Distributed Power Allocation and Scheduling for Parallel Channel Wireless Networks

Xiangping Qin, Randall A. Berry

Abstract

In this paper we develop distributed approaches for power allocation and scheduling in wireless access networks. We consider a model where users communicate over a set of parallel multi-access fading channels, as in an orthogonal frequency division multiple access (OFDMA) system. At each time, each user must decide which channels to transmit on and how to allocate its power over these channels. We give distributed power allocation and scheduling policies, where each user's actions depend only on knowledge of their own channel gains. Assuming a collision model for each channel, we characterize an optimal policy which maximizes the system throughput and also give a simpler sub-optimal policy. Both policies are shown to have the optimal scaling behavior in several asymptotic regimes.

Index Terms

Multi-user diversity, OFDM, opportunistic scheduling, random access.

I. Introduction

It is well established that dynamically allocating transmission rate and power can improve the performance of wireless networks. In this paper, we consider these approaches for the uplink in a wireless network, modeled as a fading multiple access channel. For such channels, power allocation and scheduling have received much attention. For example, [1]–[5] consider these problems in the context of the information theoretic capacity region of a multi-access fading channel under various assumptions. In other work, such as [6], [7], adaptive bit and power

This work was supported in part by the Northwestern-Motorola Center for Communications and by NSF CAREER award CCR-0238382. The material in this paper was presented in part at the 3rd Intl. Symposium on Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks (WiOpt) Trentino, Italy, April 3-7 2005.

X. Qin is now with Samsung Information Systems America, 75 West Plumeria Rd., San Jose, CA 95134 USA, email: x.qin@samsung.com. This work was performed while she was with the Department of EECS, Northwestern University.

R. Berry is with the Dept. of EECS, Northwestern University, 2145 Sheridan Rd., Evanston, IL 60208 USA, e-mail: rberry@eecs.northwestern.edu.

allocation are studied in the context of an OFDMA system. In these cases, optimally allocating resources requires a centralized controller with knowledge of every user's channel state. Because of the required overhead and delays involved, it may not be feasible for a centralized controller to acquire this information in a fast-fading environment or a system with a large number of users. Here, we instead consider approaches where each transmitter allocates its transmission rate and power based only on knowledge of its own channel conditions. This can be obtained, for example, via a single pilot signal broadcast by the receiver in a time-division duplex system [8]. This requires much less overhead, but since each user has incomplete information, a distributed approach for power allocation and scheduling is required. In prior work [8]–[10], we have considered such approaches for the case where all users communicate over a single flat-fading channel. In particular, in [8], [9], an approach based on the Aloha protocol is given, where each user randomly transmits with a probability based on its own channel gain. It is shown that as the number of users increases, the throughput of such a system scales at the same rate as that obtained by an optimal centralized controller. In this paper, we extend this approach to the case where each user may transmit over multiple ”parallel” channels. For example, each channel may model a subcarrier or group of subcarriers in an OFDMA system, such as the IEEE 802.16 standard. A new consideration here is that each user must now decide how to allocate its transmission power across the available channels.

In the next section, we begin by describing our basic model for an uplink with k parallel channels. Using this model, we formulate a distributed power allocation and scheduling problem with a finite number of users and characterize the optimal solution to this problem. We next give a simplified allocation scheme and analyze the performance of both the optimal and simplified schemes in three asymptotic regimes: (*i*) the number of users increases with a fixed number of channels, (*ii*) the number of channels increase with a fixed number of users, and (*iii*) both the number of channels and the number of users increase with fixed ratio. In each case, we characterize the asymptotic growth rate. Both approaches are shown to achieve the same order of growth as an optimal centralized approach in each asymptotic regime. The asymptotic ratio of the throughput of the distributed approaches to that of the optimal centralized approach is given in each regime. We also compare the performance of several approaches that require a small amount of additional coordination among the users. Finally, some simulation results are given.

II. Model Description

We consider a model of n users communicating to a single receiver over k parallel channels. Each channel between each user and the receiver is modeled as a time-slotted, block-fading channel with frequency-flat fading and bandwidth W_c. At each time t, the received signal on the jth channel is given by

$$y_j(t) = \sum_{i=1}^{n} \sqrt{H_{ij}(t)} x_{ij}(t) + z_j(t), \tag{1}$$

where $x_{ij}(t)$ and $H_{ij}(t)$ are the transmitted signal and channel gain for the ith user on channel j, and $z_j(t)$ is additive white Gaussian noise with power spectral density $\frac{N_0}{2}$. To simplify notation we assume that $N_0 W_c = 1$. The channel gains are assumed to be fixed during each time-slot and to randomly vary between time-slots, *i.e.* $H_{ij}(t) = H_{ij}$ for all $t \in [mT, (m+1)T]$, where T is the length of a time-slot. Here, $\{H_{ij}\}_{i=1,..,n, j=1,..,k}$ are assumed to be independent and identically distributed (i.i.d.) across both the users and channels with a continuous probability density $f_H(h)$ on $[0, \infty)$.[1] We assume that $\mathbb{E}(H_{i,k}) < \infty$ and that $f_H(h) > 0$, for all $h > 0$ and is differentiable. It follows that the corresponding distribution function $F_H(h)$ is strictly increasing and twice differentiable. Let $\bar{F}_H(h) = 1 - F_H(h)$ denote the channel gain's complimentary distribution function. For example, if each channel experiences Rayleigh fading, then H will be exponentially distributed, and so $\bar{F}_H(h) = e^{-h/h_0}$, where $h_0 = \mathbb{E}(H_{i,k})$.

Figure 1 around here.

We focus on the case where at the start of each slot, each user i has perfect knowledge of $H_{i1}, ..., H_{ik}$, but no knowledge of the channel gains for any other users. For convenience, we drop the user subscript and let $\mathbf{H} = (H_1, ..., H_k)$ denote the vector of channel gains for an arbitrary user. Given the current realization of $\mathbf{H}$, each user must determine on which channels to transmit and how to allocate its power. This is represented by the user's power allocation $\mathbf{P}(\mathbf{h}) = (P_1(\mathbf{h}), P_2(\mathbf{h}), ..., P_k(\mathbf{h}))$, where $P_j(\mathbf{h})$ indicates the power allocated to channel j given that $\mathbf{H} = \mathbf{h}$.[2] This power allocation must satisfy a total power constraint of $\check{P}$ across all channels in each time-slot, *i.e.*, $\sum_j P_j(\mathbf{h}) \leq \check{P}$, for all $\mathbf{h}$. No cooperation exists among users. In particular,

[1] In an OFDM system different sub-carriers will typically experience correlated fading. However, if each channel is a large enough group of sub-carriers, then this independence assumption is reasonable.

[2] If a user does not transmit on channel j, then $P_j(\mathbf{h}) = 0$.

all users are required to employ the same power allocation and transmission scheme; *i.e.*, they can not cooperate in selecting these allocations.

During each time-slot, we assume that at most one user can successfully transmit on each channel. If more than one user transmits on a given channel, a collision occurs and no packets are received. However, a packet sent over another channel without a collision will still be received. In other words, if a user simultaneously transmits on multiple channels, then the information sent over each channel is independently encoded, so that a packet sent on one channel may be decoded even if a collision occurs on another. Given that only one user transmits on channel j, let $R(\gamma_j)$ indicate the rate at which the user can reliably transmit as a function of the received power $\gamma_j = h_j P_j(\mathbf{h})$. We assume that $R(\gamma) := \log(1+\gamma)$, which is proportional to the Shannon capacity of the channel during a given time-slot. Provided that each time-slot is long enough to permit the use of sophisticated codes, this will give a reasonable indication of the transmission rate in a practical system. During a time-slot, each user will encode $R(\gamma_j)T$ bits into a packet to be transmitted on channel j. We assume that there is no coding done over successive time-slots. Also, we do not consider any multiuser reception or power capture effects when multiple users transmit on a channel. Given that users are received at different power levels, ignoring such effects is a questionable assumption. This is done mainly to simplify our analysis. Such techniques can be incorporated into our model and will result in improved performance.[3]

III. Optimal Distributed Power Allocation

In the above model, the power allocation $\mathbf{P}(\mathbf{h})$ for each user needs to be determined. To begin, consider the simplest case, where there is only $n = 1$ user who must allocate its power over the k available channels. In this case, for each channel realization $\mathbf{h}$, the power allocation that maximizes a user's throughput is the well-known "water-filling" allocation,

$$P_j^{wf}(\mathbf{h}) = \left(\lambda - \frac{1}{h_j}\right)^+, \tag{2}$$

where λ is chosen so that $\sum_{j=1}^k P_j^{wf}(\mathbf{h}) = \check{P}$ (see, e.g. [1]).

