Power Control and Multiuser Diversity in Multiple Access Channels with Two Time-Scale Fading

Alex S. C. Leong, Jamie S. Evans and Subhrakanti Dey

ARC Special Research Centre for Ultra-Broadband Information Networks (CUBIN)

Department of Electrical and Electronic Engineering

University of Melbourne, Vic. 3010, Australia

{a.leong, j.evans, s.dey}@ee.unimelb.edu.au [*][†]

Abstract

We derive the optimal power control strategy to maximize the sum rate of a multiple access channel with two time-scale fading, where transmitters have access to each of the other users' 'slow' fading information and the statistics of the 'fast' fading, but no knowledge of the instantaneous fast fading states. Assuming identical fast fading distributions for all users, it is found that the optimal strategy is to let at most one user transmit, with the user transmitting the one with the 'best' slow fading conditions. An example with users undergoing lognormal shadowing and Rayleigh fast fading is considered, and capacity comparisons made. Simple sub-optimal power control schemes which provide close to optimal performance in certain favorable channel conditions are also proposed and analysed.

1 Introduction

The time-varying fading channel provides considerable challenges to communications engineers, who must try to meet certain often conflicting performance requirements such as low delay and high data rates. One commonly used power control method to combat fading is channel inversion, which aims to keep the power of the received signal at a constant level, and is suitable for delay sensitive applications such as voice. However, for other applications such as data where it is desirable to achieve high transmission rates, channel inversion could be far from optimal. In such situations, considering the long term average throughput or ergodic capacity may be more appropriate.

The ergodic capacity of a single-user fading channel with perfect channel side information at the transmitter and receiver was considered in [9]. The multi-user sum capacity

was considered in [10], and a characterisation of the capacity region for the multi-access fading channel was given in [12]. Alternative situations including various different notions of capacity studied by other authors are described in [4].

Much of the previous work implicitly assumes however that the fading is of a single time-scale, whether it be fast or slow. In reality, fading is often due to multiple components, some of which could vary on widely different time-scales [13]. For example, we might have a slowly time-varying component due to shadowing effects, together with a much faster component due to multipath. In [6], power control schemes to optimize the ergodic capacity, outage capacity and capacity with outage constraints were derived, for a single user channel with a two time-scale fading process. Here the receiver is assumed to be able to track both the fast and slow fading components, but the transmitter makes power allocation decisions based on the slow fading component only (plus some knowledge of the fast fading statistics).

In this paper we extend the ergodic capacity results of [6] to a multi-access two time-scale fading channel. Assuming the fast fading distributions to be identical for all users, we find in Section 2 that the optimal strategy is, as in [10], to let at most one user transmit, but now the user transmitting is the one with the best *slow* fading conditions. We thus still have a form of multiuser diversity, but now we can only take advantage of channel fluctuations occurring at a slower time scale. In Section 3 we look at a specific example, with lognormal shadowing and Rayleigh fading for the slow and fast fading components respectively, in both the symmetric and asymmetric cases. In the symmetric case, numerical comparisons are made with the capacities achievable for our scheme, together with schemes with no power adaptation and power adaptation with full channel side information. The power allocation law derived here does not take on a closed form, so in Section 4 we look at two simpler sub-optimal schemes that are suitable for the

1

symmetric users case, and which provides close to optimal performance when the number of users or average received SNR is high. Results relating the optimal and one of the sub-optimal policies are proved in Section 5. Finally, Section 6 presents some asymptotic results for the sum capacity as the number of users goes to infinity.

2 Power control with two time-scale fading

2.1 System model

We will consider the uplink of a single cell multi-user communications system, with frequency-flat fading. The signal received by the base station at time n can be expressed as

