. Kelly, A. Maulloo, and D. Tan, "Rate control in communication networks: Shadow prices, proportional fairness, and stability," *J. Oper. Res. Soc.*, vol. 49, pp. 237–252, 1998.

[4] J. K. MacKie-Mason and H. R. Varian, "Pricing congestible network resources," *IEEE J. Sel. Areas Commun.*, vol. 13, no. 7, pp. 1141–1149, Sept. 1995.

[5] E. Altman, T. Basar, T. Jimenez, and N. Shimkin, "Competitive routing in networks with polynomial cost," in *Proc. IEEE Conf. Computer Communications*, Mar. 2000, pp. 1586–1593.

[6] V. Shah, N. B. Mandayam, and D. J. Goodman, "Power control for wireless data based on utility and pricing," in *Proc. IEEE Personal, Indoor, Mobile Radio Communications Conf.*, 1998, pp. 1427–1432.

[7] D. J. Goodman and N. B. Mandayam, "Power control for wireless data," *IEEE Personal Commun.*, vol. 7, no. 2, pp. 48–54, Apr. 2000.

[8] C. U. Saraydar, N. B. Mandayam, and D. J. Goodman, "Pricing and power control in a multicell wireless data network," *IEEE J. Sel. Areas Commun.*, vol. 19, no. 10, pp. 1883–1892, Oct. 2001.

[9] P. Liu, M. Honig, and S. Jordan, "Forward link CDMA resource allocation based on pricing," in *Proc. IEEE Wireless Communications Networking Conf.*, vol. 2, Sep. 2000, pp. 619–623.

[10] C. Zhou, M. L. Honig, and S. Jordan, "Two-cell power allocation for wireless data based on pricing," in *Proc. 39th Annu. Allerton Conf.*, Monticello, IL, Oct. 2001, pp. 1088–1097.

[11] L. Song and N. B. Mandayam, "Hierarchical sir and rate control on the forward link for CDMA data users under dalay and error constraints," *IEEE J. Sel. Areas Commun.*, vol. 19, no. 10, pp. 1871–1882, Oct. 2001.

[12] J. Chuang and N. Sollenberger, "Beyond 3G: Wideband wireless data access based on OFDM and dynamic packet assignment," *IEEE Commun. Mag.*, vol. 38, no. 7, pp. 78–87, July 2000.

[13] S. Nanda, K. Balachandran, and S. Kumar, "Adaptation techniques in wireless packet data services," *IEEE Commun. Mag.*, vol. 38, no. 1, pp. 54–64, Jan. 2000.

[14] A. J. Goldsmith and S. G. Chua, "Variable-rate variable-power MQAM for fading channel," *IEEE Trans. Commun.*, vol. 45, no. 10, pp. 1218–1230, Oct. 1997.

[15] P. Viswanath, D. N. C. Tse, and R. L. Laroia, "Opportunistic beamforming using dumb antennas," *IEEE Trans. Inf. Theory*, vol. 48, no. 6, pp. 1277–1294, June 2002.

[16] D. Tse and S. Hanly, "Multi-access fading channels: Part I: Polymatroid structure, optimal resource allocation and throughput capacities," *IEEE Trans. Inf. Theory*, vol. 44, no. 7, pp. 2796–2815, Nov. 1998.

[17] L. Li and A. J. Goldsmith, "Optimal resource allocation for fading broadcast channels- Part I: Ergodic capacity," *IEEE Trans. Inf. Theory*, vol. 47, no. 3, pp. 1083–1102, Mar. 2001.

[18] A. J. Goldsmith and M. Effros, "The capacity region of broadcast channels with intersymbol interference and colored Gaussian noise," *IEEE Trans. Inf. Theory*, vol. 47, no. 1, pp. 219–240, Jan. 2001.

[19] C. Y. Wong, R. S. Cheng, K. B. Letaief, and R. D. Murch, "Multiuser OFDM with adaptive subcarrier, bit, and power allocation," *IEEE J. Sel. Areas Commun.*, vol. 17, no. 10, pp. 1747–1758, Oct. 1999.

[20] W. Rhee and J. M. Cioffi, "Increase in capacity of multiuser OFDM system using dynamic subcarrier allocation," in *Proc. IEEE Vehicular Technology Conf.*, 2000, pp. 1085–1089.

[21] W. Yu and J. M. Cioffi, "FDMA capacity of Gaussian multiple-access channels with ISI," *IEEE Trans. Commun.*, vol. 50, no. 1, pp. 102–111, Jan. 2002.