When there are multiple users, if more than one user transmits on a channel, a collision results and no data is received. We consider a natural extension of the Aloha-based approach in [8],

[3]Moreover, since we achieve order optimal performance without these; such techniques can not improve the growth rate, only the first order constant.

where each user transmits on each channel with a certain probability p. Since each channel is i.i.d., it is reasonable to require that each user transmits with the same probability p in each slot and on each channel. The probability of some user successfully transmitting on one channel is then $np(1-p)^{n-1}$. Given this probability, for each channel j, each user chooses a subset $\mathcal{H}_j$ of the possible realizations of $\mathbf{H}$ with $\Pr(\mathbf{H} \in \mathcal{H}_j) = p$. The user then only transmits on channel j when $\mathbf{H} \in \mathcal{H}_j$. To maximize the total throughput, each user will choose channel states in each set $\mathcal{H}_j$ that can achieve higher transmission rates. However, the transmission rate that can be achieved also relies on the specific power allocation, *e.g.* if a state $\mathbf{h}$ is in both $\mathcal{H}_j$ and $\mathcal{H}_l$, the user must allocate power across both channels, while if $\mathbf{h}$ is in only one set, the user can use all the available power on the corresponding channel. For a given power allocation, $P_j(\mathbf{h})$, the expected transmission rate on channel j, conditioned on a user successfully transmitting on that channel is given by

$$\mathbb{E}_{\mathbf{H}}\left(R(H_j P_j(\mathbf{H}))\middle|\mathbf{H} \in \mathcal{H}_j\right) = \mathbb{E}_{\mathbf{H}}\left(R(H_j P_j(\mathbf{H}))\middle|P_j(\mathbf{H}) > 0\right),$$

where we have used that the channels are independent across users. We now specify the following *distributed optimal throughput problem*:

$$\begin{aligned}
&\max_{\mathbf{P}(\mathbf{H}),p} np(1-p)^{n-1}\sum_{j=1}^{k}\mathbb{E}_{\mathbf{H}}\left(R(H_j P_j(\mathbf{H}))\middle|P_j(\mathbf{H}) > 0\right)\\
&\text{s.t.} \quad \sum_{j=1}^{k} P_j(\mathbf{h}) \leq \check{P},\ \forall\, \mathbf{h} \\
&\qquad\ \Pr\{P_j(\mathbf{H}) > 0\} = p,\ j = 1, ..., k.
\end{aligned} \tag{3}$$

The objective in (3) is the average sum throughput for all n users over all k channels. This is optimized over the transmission probability p and the power allocation $(P_1(\mathbf{H}), P_2(\mathbf{H}), ..., P_k(\mathbf{H}))$, which is used by each user. The second constraint ensures that the sets $\mathcal{H}_j$ all have probability p. When this constraint is met, it follows that

$$p\mathbb{E}_{\mathbf{H}}\left(R(H_j P_j(\mathbf{H}))\middle|P_j(\mathbf{H}) > 0\right) = \mathbb{E}_{\mathbf{H}}\left(R(H_j P_j(\mathbf{H}))\right).$$

Hence, the objective in (3) can also be written as

$$n(1-p)^{n-1}\sum_{j=1}^{k}\mathbb{E}_{\mathbf{H}}\left(R(H_j P_j(\mathbf{H}))\right). \tag{4}$$

In the remainder of this section, we will characterize the solution to (3) and give an algorithm for determining this solution. We begin with some preliminary results. First, recall that when $n = 1$ the optimal solution to (3) is for the user to use the water-filling power allocation, $\mathbf{P}^{wf}(\mathbf{h})$, for all channel realizations $\mathbf{h}$. Let

$$p^{wf} := \Pr(P_j^{wf}(\mathbf{H}) > 0) \tag{5}$$

be the probability that a user will transmit on a given channel j when using this allocation. Since the channel gains are i.i.d., this will be the same for all channels j. For $n > 1$ users, the transmission probability also influences the probability of success. The next lemma states that for any number of users the optimal p will never be larger than p^{wf}.

Lemma 1: For any number of users, n, the optimal solution to (3) will satisfy $p \leq p^{wf}$.

The proof is given in Appendix A. Note that when $n > 1$, the optimal p may be strictly less than p^{wf} to decrease the probability of collisions.

The next property of the optimal solution to (3) involves the symmetry of the allocation. We define a power allocation $\mathbf{P}(\mathbf{h})$ to be *symmetric* if the power allocation for any permutation of a channel realization $\mathbf{h}$ is equal to the same permutation of $\mathbf{P}(\mathbf{h})$. For example if $k = 2$, then for a symmetric power allocation $P_1(h_a, h_b) = P_2(h_b, h_a)$ for all h_a, h_b. Note that $\mathbf{P}^{wf}(\mathbf{h})$ is always symmetric.

Lemma 2: For any n and k, there exists an optimal power allocation for (3) that is symmetric.

We omit the proof due to space considerations. For a given channel realization $\mathbf{h}$, let $(h_{(1)}, h_{(2)}, ..., h_{(k)})$ denote the ordered channel gains from the largest to the smallest, with any ties broken arbitrarily. If a power allocation is symmetric it will just depend on this ordered sequence in each time-slot. Given this ordered sequence, for $j \leq l \leq k$, let $R^l_{(j)}(\mathbf{h})$ denote the rate achievable over the jth best channel when the transmitter uses the optimal (water-filling) power allocation over only the l best channels. In other words, $R^l_{(j)}(\mathbf{h}) = \log(1 + P_{(j)}(\mathbf{h})h_{(j)})$, where $P_{(j)}(\mathbf{h}) = (\lambda - \frac{1}{h_{(j)}})^+$ and λ is chosen such that $\sum_{j=1}^{l} P_{(j)}(\mathbf{h}) = \check{P}$.

Lemma 3: As l increases, $\sum_{i=1}^{l} R^l_{(i)}(\mathbf{h}) - \sum_{i=1}^{l-1} R^{l-1}_{(i)}(\mathbf{h})$ is non-increasing.

The proof is given in Appendix B. Given a "threshold rate" $R_{th} > 0$ for each channel realization $\mathbf{h}$, we introduce the following problem:

$$\begin{aligned} &\max_{l=1,\ldots,k} l \\ \text{s.t.} \quad &\sum_{i=1}^{l} R_{(i)}^{l}(\mathbf{h}) - \sum_{i=1}^{l-1} R_{(i)}^{l-1}(\mathbf{h}) \geq R_{th} \end{aligned} \tag{6}$$

If this problem has no feasible solution, we define the solution to be $l = 0$. When $k = 1$, the constraint in (6) is $R_{(1)}^{1}(\mathbf{h}) \geq R_{th}$, i.e., the rate when only transmitting on the best channel should be greater than R_{th}. For $k = 2$, the constraint in (6) becomes $R_{(1)}^{2}(\mathbf{h}) + R_{(2)}^{2}(\mathbf{h}) - R_{(1)}^{1}(\mathbf{h}) \geq R_{th}$, which means that the increase in the total rate from using the best two channels versus only using the best channel should be greater than R_{th}. In general, the objective of (6) is to find the maximal number of channels l, such that the gain in the sum rate from transmitting on the l best channels instead of only the $l-1$ best channels is at least R_{th}. Note from Lemma 3 it follows that if l^* solves (6), then any $l < l^*$ will also satisfy the constraint in (6).

For a given R_{th}, let $\mathbf{P}^{R_{th}}(\mathbf{h})$ be the power allocation that corresponds to solving (6) for each channel realization $\mathbf{h}$; i.e. this will be a water-filling allocation over the l best channels, where l is the solution to (6) for each given realization (note l may change with each realization). The following proposition relates this to the solution of (3).

Proposition 1: There exists a constant $R_{th} > 0$ such that $\mathbf{P}^{R_{th}}(\mathbf{h})$ is also the optimal solution to (3).