$$Y(n) = \sum_{i=1}^{K} \sqrt{G_i(n)F_i(n)} X_i(n) + Z(n).$$

Here K denotes the number of users in the cell and $X_i(n)$ the transmitted information of user i, with $X_i(n)$ and $X_j(n)$ assumed to be independent for all $i \neq j$, and each user subject to an average power constraint $\mathcal{P}$. $\{Z(n)\}$ is a white Gaussian noise process, with zero mean and variance N_0. $G_i(n)$ and $F_i(n)$ will be regarded as the random slow and fast channel fading gains respectively of user i, both assumed to be continuous positive random variables with unbounded support and finite means, and with $F_i(n)$ possibly changing on a much faster time-scale than $G_i(n)$. $\{G_i(n)\}$ and $\{F_i(n)\}$ will be assumed to be jointly stationary and ergodic for the time-scales concerned, with $G_i(n)$ and $F_j(n)$ independent for all i and j, $G_i(n)$ and $G_j(n)$ independent $\forall i \neq j$, and $F_i(n)$ and $F_j(n)$ independent $\forall i \neq j$. In order to derive the power control strategy in section 2.3, we make the additional assumption that the fast fading distributions for the $F_i(n)$'s are identical for all users i. For brevity, the time index n will be suppressed in the ensuing discussions.

2.2 Problem formulation

This paper considers the situation where the base station has perfect knowledge of the instantaneous slow and fast fading gains for all of the users, whilst the transmitting users have perfect knowledge of each of the users' slow fading G_i's, but no knowledge of any of the instantaneous fast fading F_i's. We do however assume that the transmitters know the statistics of the F_i's. Such a situation could occur either because the transmitter is unable to track all the fast components or chooses not to do so, eg. computational reasons. The aim then is to find the power control strategy which will maximize the capacity in an information theoretic sense.

Let g_i and f_i represent realisations of G_i and F_i respectively. Let $\boldsymbol{G} = (G_1, G_2, \ldots, G_K)$, $\boldsymbol{g} = (g_1, g_2, \ldots, g_K)$, $\boldsymbol{F} = (F_1, F_2, \ldots, F_K)$, and $\boldsymbol{f} = (f_1, f_2, \ldots, f_K)$. Denote the power control by $\boldsymbol{\mu}(\boldsymbol{G}) = (\mu_1(\boldsymbol{G}), \mu_2(\boldsymbol{G}), \ldots, \mu_K(\boldsymbol{G}))$, where $\mu_i(\boldsymbol{G})$ is the power control law for user i, a function (assumed to be non-negative) of the slow fades $\boldsymbol{G}$. We constrain the average transmitted power of each user i to be no greater than $\mathcal{P}$, ie. $\mathbb{E}_{\boldsymbol{G}}[\mu_i(\boldsymbol{G})] \leq \mathcal{P}, \forall i$.

We use the notation $\mathbb{E}_{\mathbf{X}}[\bullet]$ to denote the expectation with respect to $\mathbf{X}$, and $\mathbb{E}_{\mathbf{X}|\mathbf{Y}}[\bullet|\mathbf{Y}]$ to denote the conditional expectation with respect to $\mathbf{X}$ given $\mathbf{Y}$. $p_{\mathbf{X}}(\mathbf{x})$ and $P_{\mathbf{X}}(\mathbf{x})$ will be used to represent the pdf and cdf respectively.

In this paper we will only address the problem of maximizing the sum rate, rather than dealing with general capacity regions as in eg. [12]. The following definition of the sum capacity in two time-scale fading will be used.

Definition 1.

$$C_{tts} = \max_{\boldsymbol{\mu}(\boldsymbol{G})} \mathbb{E}_{\boldsymbol{G}} \left[\mathbb{E}_{\boldsymbol{F}|\boldsymbol{G}} \left[\frac{1}{2} \log_2 \left(1 + \frac{\sum_{i=1}^{K} \mu_i(\boldsymbol{G}) G_i F_i}{N_0} \right) \right] \right]$$

with $\mathbb{E}_{\boldsymbol{G}}[\mu_i(\boldsymbol{G})] \leq \mathcal{P}$ and $\mu_i(\boldsymbol{G}) \geq 0$ for $i = 1, \ldots, K$.

Some motivation for Definition 1 is as follows. Suppose the slow fading components for each user remains constant over a block of say N symbols, but varies from block to block. Then the maximum sum rate achievable over each block, what might be termed the 'block-ergodic sum capacity' (a similar notion can be found in [6]), is given by $\mathbb{E}_{\boldsymbol{F}|\boldsymbol{G}} \left[\frac{1}{2} \log_2 \left(1 + \frac{\sum_{i=1}^{K} \mu_i(\boldsymbol{G}) G_i F_i}{N_0} \right) \right]$, which is an expectation over the fast fading $\boldsymbol{F}$, given the slow fading $\boldsymbol{G}$. The expected value of this block-ergodic sum capacity, with the expectation now taken over the slow fading $\boldsymbol{G}$, then gives us Definition 1.