[22] L. M. Hoo, B. Halder, J. Tellado, and J. M. Cioffi, "Multiuser transmit optimization for multicarrier broadcast channels: Asymptotic FDMA capacity region and algorithms," *IEEE Trans. Commun.*, submitted for publication.

[23] G. Song and Y. (G.) Li, "Utility-based boint physical-MAC layer optimization in OFDM," in *Proc. IEEE Global Communications Conf.*, vol. 1, Nov. 2002, pp. 671–675.

[24] ——, "Adaptive subcarrier and power allocation in OFDM based on maximizing utility," in *Proc. IEEE Vehicular Technology Conf.*, vol. 2, Apr. 2003, pp. 905–909.

[25] ——, "Adaptive resource allocation based on utility optimization in OFDM networks," in *Proc. IEEE Global Communications Conf.*, Dec. 2003, pp. 586–590.

[26] Y. (G.) Li, "Pilot-symbol-aided channel estimation for OFDM in wireless systems," *IEEE Trans. Veh. Tech.*, vol. 49, no. 4, pp. 1207–1215, Jul. 2000.

[27] X. Qiu and K. Chawla, "On the performance of adaptive modulation in cellular systems," *IEEE Trans. Commun.*, vol. 47, no. 6, pp. 884–895, Jun. 1999.

[28] Z. Jiang, Y. Ge, and Y. (G.) Li, "Max-utility wireless resource management for best effort traffic," *IEEE Trans. Wireless Commun.*, vol. 4, no. 1, pp. 100–111, Jan. 2005.

[29] C. E. Shannon, "Communication in the presence of noise," *Proc. IRE*, vol. 37, pp. 10–21, Jan. 1949.

[30] L. Hoo, J. Tellado, and J. Cioffi, "FDMA-based multiuser transmit optimization for broadcast channels," in *Proc. IEEE Wireless Communications Networking Conf.*, vol. 2, Chicago, IL, Sep. 2000, pp. 597–602.

[31] D. N. Tse, "Optimal power allocation over parallel Gaussian broadcast channel," in *Proc. IEEE Int. Symp. Information Theory*, Ulm, Germany, June 1997, pp. 27–27.

[32] R. T. Rockafellar, *Convex Analysis*. Princeton, NJ: Princeton Univ. Press, 1970.

[33] D. Bertsekas and R. Gallager, *Data Networks*. Englewood Cliffs, NJ: Prentice-Hall, 1987.

[34] G. L. Stüber, *Principles of Mobile Communication*, 2nd ed. Norwell, MA: Kluwer, 2000.

[35] G. Song and Y. (G.) Li, "Cross-layer optimization for OFDM wireless network—Part II: Algorithm development," *IEEE Trans. Wireless Commun.*, vol. 4, no. 2, pp. 625–634, Mar. 2005.

[36] W. C. Jakes, Ed., *Microwave Mobile Communication*, 2nd ed. Piscataway, NJ: IEEE Press, 1994.

Guocong Song (S'02) received the B.S. and M.S. degrees from the Department of Electronic Engineering, Tsinghua University, Beijing, China, in 1997 and 2000, respectively. He is currently working toward the Ph.D. degree at the School of Electrical and Computer Engineering, Georgia Institute of Technology (Georgia Tech), Atlanta.

Since May 2001, he has been a Research Assistant with the School of Electrical and Computer Engineering, Georgia Tech. His research interests include wireless communications and networking, with a current focus on cross-layer design and optimization for wireless networks, centralized and distributed resource allocation, scheduling and MAC in multichannel networks, and QoS provisioning.

Ye (Geoffrey) Li (S'93–M'95–SM'97) received the B.S.E. and M.S.E. degrees in wireless engineering from Nanjing Institute of Technology, Nanjing, China, in 1983 and 1986, respectively, and the Ph.D. degree in electrical engineering from Auburn University, Auburn, AL, in 1994 .

After spending several years at AT&T Labs—Research, Redbank, NJ, he joined the Georgia Institute of Technology, Atlanta, as an Associate Professor in 2000. His general research interests include statistical signal processing and wireless communications. In these areas, he has written over 100 papers published in referred journals and presented in various international conferences. He also has over ten US patents granted or pending.

Dr. Li once served as a Guest Editor for two special issues on Signal Processing for Wireless Communications for the IEEE JOURNAL OF SELECTED AREAS IN COMMUNICATIONS and is currently an Editorial Board Member of *EURASIP Journal on Applied Signal Processing*. He is serving as an Editor for Wireless Communication Theory for the IEEE TRANSACTIONS ON COMMUNICATIONS. He organized and chaired many international conferences, including Vice-Chair of *IEEE 2003 International Conference on Communications*.