The proof is given in Appendix C. This proposition specifies the form of the optimal power allocation; the corresponding transmission probability is given by $p = \Pr(P_i^{R_{th}}(\mathbf{H}) > 0)$. It follows from this proposition that the optimal solution to (3) can be found by solving (6) for a given R_{th}, and then iteratively searching for the optimal R_{th}. Solving (6) for a given R_{th} and channel realization $\mathbf{h}$ can be done via Algorithm 1, which determines the set of channels from $\mathbf{h}$ that are transmitted on.

After each iteration in Algorithm 1, according to Lemma 3, the rate gain $d - d_{-1}$ decreases. Therefore, after at most k steps, the algorithm converges, and it converges to the optimal solution to (6). Note that a feasible solution might not exist for some channel realizations, in which case the algorithm returns $\mathcal{W} = \emptyset$. For a given R_{th}, this algorithm can be used to find the corresponding power allocation. The optimal value of R_{th} can then be found via a numeric search; however, we note that this search is now only a one-dimensional search, instead of a k-

Algorithm 1 k-best channels $(\mathbf{h}, R_{th})$

initialize:

$\mathcal{M} = \{1, \dots, k\}$

$j = \arg\max_{i \in \mathcal{M}} h_i$

$\mathcal{W} = \{j\}$

$d_{-1} = 0$

Water-fill over channels $\{h_i : i \in \mathcal{W}\}$ giving sum rate d.

if $d < R_{th}$ **then**

$\mathcal{W} = \emptyset$

else

while $d - d_{-1} > R_{th}$ **do**

$d_{-1} = d$

$\mathcal{M} = \mathcal{M}/\mathcal{W}$

$j = \arg\max_{i \in \mathcal{M}} h_i$

$\mathcal{W} = \{j\} \bigcup \mathcal{W}$

Water-fill over channels $\{h_i : i \in \mathcal{W}\}$ giving sum rate d.

end while

end if

return $\mathcal{W}$

dimensional search over the possible power allocations.[4] An example of the resulting allocation for a system with $k = 2$ channels is given in Fig. 2, which shows the channel states (h_1, h_2) during which a user will transmit for a given transmission probability p. For convenience, the axis in this figure are the inverse of the channel gains, i.e. $\frac{1}{h_1}$ and $\frac{1}{h_2}$. The double crossed area indicates the channel states for which a user should transmit over both channels; the single crossed areas are when the user should transmit on only one channel. The area of each single crossed region has a total probability of p.

Figure 2 around here.

[4]Note that the objective in (6) is the average throughput over all channel states; evaluating this requires finding the power allocation for each state, e.g. executing algorithm 1 for each state. For a large number of states, this can be approximated via Monte Carlo simulation.

For a given n and k, the optimal power allocation could be determined offline using the above procedure. However, for a large number of channels k this will result in a large computational cost. In the next section, we introduce a simpler sub-optimal algorithm and analyze its performance.

IV. Sub-optimal Power Allocation and Asymptotic Analysis

We consider a simplified distributed scheme, where instead of finding a threshold rate R_{th} and solving (6), we set a threshold h_{th} on the channel gain. Each user then transmits on the kth channel when its gain is greater than h_{th}, resulting in the transmission probability $p = \bar{F}_H(h_{th})$. If a user has more than one channel whose gain is higher than the threshold, then the total power $\check{P}$ will be allocated equally to each of these channels.[5] Given that a user transmits on i channels, we assume it transmits at a constant rate of $R_i(p) := R\left(\bar{F}_H^{-1}(p)\frac{\check{P}}{i}\right)$ on each channel. This is a lower bound on the achievable rate and simplifies our analysis.

The total throughput using this scheme is a function of k, n and p. For $i = 1, ..., k$, let $q_{k,p}(i)$ be the probability one user has i channels above the threshold $h_{th} = \bar{F}_H^{-1}(p)$, *i.e.*,

$$q_{k,p}(i) = \binom{k}{i}(p)^i(1-p)^{k-i}.$$

Among these i channels, for $j = 1, ..., i$, let $\omega_{p,i}(j)$ be the probability a user transmits successfully on exactly j channels, *i.e.* the probability there is no collision on exactly j channels, given that i are above the threshold. This is given by

$$\omega_{p,i}^n(j) = \binom{i}{j}\left[(1-p)^{n-1}\right]^j\left[1-(1-p)^{n-1}\right]^{i-j}.$$

The average sum throughput of this system is then given by

$$s(k,n,p) = n\sum_{i=1}^{k} q_{k,p}(i)\sum_{j=1}^{i}\omega_{p,i}^n(j)jR_i(p).$$

Note that $\omega_{p,i}^n(j)$ is a Binomial probability mass function (p.m.f.) and so $\sum_{j=1}^{i}\omega_{p,i}^n(j)j = (1-p)^{n-1}i$. Therefore,

$$s(k,n,p) = n(1-p)^{n-1}\sum_{i=1}^{k}\binom{k}{i}(p)^i(1-p)^{k-i}iR_i(p). \quad (7)$$

[5]Other equal power allocation approaches for multi-carrier systems have been studied in [11] for a single user channel and in [12] for a downlink channel.

The next lemma gives upper and lower bounds on $s(k,n,p)$,

Lemma 4: For all k, n and p,

$$\log\left(1+\frac{\check{P}\bar{F}_H^{-1}(p)}{p(k-1)+1}\right) \leq \frac{s(k,n,p)}{n(1-p)^{n-1}kp} \leq \log\left(1+\frac{\check{P}\bar{F}_H^{-1}(p)[1-(1-p)^k]}{kp}\right).$$

The proof is given in Appendix D.

Using these bounds, we next consider how the sum throughput of this scheme and the optimal scheme scales in the three asymptotic regimes given in the introduction. We define two sequences $f(m)$ and $g(m)$ to be *asymptotically equivalent*, denoted by $f(m) \asymp g(m)$, if $\lim_{m\to\infty}\frac{f(m)}{g(m)} = c$. In the special case where $c = 1$, we say that they are *strongly asymptotically equivalent* and denote this by $f(m) \stackrel{\scriptscriptstyle{-}}{\asymp} g(m)$. This implies that both sequences asymptotically grow at the same rate and have the same first order constant. In each regime, we show that this simplified scheme is strongly asymptotically equivalent to the optimal distributed algorithm (*i.e.*, the solution to (3)). For these results, we need to make an additional assumption on the tail of the fading distribution. Specifically, we assume that as $h \to \infty$,

$$f_H(h) \asymp f_H'(h), \tag{8}$$

where $f_H'(h) = \frac{d}{dh}f_H(h)$. This is satisfied by any fading distribution that has an exponential tail, which is the case for most common fading models such as Rayleigh, Ricean and Nakagami fading. The follow lemma summarizes several other useful properties of such a fading distribution.

Lemma 5: For any continuous, differentiable fading density f_H that satisfies (8), then the following conditions hold: (*a*.) $\bar{F}_H(h) \asymp f_H(h)$, (*b*.) $\lim_{h\to\infty}\frac{\bar{F}_H(h)}{hf_H(h)} = 0$, and (*c*.) $\lim_{h\to\infty}\frac{d}{dh}\left[\frac{\bar{F}_H(h)}{f_H(h)}\right] = 0$.

The proof of these conditions follows directly from evaluating the limits using L'Hospital's rule. Conditions (*b*) and (*c*) were used in [9] to characterize the asymptotic performance of a single channel system.

We also compare the throughput achieved by these distributed approaches to an optimal centralized system that schedules the users to maximize the throughput in every slot (still assuming at most one user can transmit on each channel). This is given by solving[6]

$$\begin{aligned} &\max_{\{P_{ij},c_{ij}\}} \sum_{i=1}^{n}\sum_{j=1}^{k} R(P_{ij}c_{ij}h_{ij}) \\ &\text{s.t.} \quad \sum_{j=1}^{k} P_{ij}c_{ij} = \check{P},\ \forall i, \\ &\qquad\quad \sum_{i=1}^{n} c_{ij} \leq 1,\ \forall j,\ c_{nk}\in\{0,1\},\ \forall i,j, \end{aligned} \tag{9}$$

during each time-slot. Here, the integer variables, c_{ij}, indicate when user i is assigned to channel j; the second constraint ensures that at most one user is assigned to each channel. Let $s_{ct}(k,n)$ be the sum throughput obtained by the optimal centralized scheduling policy that solves (9) in each time-slot, averaged over the channel distributions. Denote the throughput of the optimal distributed policy by $s^*(k,n)$ and the optimal throughput of the threshold-based algorithm by $s(k,n,p^*)$, where p^* is the transmission probability that optimizes $s(k,n,p)$. For all n and k, from their definitions we have,

$$s\left(k,n,\tfrac{1}{n}\right) \leq s(k,n,p^*) \leq s^*(k,n) \leq s_{ct}(k,n), \tag{10}$$

where the first term is the throughput with a transmission probability of $1/n$.