We write more explicitly the equivalent optimization problem:

$$\max_{\boldsymbol{\mu}(\boldsymbol{G})} \mathbb{E}_{\boldsymbol{G}} \left[\mathbb{E}_{\boldsymbol{F}|\boldsymbol{G}} \left[\ln \left(1 + \frac{\sum_{i=1}^{K} \mu_i(\boldsymbol{G}) G_i F_i}{N_0} \right) \right] \right] \quad (1)$$

subject to

$$\mathbb{E}_{\boldsymbol{G}}[\mu_i(\boldsymbol{G})] = \mathcal{P}, i = 1, \ldots, K \quad (2)$$

$$\mu_i(\boldsymbol{G}) \geq 0, i = 1, \ldots, K \quad (3)$$

We have changed the constraint (2) to an equality, but this will not affect the optimal solution since allocating more power will always increase the objective.

2.3 Optimal power control strategies

Here we will derive the solution to (1)-(3). Firstly, the objective can be shown to be a concave function of $\boldsymbol{\mu}$, by

2

using the concavity of the logarithm and noting that taking expectations will retain the concavity (provided certain mild technical assumptions on the distributions are satisfied). We also have

$$
\frac{\partial}{\partial \mu_i} \mathbb{E}_{\boldsymbol{F}|\boldsymbol{G}} \left[\ln \left(1 + \frac{\sum_{i=1}^K \mu_i(\boldsymbol{G}) G_i F_i}{N_0} \right) \right]
$$

$$
= \mathbb{E}_{\boldsymbol{F}|\boldsymbol{G}} \left[\frac{\partial}{\partial \mu_i} \left(\ln \left(1 + \frac{\sum_{i=1}^K \mu_i(\boldsymbol{G}) G_i F_i}{N_0} \right) \right) \right]
$$

$$
= \mathbb{E}_{\boldsymbol{F}|\boldsymbol{G}} \left[\frac{G_i F_i}{N_0 + \sum_{k=1}^K \mu_k(\boldsymbol{G}) G_k F_k} \right]
$$

where taking the partial derivative inside the expectation is justified because (see eg. [3])

$$
\left| \frac{g_i f_i}{N_0 + \sum_{k=1}^K \mu_k(\boldsymbol{g}) g_k f_k} p_{\boldsymbol{F}|\boldsymbol{G}}(\boldsymbol{f}|\boldsymbol{g}) \right| \leq \frac{g_i f_i}{N_0} p_{\boldsymbol{F}|\boldsymbol{G}}(\boldsymbol{f}|\boldsymbol{g})
$$

and $\mathbb{E}_{\boldsymbol{F}|\boldsymbol{G}} \left[\frac{G_i F_i}{N_0} \right] < \infty$ (since $\mathbb{E}[F_i] < \infty$ by assumption).

Introducing the Lagrange multipliers λ_i for each of the constraints (2) and applying the Kuhn-Tucker optimality conditions [11], we obtain for $i = 1, \ldots, K$:

$$
\mathbb{E}_{\boldsymbol{F}|\boldsymbol{G}} \left[\frac{G_i F_i}{N_0 + \sum_{k=1}^K \mu_k(\boldsymbol{G}) G_k F_k} \right] \begin{cases} = \lambda_i, & \mu_i(\boldsymbol{G}) > 0 \\ \leq \lambda_i, & \mu_i(\boldsymbol{G}) = 0 \end{cases}
$$
(4)

We now have the following results:

Lemma 1. *At most one user should transmit during any slow fading realisation $\boldsymbol{g}(n) \equiv \boldsymbol{g}$.*

Proof. For a given slow fading realisation $\boldsymbol{g}$, suppose that user i is transmitting, ie. $\mu_i(\boldsymbol{g}) > 0$. Then from (4), we have $\mathbb{E}_{\boldsymbol{F}|\boldsymbol{g}} \left[\frac{g_i F_i}{N_0 + \sum_{k=1}^K \mu_k(\boldsymbol{g}) g_k F_k} \right] = \lambda_i$ or

$$
\frac{g_i}{\lambda_i} \mathbb{E}_{\boldsymbol{F}|\boldsymbol{g}} \left[\frac{F_i}{N_0 + \sum_{k=1}^K \mu_k(\boldsymbol{g}) g_k F_k} \right] = 1.
$$