First, we consider the case where k is fixed and n increases.

Proposition 2: Given any finite k, as $n\to\infty$, $s(k,n,\frac{1}{n})$, $s(k,n,p^*)$, $s^*(k,n)$ and $\frac{1}{e}s_{ct}(k,n)$ are all strongly asymptotically equivalent to $\frac{k}{e}\log\left(1+\check{P}\bar{F}_H^{-1}(\frac{1}{n})\right)$.

The proof is given in Appendix E. This proposition states that asymptotically there is no difference in the first-order performance compared to the optimal distributed approach when using the simplified scheme or from choosing $p=\frac{1}{n}$ instead of the optimal p^*. The throughput for each distributed approach asymptotically increases like $\frac{k}{e}\log(1+\check{P}\bar{F}_H^{-1}(\frac{1}{n}))$, as does $\frac{1}{e}$ times the throughput with the optimal centralized scheduler. In other words, the distributed approaches all grow at the same rate as the centralized approach and asymptotically the ratio of their throughputs approach $\frac{1}{e}$, the contention loss in a standard slotted Aloha system. As an example, for the case of i.i.d. Rayleigh fading on each channel the throughput in each case will increase at rate $O(\log(\log(n)))$.

[6]This is similar to a problem studied for centralized OFDM systems in [6].

The second regime we consider is when n is fixed and k increases.

Proposition 3: Given any finite n, as $k \to \infty$, $s(k,n,p^*)$, $s^*(k,n)$, $s_{ct}(k,n)$ are all strongly asymptotically equivalent to $n\check{P}\bar{F}_H^{-1}(\frac{1}{k})$.

The proof is given in Appendix F. This proposition states that again the threshold based approach is strongly asymptotically equivalent to the optimal distributed approach. In this case, it is also asymptotically equivalent to the optimal centralized system; i.e. there is no loss of $\frac{1}{e}$. Intuitively, this is because as the number of channels increases, the probability of collision becomes negligible. In this case, for a Rayleigh fading channel each of these terms grows like $O(\log(k))$ as $k \to \infty$, with a first order constant that is linear in n.

The last regime we consider is where both k and n increase with fixed ratio $\frac{k}{n} = \beta$.

Proposition 4: If $\frac{k}{n} = \beta$, as $n \to \infty$, $s(\beta n, n, \frac{1}{n})$, $s(\beta n, n, p^*)$, $s^*(\beta n, n)$ and $\frac{1}{e}s_{ct}(\beta n, n)$ are all strongly asymptotically equivalent to $\beta n e^{-1}\log\left(1 + \check{P}\bar{F}_H^{-1}(\frac{1}{n})\right)$.

The proof uses similar ideas to Proposition 2 and is given in Appendix G. As in Proposition 2, once again compared to the centralized scheme there is an asymptotic penalty of $1/e$ due to the contention, and a transmission probability of $p = \frac{1}{n}$ is asymptotically optimal for the distributed system. For Rayleigh fading channels the throughput now grows like $O(n\log(\log(n)))$, as $n \to \infty$, with a first order constant that is linear in β.

V. Comparison with Other Distributed Approaches

We next compare the distributed approaches to several schemes that require minimal coordination for assigning different power allocation policies to different sets of users. First, assume that $\frac{k}{n} = \beta$, where β is a positive integer. In this case, a "non-collision scheme" is to assign β channels to each user for all time. Let $s_{nc}(k,n)$ denote the average sum throughput of this scheme. It follows that $s_{nc}(k,n) = ns_{ct}(\beta,1)$. Comparing this to $s(\beta n, n, p^*)$, we have

$$\lim_{n\to\infty}\frac{s_{nc}(\beta n,n)}{s(\beta n,n,p^*)} = \lim_{n\to\infty}\frac{s_{ct}(\beta,1)}{\beta e^{-1}\log(1+\check{P}\bar{F}_H^{-1}(\frac{1}{n}))} = 0,$$

where the first equality follows from Proposition 4. In other words, when enough users are present, this non-collision scheme will perform worse than the simplified distributed approach. This is because the non-collision scheme cannot exploit any "multi-user" diversity. Hence, it has a constant throughput as n increases, while $s(\beta n, n, p^*)$ is unbounded. On the other hand, when n is fixed and β (*i.e.* k) increases, then from Proposition 3, $s_{ct}(\beta,1) \asymp \check{P}F_H^{-1}(\frac{1}{\beta})$. Hence, as $k \to \infty$

$$s_{nc}(k,n) \asymp n\check{P}F_H^{-1}(\frac{1}{\beta}) \asymp n\check{P}F_H^{-1}(\frac{1}{k}) \asymp s(n,k,p^*).$$

In this case, both approaches see increased frequency diversity as k increases and are asymptotically equivalent.

Next, assume $\alpha = \frac{n}{k} > 1$. In this case, an approach with fewer collisions is to assign α users to each channel for all time. These α users then use the optimal distributed protocol for a single channel. This results in a throughput of $s_{lc}(k,n) = ks^*(1,\alpha)$. Using Proposition 4, it follows that for fixed α, as n (and k) increase, we have

$$\lim_{n\to\infty} \frac{s_{lc}(n/\alpha,n)}{s(n/\alpha,n,p^*)} = \lim_{n\to\infty} \frac{s^*(1,\alpha)}{e^{-1}\log(1+\check{P}\bar{F}_H^{-1}(\frac{1}{n}))} = 0.$$

Again, the new scheme cannot fully exploit the available diversity. However, if k is fixed so that α increases with n, then from Proposition 2, $s^*(1,\alpha) \asymp \frac{1}{e}\check{P}\log(F_H^{-1}(\frac{1}{\alpha}))$. Hence,

$$s_{lc}(k,n) \asymp \frac{K}{e}\check{P}h_0 \log\left(F_H^{-1}(\frac{1}{\alpha})\right) \asymp \frac{K}{e}\check{P}h_0 \log(F_H^{-1}(\frac{1}{n}))) \asymp s(k,n,\frac{1}{n}).$$

In this case, the throughput ratio of the two schemes approaches to 1 asymptotically because both exploit increasing multiuser diversity as n increases.

To summarize, we have seen that grouping users to avoid contention is not desirable unless the users can still exploit the available diversity. Even when the users can exploit the diversity, such approaches do not improve on the first order asymptotic performance.

VI. Numerical Examples

In this section, we give some numerical examples to illustrate the performance of the optimal and simplified distributed algorithms with a finite number of channels and users. All the results in this section are for an i.i.d. Rayleigh fading model, with $E(H_{ij}) = 1$, and a total power constraint of $\check{P} = 1$. The performance is averaged over multiple channel realizations. Figure 3 shows the average throughput achieved by the optimal distributed power allocation scheme from Section III compared to the simplified power allocation scheme in Section IV. The throughput of both approaches is shown as a function of the number of users for a system with $k = 10$ channels. As the number of users increases, both throughputs increase and the difference between the two curves decreases.

Figure 4 shows upper and lower bounds on the ratio of the average throughput of the optimal distributed scheme $s^*(k,n)$ to the centralized scheme $s_{ct}(k,n)$ defined in (9) as a function of the number of users, for $k = 5$ and 10 channels. Calculating $s_{ct}(k,n)$ requires solving the optimization problem in (9) for every channel realization, which is complicated due to the integer constraints. Instead we compare $s^*(k,n)$ to upper and lower bounds on $s_{ct}(k,n)$. We upper bound $s_{ct}(k,n)$ by relaxing the total power constraint on the channels, $\sum_k P_{nk}c_{nk} = \check{P}$. Instead, we allow each user to transmit with $P_{nk} = \check{P}$ over each channel. The maximum throughput is achieved for this relaxed system by letting the best user on each channel transmit at each time. We take the resulting throughput as our upper bound. To lower bound $s_{ct}(k,n)$, we still choose the best user to transmit on each channel, but if one user is chosen to transmit on more than one channels, its power $\check{P}$ is divided equally across these channels. The resulting throughput is then a lower bound on $s_{ct}(k,n)$. Figure 4 shows that as the number of users increases, the two bounds approach each other. The reason is that the probability that one user is chosen to transmit on more than one channel is small for a larger number of users. It can be seen that the ratio of the throughputs of the distributed to the optimal scheme is decreasing as the number of users increases and is larger than the limiting value of $1/e$ (see Proposition 2) for all finite n. As the number of the channels, k, increases, the throughput ratio also increases for a fixed n. This is due to the increased frequency diversity with more channels.

Figure 5 shows upper and lower bounds on the ratio of the throughput of the optimal distributed scheme to that of the optimal centralized approach as the number of channels increases, for a system with $n = 5$ and 10 users. In this case, we upper bound $s_{ct}(k,n)$ by the information theoretic capacity of this multi-access system. In other words, joint decoding is used when multiple users transmit on the same channel. We use the iterative water-filling algorithm from [4] to obtain this capacity. One channel can be assigned to multiple users to achieve the capacity. By only allowing the user who has the best channel to transmit on that channel, we obtain a lower bound of the system. Figure 5 shows that as the number of channels increases, the two bounds quickly converge. The throughput ratio increases as the number of channels increases. From Proposition 3, as k increases, these bounds should approach 1. In this asymptotic regime, the convergence appears to be much slower than in Figure 4.

Figure 3 around here.

Figure 4 around here.

Figure 5 around here.

VII. Summary

In this paper we have presented distributed algorithms for scheduling and power allocation in a parallel channel wireless network, where each user only has knowledge of its own channel gains. Using a contention model, an optimal distributed algorithm is characterized. A simplified distributed approach is also given. In three different asymptotic regimes, the simplified algorithm is shown to be asymptotically equivalent to the optimal distributed algorithm. Both algorithms are also shown to scale at the same rate as the optimal centralized scheduler. These results suggest that it is possible to develop near optimal approaches for scheduling and power allocation without requiring a centralized controller with complete channel knowledge. There are several important issues that we have not addressed here. For example, we have not considered asymmetric models, where the fading is not identically distributed across the channels or the users, or models where the fading is correlated across the channels. We also assumed that each user knows the fading distribution; in practice, an adaptive approach would be required to estimate this distribution.

Appendix A
Proof of Lemma 1:

To establish a contradiction, suppose that for a given number of users, n, the optimal p satisfies $p > p^{wf}$. This will require that each user transmits on a larger set of channels than they would under $\mathbf{P}^{wf}(\mathbf{h})$. Let $\tilde{\mathbf{P}}(\mathbf{h})$ denote the resulting power allocation. Since $\mathbf{P}^{wf}(\mathbf{h})$ maximizes the sum throughput for one user, it must be that

$$\sum_{j=1}^{k} \mathbb{E}_{\mathbf{H}} \left(R(H_j P_j^{wf}(\mathbf{H})) \right) > \sum_{j=1}^{k} \mathbb{E}_{\mathbf{H}} \left(R(H_j \tilde{P}_j(\mathbf{H})) \right). \tag{11}$$

Also since $p > p^{wf}$,

$$n(1-p^{wf})^{n-1} > n(1-p)^{n-1}. \tag{12}$$

Substituting (11) and (12) into (4), it follows that using $\mathbf{P}^{wf}(\mathbf{h})$ results in a larger throughput. Hence, p cannot be optimal. ∎

APPENDIX B
PROOF OF LEMMA 3:

Given a channel realization $\mathbf{h}$, consider the optimal (water-filling) power allocation over the following $2(l-1)$ channels $h_{(1)}, h_{(1)}, h_{(2)}, h_{(2)}, ..., h_{(l-1)}, h_{(l-1)}$, with total power $2\check{P}$. The resulting sum rate will be $2\sum_{i=1}^{l-1} R_{(i)}^{l-1}(\mathbf{h})$. Next, consider

$$\sum_{i=1}^{l} R_{(i)}^{l}(\mathbf{h}) + \sum_{i=1}^{l-2} R_{(i)}^{l-2}(\mathbf{h}),$$

this rate can be achieved by some power allocation over $2(l-1)$ channels with channel gains $h_{(1)}, h_{(1)}, ..., h_{(l-2)}, h_{(l-2)}, h_{(l-1)}, h_{(l)}$, which satisfies the same power constraint $2\check{P}$. Consider using the same power allocation as in the first case. Since $h_{(l)} \leq h_{(l-1)}$, and we are using the optimal power allocation for the first case, it can be seen that

$$2\sum_{i=1}^{l-1} R_{(i)}^{l-1}(\mathbf{h}) \geq \sum_{i=1}^{l} R_{(i)}^{l}(\mathbf{h}) + \sum_{i=1}^{l-2} R_{(i)}^{l-2}(\mathbf{h}).$$

Therefore, rearranging terms we have

$$\sum_{i=1}^{l-1} R_{(i)}^{l-1}(\mathbf{h}) - \sum_{i=1}^{l-2} R_{(i)}^{l-2}(\mathbf{h}) \geq \sum_{i=1}^{l} R_{(i)}^{l}(\mathbf{h}) - \sum_{i=1}^{l-1} R_{(i)}^{l-1}(\mathbf{h}).$$

■

APPENDIX C
PROOF OF PROPOSITION 1:

First note that for a given rate threshold R_{th}, $\mathbf{P}^{R_{th}}(\mathbf{h})$ will result in a transmission probability, $p(R_{th}) = \Pr(P_i^{R_{th}}(\mathbf{H}) > 0)$.[7] This probability will satisfy $p(0) = p^{wf}$, where p^{wf} is defined in (5), and $p(R_{th})$ will be a monotonically decreasing function of R_{th}.

Let $\mathbf{P}(\mathbf{h})$ be the optimal power allocation for problem (3) with a given n and k, and let p^* be the corresponding transmission probability. From Lemma 1, $p^* \leq p^{wf}$, and from Lemma 2, we can assume that $\mathbf{P}(\mathbf{h})$ is symmetric. It can be easily shown that the optimal allocation must also satisfy the following two properties: first, given an ordered channel realization $(h_{(1)}, h_{(2)}, ..., h_{(k)})$, if the ith best channel is allocated positive power, then the channels $h_{(1)}, .., h_{(i-1)}$ whose gains are no worse than $h_{(i)}$ should also be allocated positive power; second, for any set of channel

[7]By construction, it can be seen that this probability will be the same for each channel i.

states $h_{(1)}, .., h_{(l)}$ allocated positive power, in order to maximize the total transmission rate, the water-filling power allocation should be used over these states, resulting in rate $\sum_{i=1}^{l} R^{l}_{(i)}(\mathbf{h})$. Also note that from Lemma 3, if $\sum_{i=1}^{l} R^{l}_{(i)}(\mathbf{h}) - \sum_{i=1}^{l-1} R^{l-1}_{(i)}(\mathbf{h}) < R_{th}$, then $\sum_{i=1}^{m} R^{m}_{(i)}(\mathbf{h}) - \sum_{i=1}^{m-1} R^{m-1}_{(i)}(\mathbf{h}) < R_{th}$, for all $m > l$.

Let $\mathbb{H}$ be the k dimensional space of possible channel state vectors $\mathbf{h}$. Also, for $l = 1, \ldots, k$, let $s_l(\mathbf{h}) = \sum_{i=1}^{l} R^{l}_{(i)}(\mathbf{h})$ and $s_0(\mathbf{h}) = 0$. We complete the proof by contradiction. Namely, let R^*_{th} be the rate threshold that satisfies $p(R^*_{th}) = p^*$, and assume that $\mathbf{P}(\mathbf{h})$ is not equal to the solution to (6) with this rate threshold for some set of channel states with positive probability.[8] In particular, for some channel j, there must exist a set of states $\mathcal{N} \in \mathbb{H}$ such that for all $\mathbf{h} \in \mathcal{N}$, j is the lth best channel (for any l) with $s_l(\mathbf{h}) < s_{l-1}(\mathbf{h}) + R^*_{th}$, and j is allocated positive power. Likewise, to ensure that $Pr\{P_j(\mathbf{H}) > 0\} = p^*$, there must be another set of states $\mathcal{B} \in \mathbb{H}$ (with the same probability) such that when j is the lth best state and $s_l(\mathbf{h}) > s_{l-1}(\mathbf{h}) + R^*_{th}$, then j is not allocated positive power. However, by transmitting on channel j in $\mathcal{B}$ instead of $\mathcal{N}$, and using the above properties of the optimal allocation, it can be shown that the total throughput will increase, which contradicts this power allocation being optimal. ■

APPENDIX D
PROOF OF THE LEMMA 4:

Let

$$R_k = \sum_{i=1}^{k} \binom{k}{i} (p)^i (1-p)^{k-i} i \log\left(1 + \frac{\check{P}\bar{F}_H^{-1}(p)}{i}\right)$$
$$= kp \sum_{j=0}^{k-1} \binom{k-1}{j} p^j (1-p)^{k-1-j} \log\left(1 + \frac{\check{P}\bar{F}_H^{-1}(p)}{j+1}\right).$$

The function $g(x) = \log(1 + \frac{\check{P}\bar{F}_H^{-1}(p)}{x+1})$ can be shown to be a convex function, for $x > 0$. Also

$$\Pi_j = \binom{k-1}{j} p^j (1-p)^{k-1-j}, j = 0..k-1$$

is a p.m.f. Hence, using Jenson's inequality, we have

$$R_k \geq kp \log\left(1 + \frac{\check{P}\bar{F}_H^{-1}(p)}{p(k-1)+1}\right).$$

[8]Note since $p^* \leq p^{wf}$ such a R^*_{th} must exist.