Consider another user j, where $j \neq i$. Since the G_i's are continuous distributions, we will have $\frac{g_j}{\lambda_j} \neq \frac{g_i}{\lambda_i}$ almost surely, and so

$$
\frac{g_j}{\lambda_j} \mathbb{E}_{\boldsymbol{F}|\boldsymbol{g}} \left[\frac{F_j}{N_0 + \sum_{k=1}^K \mu_k(\boldsymbol{g}) g_k F_k} \right]
$$

$$
= \frac{g_j}{\lambda_j} \mathbb{E}_{\boldsymbol{F}|\boldsymbol{g}} \left[\frac{F_i}{N_0 + \sum_{k=1}^K \mu_k(\boldsymbol{g}) g_k F_k} \right]
$$

$$
\neq \frac{g_i}{\lambda_i} \mathbb{E}_{\boldsymbol{F}|\boldsymbol{g}} \left[\frac{F_i}{N_0 + \sum_{k=1}^K \mu_k(\boldsymbol{g}) g_k F_k} \right]
$$

$$
= 1
$$

where the first equality comes from the F_i's being identically distributed. From the optimality conditions (4), this implies that user j cannot also be transmitting if user i is transmitting. $\square$

Lemma 2. *Necessary and sufficient conditions for user i to transmit are that $g_i > \frac{N_0}{\mathbb{E}[F_i]} \lambda_i$ and $\frac{g_i}{\lambda_i} > \frac{g_j}{\lambda_j}$ for all $j \neq i$.*

Proof. We prove sufficiency first. Given $\frac{g_i}{\lambda_i} > \frac{g_j}{\lambda_j}, \forall j \neq i$, it is straightforward to show by a contradiction argument that $\mu_j(\boldsymbol{g}) = 0, \forall j \neq i$. The optimality conditions (4) for user i therefore become

$$
\mathbb{E}_{F_i|\boldsymbol{g}} \left[\frac{g_i F_i}{N_0 + \mu_i(\boldsymbol{g}) g_i F_i} \right] \begin{cases} = \lambda_i, & \mu_i(\boldsymbol{g}) > 0 \\ \leq \lambda_i, & \mu_i(\boldsymbol{g}) = 0 \end{cases}
$$
(5)

Assume that $\mu_i(\boldsymbol{g}) = 0$. Then from (5) we have $\frac{g_i}{N_0} \mathbb{E}[F_i] \leq \lambda_i$ or $g_i \leq \frac{N_0}{\mathbb{E}[F_i]} \lambda_i$. Hence $g_i > \frac{N_0}{\mathbb{E}[F_i]} \lambda_i \Rightarrow \mu_i(\boldsymbol{g}) > 0$ (when $\frac{g_i}{\lambda_i} > \frac{g_j}{\lambda_j}, \forall j \neq i$), which proves sufficiency.

For necessity, let $\mu_i(\boldsymbol{g}) > 0$. We have from (4) and the proof of Lemma 1 that

$$
\frac{g_j}{\lambda_j} \mathbb{E}_{\boldsymbol{F}|\boldsymbol{g}} \left[\frac{F_i}{N_0 + \sum_{k=1}^K \mu_k(\boldsymbol{g}) g_k F_k} \right]
$$

$$
< 1
$$

$$
= \frac{g_i}{\lambda_i} \mathbb{E}_{\boldsymbol{F}|\boldsymbol{g}} \left[\frac{F_i}{N_0 + \sum_{k=1}^K \mu_k(\boldsymbol{g}) g_k F_k} \right]
$$

or $\frac{g_j}{\lambda_j} < \frac{g_i}{\lambda_i}, \forall j \neq i$. Using (5) again, we obtain

$$
\lambda_i = \mathbb{E}_{F_i|\boldsymbol{g}} \left[\frac{g_i F_i}{N_0 + \mu_i(\boldsymbol{g}) g_i F_i} \right] < \mathbb{E}_{F_i|\boldsymbol{g}} \left[\frac{g_i F_i}{N_0} \right] = \frac{g_i}{N_0} \mathbb{E}[F_i]
$$