Therefore,

$$s(k,n,p) \geq n(1-p)^{n-1} kp \log\left(1 + \frac{\check{P}\bar{F}_H^{-1}(p)}{p(k-1)+1}\right),$$

which gives the desired lower bound.

Next, we derive the upper bound. Let $f(i) = i\log(1 + \frac{\check{P}\bar{F}_H^{-1}(p)}{i})$, so that

$$\begin{aligned} R_k &= \sum_{i=1}^{k} \binom{k}{i} (p)^i (1-p)^{k-i} f(i) \\ &= \frac{\sum_{i=1}^{k} \binom{k}{i} (p)^i (1-p)^{k-i} f(i)}{\sum_{i=1}^{k} \binom{k}{i} (p)^i (1-p)^{k-i}} \sum_{i=1}^{k} \binom{k}{i} (p)^i (1-p)^{k-i}. \end{aligned}$$

The function $f(i)$ can be shown to be concave for $i > 0$. Taking expectation with respect to the p.m.f.

$$\Pi_i = \frac{\binom{k}{i} (p)^i (1-p)^{k-i}}{\sum_{i=1}^{k} \binom{k}{i} (p)^i (1-p)^{k-i}}, i = 1, \ldots, k,$$

and again using Jenson's inequality, we have

$$\begin{aligned} R_k &= \mathbb{E} f(i)[1-(1-p)^k] \\ &\leq f(\mathbb{E} i)[1-(1-p)^k] \\ &= f\left(\frac{kp}{1-(1-p)^k}\right)[1-(1-p)^k] \\ &= kp\log\left(1 + \frac{\check{P}\bar{F}_H^{-1}(p)[1-(1-p)^k]}{kp}\right). \end{aligned}$$

Therefore, we have the upper bound

$$s(k,n,p) \leq n(1-p)^{n-1} kp \log\left(1 + \frac{\check{P}\bar{F}_H^{-1}(p)[1-(1-p)^k]}{kp}\right).$$

■

APPENDIX E

PROOF OF PROPOSITION 2:

From (10), to prove that $s\left(k,n,\frac{1}{n}\right)$, $s(k,n,p^*)$, and $s^*(k,n)$ are all asymptotically equivalent to $\frac{k}{e}\log\left(1 + \check{P}\bar{F}_H^{-1}(\frac{1}{n})\right)$, it is sufficient to show that

$$\lim_{n\to\infty} \frac{s(k,n,\frac{1}{n})}{\frac{k}{e}\log\left(1 + \check{P}\bar{F}_H^{-1}(\frac{1}{n})\right)} \geq 1, \tag{13}$$

and

$$\lim_{n\to\infty} \frac{s^*(k,n)}{\frac{k}{e}\log\left(1+\check{P}\bar{F}_H^{-1}(\frac{1}{n})\right)} \leq 1. \tag{14}$$

From Lemma 4,

$$s(k,n,\tfrac{1}{n}) \geq (1-\frac{1}{n})^{n-1} k \log\left(1+\frac{\check{P}\bar{F}_H^{-1}(\frac{1}{n})}{n^{-1}(k-1)+1}\right).$$

Letting $s_l(k,n,\frac{1}{n})$ denote this lower bound, it can be seen that

$$\lim_{n\to\infty} \frac{s_l(k,n,\frac{1}{n})}{\frac{k}{e}\log\left(1+\check{P}\bar{F}_H^{-1}(\frac{1}{n})\right)} = 1.$$

It follows that (13) holds.

To show that (14) holds, consider a new system with k parallel channels, that is identical to the original system, except each user now has peak power constraint of $\check{P}$ for *each* channel, instead of having the total power across all the channels constrained by $\check{P}$. Denote the optimal distributed throughput of this new parallel system by $s_p^*(k,n)$. It follows that $s^*(k,n) \leq s_p^*(k,n) = ks^*(1,n)$. In [9], it is shown that for a single channel system, if $f_H(h)$ satisfies condition (*b*) in Lemma 5, then $s^*(1,n) \asymp \frac{1}{e}\log(1+\check{P}\bar{F}_H^{-1}(\frac{1}{n}))$. By assumption, $f_H(h)$ satisfies (8), and hence this condition. Combining the above observations, (14) follows.

Next, we show that $s(k,n,\frac{1}{n}) \asymp \frac{1}{e}s_{ct}(k,n)$. For this we again consider the parallel channel system where each user has a peak power constraint of $\check{P}$ for each channel. In this case, the optimal centralized system would simply schedule the user with the best channel gain to transmit on each channel using the maximum power $\check{P}$. Let $s_{ctu}(k,n)$ denote the throughput of this new system; this will clearly upper bound $s_{ct}(k,n)$ and will satisfy $s_{ctu}(k,n) = ks_{ct}(1,n)$. Again referring to [9], it is shown that for a single channel system, if $f_H(h)$ satisfies condition (*c*) in Lemma 5, then $s_{ct}(1,n) \asymp \log(1+\check{P}\bar{F}_H^{-1}(\frac{1}{n}))$. Likewise, the above results show that $s(k,n,\frac{1}{n}) \asymp \frac{k}{e}\log(1+\check{P}\bar{F}_H^{-1}(\frac{1}{n}))$. Combining these, it follows that

$$\lim_{n\to\infty} \frac{s(k,n,\frac{1}{n})}{s_{ct}(k,n)} \geq \frac{1}{e}. \tag{15}$$

To complete the proof we lower bound $s_{ct}(k,n)$. In this case, we consider a sub-optimal centralized scheduler for the original system. Specifically, the system schedules the user who has the best channel to transmit on each channel. However, if one user is scheduled for more than one channel, it will split its total power and allocate equal power to each channel. Denote the throughput of this model by $s_{ctl}(k,n)$; this is given by

$$s_{ctl}(k,n) = n\sum_{i=1}^{k}\binom{k}{i}\left(\frac{1}{n}\right)^{i}\left(1-\frac{1}{n}\right)^{k-i} i\mathbb{E}\left(\log\left(1+\frac{\check{P}H_{max}}{i}\right)\right).$$

Clearly, $s_{ctl}(k,n) \leq s_{ct}(k,n)$.

It can be shown (see [9]) that if $f_H(h)$ satisfies condition (*c*) in Lemma 5, then for any i, as $n \to \infty$,

$$\mathbb{E}_{\mathbf{H}}\left(\log\left(1+\frac{\check{P}H_{max}}{i}\right)\right) \asymp \log\left(1+\frac{\check{P}\bar{F}_H^{-1}(\frac{1}{n})}{i}\right).$$

Comparing this to (7), we have that $s(k,n,\frac{1}{n}) \asymp \frac{1}{e}s_{ctl}(k,n)$, and so

$$\lim_{n\to\infty}\frac{s(k,n,\frac{1}{n})}{s_{ct}(k,n)} \leq \frac{1}{e}.$$

Combining this with (15), it follows that $s(k,n,\frac{1}{n}) \asymp \frac{1}{e}s_{ct}(k,n)$, as desired. ■

APPENDIX F

PROOF OF PROPOSITION 3:

Again from (10), it is sufficient to show that

$$\lim_{n\to\infty}\frac{s_{ct}(k,n)}{n\check{P}\bar{F}_H^{-1}(\frac{1}{k})} \leq 1, \tag{16}$$

and

$$\lim_{n\to\infty}\frac{s(k,n,p^*)}{n\check{P}\bar{F}_H^{-1}(\frac{1}{k})} \geq 1. \tag{17}$$

We first upper-bound $s_{ct}(k,n)$ by considering the optimal throughput in n parallel systems with one user each. For a given set of n channel realizations the n parallel systems must have a throughput no smaller than the original system, since any power allocation used in the original system is also feasible for the new system. In other words, $s_{ct}(k,n) \leq ns_{ct}(k,1)$. Next we upper bound $s_{ct}(k,1)$. Given any ordered realization of the channel gains, $(h_{(1)}, h_{(2)}, ..., h_{(k)})$, the maximum throughput in the centralized system is upper bounded by the maximum throughput over k parallel channels all with gain $h_{(1)}$, which is given by $k\log\left(1+\frac{\check{P}h_{(1)}}{k}\right)$. For all k, this satisfies

$$k\log\left(1+\tfrac{\check{P}h_{(1)}}{k}\right) \leq (\log_2 e)\check{P}h_{(1)}.$$

Combining the above observations, we have

$$s_{ct}(k,n) \leq n s_{ct}(k,1) \leq n \log_2 \mathbb{E}_H (\log_2 e) \check{P} H_{(1)}.$$

In [9], it is shown that if $f_H(h)$ satisfies condition (c) in Lemma 5, then

$$(\log_2 e)\mathbb{E}_H(\check{P} H_{(1)}) \stackrel{\sim}{\asymp} (\log_2 e)\check{P} F_H^{-1}(\tfrac{1}{k}).$$

Using this, (16) follows.

Next, we lower bound $s(k,n,p^*)$. Let $\tilde{p} = \frac{\log^2(k)}{k}$, for $k > 1$. Clearly, $s(k,n,p^*) \geq s(k,n,\tilde{p})$. From Lemma 4, $s(k,n,\tilde{p})$ is lower bounded by

$$s_l(k,n,\tilde{p}) = n \left(1 - \frac{\log^2(k)}{k}\right)^{n-1} \log^2(k) \log \left(1 - \frac{\check{P} \bar{F}_H^{-1}(\frac{\log^2(k)}{k})}{\log^2(k)\frac{k-1}{k} + 1}\right).$$

From Lemma 5, $f_H(h) \asymp \bar{F}_H(h)$. Using this it can be shown that

$$\lim_{k\to\infty} \frac{\check{P} \bar{F}_H^{-1}(\frac{\log^2(k)}{k})}{\log^2(k)\frac{k-1}{k} + 1} = 0,$$

and so, as $k \to \infty$,

$$s_l(k,n,\tilde{p}) \stackrel{\sim}{\asymp} n(\log_2 e)\check{P}\bar{F}_H^{-1}\left(\frac{\log^2(k)}{k}\right). \tag{18}$$

Now, we show that the condition $\bar{F}_H(h) \asymp f_H(h)$ implies that $\bar{F}_H^{-1}(\frac{\log^2(k)}{k}) \stackrel{\sim}{\asymp} \bar{F}_H^{-1}(\frac{1}{k})$. Since both $\lim_{k\to\infty} \bar{F}_H^{-1}(\frac{\log^2(k)}{k}) = \infty$ and $\lim_{k\to\infty} \bar{F}_H^{-1}(\frac{1}{k}) = \infty$, L'Hospital's rule can be applied, yielding

$$\begin{aligned}
\lim_{k\to\infty} \frac{\bar{F}_H^{-1}(\frac{\log^2(k)}{k})}{\bar{F}_H^{-1}(\frac{1}{k})} &= \lim_{k\to\infty} \frac{f_H(\bar{F}_H^{-1}(\frac{1}{k}))}{f_H(F_H^{-1}\left(\frac{\log^2(k)}{k}\right))} \cdot \left(\log^2(k) - 2\log(k)\right) \\
&= \lim_{k\to\infty} \frac{\bar{F}_H(\bar{F}_H^{-1}(\frac{1}{k}))}{\bar{F}_H\left(\bar{F}_H^{-1}(\frac{\log^2(k)}{k})\right)} \cdot (\log^2(k) - 2\log(k)) \\
&= 1,
\end{aligned}$$

where the second line follows from the fact that $\bar{F}_H(h) \asymp f_H(h)$. Combining this with (18), it follows that $s_l\left(k, n, \frac{\log^2(k)}{k}\right) \stackrel{\sim}{\asymp} n\check{P}\bar{F}_H^{-1}(\frac{1}{k})$, and so (17) holds, as desired. ■

APPENDIX G

PROOF OF THE PROPOSITION 4:

To prove that $s(\beta n, n, \frac{1}{n})$, $s(\beta n, n, p^*)$ and $s^*(\beta n, n)$ are all asymptotically equivalent to $\frac{\beta n}{e} \log\left(1 + \check{P}\bar{F}_H^{-1}(\frac{1}{n})\right)$, it is sufficient to show that

$$\lim_{n\to\infty} \frac{s(\beta n, n, \frac{1}{n})}{\frac{\beta n}{e} \log\left(1 + \check{P}\bar{F}_H^{-1}(\frac{1}{n})\right)} \geq 1, \tag{19}$$

and

$$\lim_{n\to\infty} \frac{s^*(\beta n, n)}{\frac{\beta n}{e} \log\left(1 + \check{P}\bar{F}_H^{-1}(\frac{1}{n})\right)} \leq 1. \tag{20}$$

From Lemma 4, $s(\beta n, n, \frac{1}{n}) \geq s_l(\beta n, n, \frac{1}{n})$, where

$$s_l(\beta n, n, \frac{1}{n}) = (1 - n^{-1})^{n-1} \beta n \log(1 + \frac{\check{P}\bar{F}_H^{-1}(\frac{1}{n})}{n^{-1}(\beta n - 1) + 1}).$$

It can be seen that,

$$\lim_{n\to\infty} \frac{s_l(\beta n, n, \frac{1}{n})}{\frac{\beta n}{e} \log\left(1 + \check{P}\bar{F}_H^{-1}(\frac{1}{n})\right)} = 1. \tag{21}$$

Therefore, (19) is satisfied.

To upper bound $s^*(\beta n, n)$ as in the proof of Proposition 2 consider a new system with k parallel channels, where each user has a peak power constraint of $\check{P}$ for each channel. Denote the optimal throughput of this new system by $s_p(k, n)$. Clearly, $s_p(k, n) \geq s^*(k, n)$ and $s_p(k, n) = ks^*(1, n)$. Again, referring to [9], it is shown that if $f_H(h)$ satisfies condition (*b*) in Lemma 5, then $s^*(1, n) \precsim \frac{1}{e} \log(1 + \check{P}\bar{F}_H^{-1}(\frac{1}{n}))$. Therefore, (20) is also true.

With the same argument as proof of Proposition 2, we also have $s(\beta n, n, \frac{1}{n}) \precsim \frac{1}{e} s_{ct}(\beta n, n)$. ■

REFERENCES

[1] D. Tse and S. Hanly, Multi-Access Fading Channels: Part I: Polymatroid Structure, Optimal Resource Allocation and Throughput Capacities, *IEEE Trans. on Information Theory,* Vol. 44 (November 1998) pp. 2796-2815.

[2] R. Knopp and P. A. Humblet, Information capacity and power control in single-cell multiuser communications, in *Proceedings of IEEE ICC '95,* Seattle, WA (June 1995).

[3] R. Cheng and S. Verdu, Gaussian Multiaccess Channels with ISI: Capacity Region and Multiuser Water-Filling, *IEEE Transactions on Information Theory*, Vol. 39 (May 1993) pp. 772-785.

[4] W. Yu, W. Rhee, S. Boyd and J. Cioffi, Iterative Water-filling for Gaussian Vector Multiple Access Channels, in *Proceedings of IEEE International Symposium on Information Theory* (2001).

[5] W. Yu and J. Cioffi, FDMA Capacity of Gaussian Multiple Access Channels with ISI, *IEEE Trans. on Communications,* vol. 50 (January 2002) pp. 102-111.

[6] C. Y. Wong, R. S. Cheng, K. B. Letaief, and R. D. Murch, Multiuser OFDM with adaptive subcarrier, bit, and power allocation, *IEEE Journal on Selected Areas in Communications*, Vol. 17 (October 1999) pp. 1747–1758.