or $g_i > \frac{N_0}{\mathbb{E}[F_i]} \lambda_i$. $\square$

Now following [5], define

$$
e_{\boldsymbol{g}}(\mu_i) = \mathbb{E}_{F_i|\boldsymbol{g}} \left[\frac{g_i F_i}{N_0 + \mu_i(\boldsymbol{g}) g_i F_i} \right]
$$
(6)

which equals λ_i when $\mu_i(\boldsymbol{g}) > 0$. It is easily seen that for a given $\boldsymbol{g}$, $e_{\boldsymbol{g}}(\mu_i)$ is monotonically decreasing in μ_i, and hence that $\mu_i(\boldsymbol{g}) = e_{\boldsymbol{g}}^{-1}(\lambda_i)$ exists. Combining all the above results, we have thus shown the following.

Theorem 3. *Under the assumptions of Section 2.1, the optimal power control law for user i, for the problem (1)-(3), is of the form*

$$
\mu_i(\boldsymbol{g}) = \begin{cases} e_{\boldsymbol{g}}^{-1}(\lambda_i), & \frac{g_i}{\lambda_i} > \frac{g_j}{\lambda_j}, g_i > \frac{N_0}{\mathbb{E}[F_i]} \lambda_i, \forall j \neq i \\ 0, & otherwise \end{cases}
$$
(7)

where $e_{\boldsymbol{g}}^{-1}(\lambda_i)$ is determined from (6) and λ_i is chosen to satisfy the power constraint (2).

3

This power control strategy, which states that at most one user should be transmitting at any time, with the user who is transmitting the one with the best slow fading conditions (appropriately scaled by its corresponding Lagrange multiplier) and in addition is above a certain threshold, are analogous to the results of [10], which addressed the case where perfect side information is available at both the transmitters and receiver. As in [10], we also have a form of multiuser diversity, except that now we can only take advantage of channel flucuations which occur at a slower time scale. Whether there is a large penalty to be paid will be considered next.

3 Lognormal-Rayleigh fading example

In this section we will look at a specific example, where the fast components undergo Rayleigh fading (so that the F_i's are exponentially distributed), and slow components undergo lognormal shadowing. We will take $\mathcal{P} = 1$, $N_0 = 1$, $F_i \sim \exp(f^*)$, $G_i \sim \text{lognormal}(\mu_{\text{dB}i}, \sigma_{\text{dB}i})$, $\forall i$. The pdfs (p's) and cdfs (P's) are [8]:

$$p_G(g) = \frac{10}{\ln 10 \sqrt{2\pi}\sigma_{\text{dB}}g} \exp\left[\frac{-(10\log_{10} g - \mu_{\text{dB}})^2}{2\sigma_{\text{dB}}^2}\right], g > 0$$

$$P_G(g) = \frac{1}{2} + \frac{1}{2}\text{erf}\left[\frac{10\ln g - \mu_{\text{dB}}\ln 10}{\sqrt{2}\sigma_{\text{dB}}\ln 10}\right], g > 0$$

$$p_F(f) = \frac{1}{f^*}\exp\left(-\frac{f}{f^*}\right), f > 0$$

$$P_F(f) = 1 - \exp\left(-\frac{f}{f^*}\right), f > 0.$$

We also choose the mean $f^* = 1$ (ie. $\mathbb{E}[F_i] = 1, \forall i$) and $\sigma_{\text{dB}i} = 8$ dB, $\forall i$, and look at the effects on the sum capacity for different values of the lognormal parameter $\mu_{\text{dB}i}$ (the dB mean, which in effect controls the average received SNR) and the number of users K.

Let us first derive a more explicit expression for calculating $\mu_i(\boldsymbol{g}) = e_{\boldsymbol{g}}^{-1}(\lambda_i)$. Note that $p_{F_i|\boldsymbol{g}}(f_i|\boldsymbol{g}) = p_{F_i}(f_i)$ since we assumed F_i and G_j to be independent for all i and j. Conditional on $\mu_i(\boldsymbol{g}) > 0$, we have from (6) that

$$\lambda_i = \mathbb{E}_{F_i|\boldsymbol{g}}\left[\frac{g_i F_i}{1 + \mu_i(\boldsymbol{g})g_i F_i}\right]$$
$$= \int_0^\infty \frac{g_i f_i}{1 + \mu_i(\boldsymbol{g})g_i f_i}\exp(-f_i)df_i$$

or

$$\lambda_i = g_i\left[\frac{1}{\mu_i(\boldsymbol{g})g_i} - \frac{1}{(\mu_i(\boldsymbol{g})g_i)^2}\exp\left(\frac{1}{\mu_i(\boldsymbol{g})g_i}\right)E_1\left(\frac{1}{\mu_i(\boldsymbol{g})g_i}\right)\right] \tag{8}$$

for $i = 1, \ldots, K$, where $E_1(x) = \int_x^\infty \frac{e^{-t}}{t}dt$ is the exponential integral of the first order [1]. Properties relating to Equation (8), such as uniqueness and limiting results, can be found in [6] (the Lagrange multipliers are defined slightly

differently however). The constraint equations (2) take the form $\int_{\lambda_i}^\infty \mu_i(\boldsymbol{g})p_{G_i}(g_i)\prod_{j\neq i}\Pr\left(\frac{G_j}{\lambda_j} < \frac{g_i}{\lambda_i}\right)dg_i = 1$ or

$$\int_{\lambda_i}^\infty \mu_i(\boldsymbol{g})p_{G_i}(g_i)\prod_{j\neq i}P_{G_j}\left(\frac{\lambda_j}{\lambda_i}g_i\right)dg_i = 1, i = 1, \ldots, K \tag{9}$$

The sum capacity (Definition 1), can be written more explicitly using the power control law (7) as the expression

$$C_{tts} = \sum_{i=1}^K \frac{1}{2\ln 2}\int_{\lambda_i}^\infty \exp\left(\frac{1}{\mu_i(\boldsymbol{g})g_i}\right)E_1\left(\frac{1}{\mu_i(\boldsymbol{g})g_i}\right)$$
$$\times p_{G_i}(g_i)\prod_{j\neq i}P_{G_j}\left(\frac{\lambda_j}{\lambda_i}g_i\right)dg_i \tag{10}$$

3.1 Symmetric case

In the symmetric users case, we assume that all the slow fading distributions G_i's are identically distributed. By symmetry, this implies that all the Lagrange multipliers are equal, ie. $\lambda_i = \lambda, \forall i$. (8) reduces to the single equation

$$\lambda = g_i\left[\frac{1}{\mu_i(\boldsymbol{g})g_i} - \frac{1}{(\mu_i(\boldsymbol{g})g_i)^2}\exp\left(\frac{1}{\mu_i(\boldsymbol{g})g_i}\right)E_1\left(\frac{1}{\mu_i(\boldsymbol{g})g_i}\right)\right] \tag{11}$$

The constraint equations (9) also simplify to the single equation

$$\int_\lambda^\infty \mu_i(\boldsymbol{g})p_G(g_i)P_G(g_i)^{K-1}dg_i = 1. \tag{12}$$

Equations (11) and (12) can be used to determine λ and $\mu_i(\boldsymbol{g})$, either numerically or via Monte Carlo methods.

We will compare the sum capacities for the following three schemes.

1. Power control using slow fading information only
The sum capacity (10) simplies to

$$C_{tts} = \frac{K}{2\ln 2}\int_\lambda^\infty \exp\left(\frac{1}{\mu_i(\boldsymbol{g})g_i}\right)E_1\left(\frac{1}{\mu_i(\boldsymbol{g})g_i}\right)$$
$$\times p_G(g_i)P_G(g_i)^{K-1}dg_i$$

where λ and $\mu_i(\boldsymbol{g})$ are determined using (11) and (12).
2. No power control
The sum capacity with no power control, C_{npc}, can be expressed as

$$C_{npc} = \mathbb{E}_{\boldsymbol{G}}\left[\mathbb{E}_{\boldsymbol{F}|\boldsymbol{G}}\left[\frac{1}{2}\log_2\left(1 + \sum_{i=1}^K G_i F_i\right)\right]\right]$$