[7] M. Ergen, S. Coleri, and P. Varaiya, QoS Aware Adaptive Resource Allocation Techniques for Fair Scheduling in OFDMA Based Broadband Wireless Access Systems, *IEEE Trans. on Broadcasting*, Vol. 49 (Dec. 2003) pp. 362-370.

[8] X. Qin and R. Berry, Exploiting Multiuser Diversity for Medium Access Control in Wireless Networks, in *Proceedings of 2003 IEEE INFOCOM*, San Francisco, CA (March 2003), pp. 1084-1094.

[9] X. Qin and R. Berry, Distributed Approaches for Exploiting Multiuser Diversity in Wireless Networks, *IEEE Transactions on Information Theory*, Vol. 52 (February 2006) pp. 392-413.

[10] X. Qin and R. Berry, Opportunistic Splitting Algorithms for Wireless Networks, in *Proc. of 2004 IEEE INFOCOM*, Hong Kong, PR China (March 2004) pp. 1662-1672.

[11] Y. Sun and M. Honig, Asymptotic Capacity of Multi-Carrier Transmission over a Fading Channel with Feedback, in *Proceedings of IEEE International Symposium on Information Theory* (2003).

[12] L. Hoo, B. Halder, J. Tellado and J. Cioffi, Multiuser Transmit Optimization for Multicarrier Broadcast Channels: Asymptotic FDMA Capacity Region and Algorithms, *IEEE Transaction on Communications*, Vol. 52 (June 2004) pp. 922-930.

PLACE
PHOTO
HERE

Xiangping Qin received the B.S. and M.S. degrees in Electrical Engineering from Tsinghua University, China in 1998 and 2000 respectively, and the PhD degree in Electrical Engineering from Northwestern University in 2005. She is currently a senior engineer at Samsung Information Systems America. In 2005/2006, She was a postdoctoral associate in the Department of Electrical and Computer Engineering at Boston University. In 2004, she was an intern on the technical staff of Intel Cooperate Technology Laboratory, Oregon. Her primary research interests include wireless communication and data networks. She is the recipient of a Walter P. Murphy Fellowship for the 2000/2001 academic year from the ECE Department at Northwestern University.

PLACE
PHOTO
HERE

Randall A. Berry received the B.S. degree in Electrical Engineering from the University of Missouri-Rolla in 1993 and the M.S. and PhD degrees in Electrical Engineering and Computer Science from the Massachusetts Institute of Technology in 1996 and 2000, respectively. In September 2000, he joined the faculty of Northwestern University, where he is currently an Associate Professor in the Department of Electrical Engineering and Computer Science. In 1998 he was on the technical staff at MIT Lincoln Laboratory in the Advanced Networks Group, where he worked on optical network protocols. His current research interests include wireless communication, data networks and information theory.

Dr. Berry is the recipient of a 2003 NSF CAREER award and the 2001-02 best teacher award from the ECE Department at Northwestern. He is currently serving on the editorial board of *IEEE Transactions on Wireless Communications* and is a guest editor of an upcoming special issue of *IEEE Transactions on Information Theory* on “Relaying and Cooperation in Networks.”

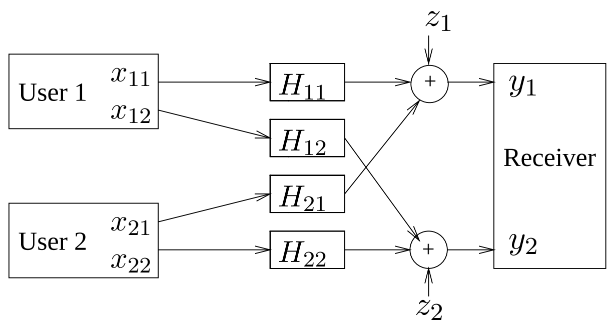

Fig. 1. Two users with $k = 2$ parallel channels.

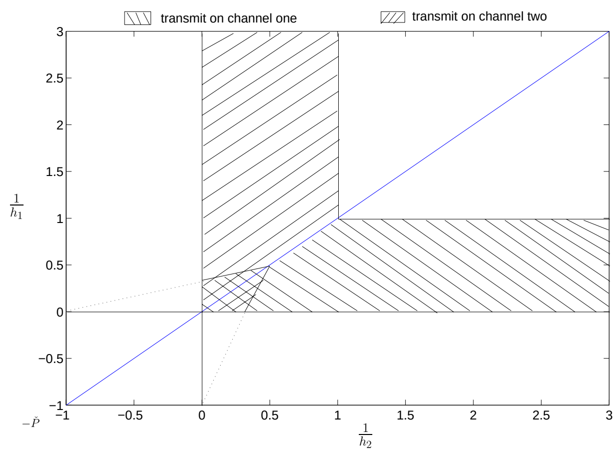

Fig. 2. Optimal power allocation for $k = 2$ parallel channels with multiple users. The double crossed area indicates when the user transmits on both channels. The single crossed area indicates transmission on only one channel.

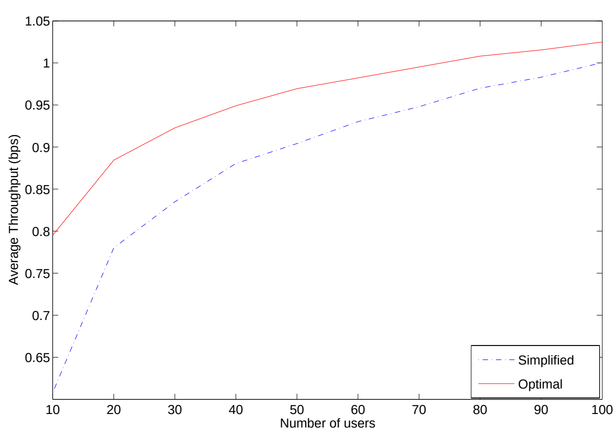

Fig. 3. Average throughput (bps) per channel of the optimal distributed scheme and the simplified distributed scheme as a function of the number of users for $k = 10$ channels.

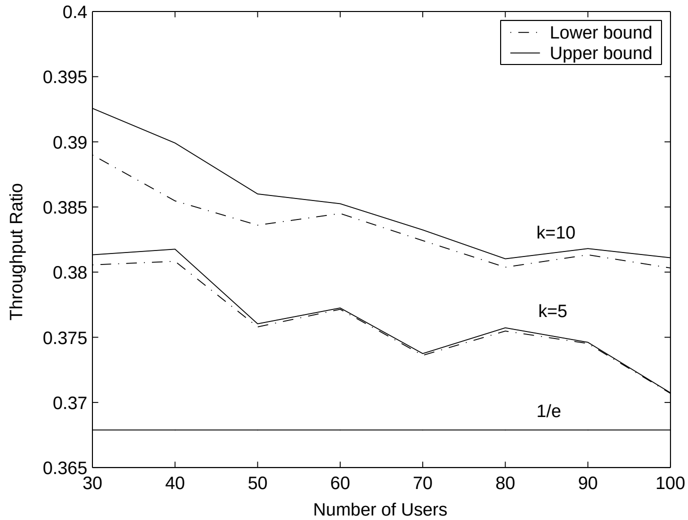

Fig. 4. Lower and upper bounds on the ratio of average throughputs of the optimal distributed scheme to the optimal centralized scheme versus the number of users, for $k = 5$ and 10 channels.

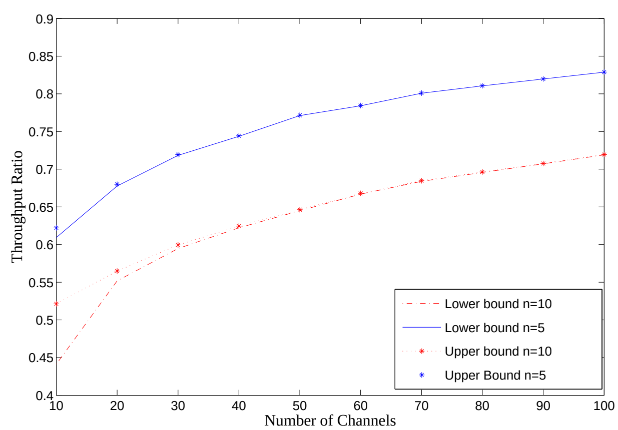

Fig. 5. Lower and upper bounds on the ratio of the average throughputs of the optimal distributed scheme to the optimal centralized scheme versus the number of channels, for $n = 5$ and 10 users.