We are unable to simplify this expression, and direct numerical integration seems to be very challenging, even for $K = 2$, thus the results presented here were generated by Monte Carlo methods.
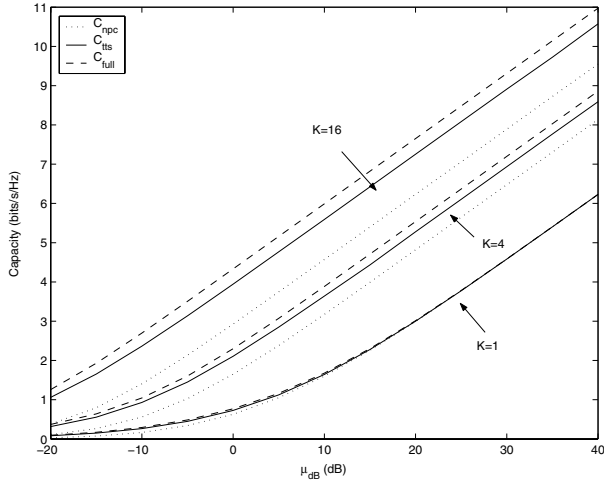
4

Figure 1. Capacity comparisons for $K = 1, 4$ and 16 users

Figure 2. Sum capacity and rates of the users, for $K = 2$ and $\mu_{\mathrm{dB1}} = 0$ dB
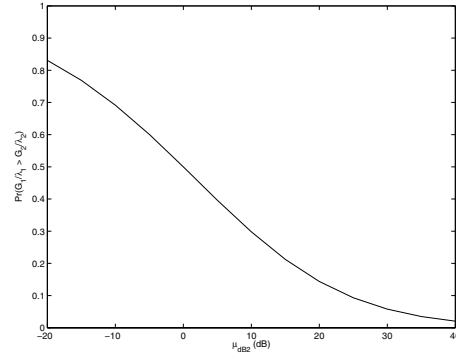
Figure 3. Probability that user 1 is better than user 2, for $K = 2$ and $\mu_{\mathrm{dB1}} = 0$ dB

3. Power control with full channel side information
Let $H_i = G_i F_i$, with h_i being a realisation of H_i. In the case where the fast and slow fading are known to both the transmitters and receiver, the sum capacity is

$$C_{full} = \frac{K}{2 \ln 2} \int_{\eta}^{\infty} \ln \left(\frac{h_i}{\eta} \right) p_H(h_i) P_H(h_i)^{K-1} dh_i$$

where η is determined from the constraint equation

$$\int_{\eta}^{\infty} \left(\frac{1}{\eta} - \frac{1}{h_i} \right) p_H(h_i) P_H(h_i)^{K-1} dh_i = 1.$$

Results for $K = 1, 4$ and 16 users, and various values of μ_{dB}, are shown in Figure 1. Monte Carlo methods were used to used to evaluate C_{npc}, while C_{tts} and C_{full} were numerically evaluated using Mathematica. It can be seen that $C_{npc} \leq C_{tts} \leq C_{full}$, which intuitively agrees with the notion that having more channel side information at the transmitters allows us to achieve better throughputs. For $K = 1$, the difference between all three schemes is negligible at large values of μ_{dB}. For $K = 4$ and $K = 16$, there is a noticeable gap between C_{tts} and C_{full} at all values considered. However, when compared to the case with no power control, the capacity gains of the scheme based on slow fading only are still substantial, particularly as the number of users is increased.

3.2 Asymmetric case

We now consider the asymmetric case, where the slow fading distributions may be different between users. Using the K equations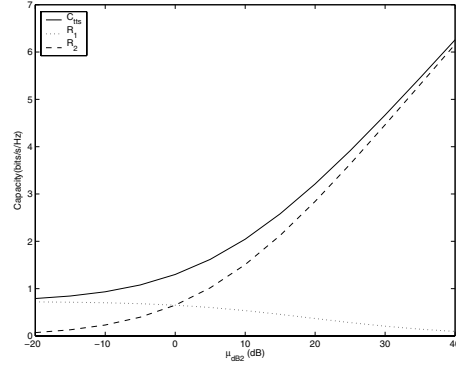 in (8) and the K equations in (9), we can numerically determine the values λ_i and hence $\mu_i(\boldsymbol{g})$. Results will be presented for the two user case. We will fix $\mu_{\mathrm{dB1}} = 0$ dB and vary μ_{dB2}. Figure 2 plots the sum capacity (10), together with the rates that each of the two users contributes to the overall sum, for different values of μ_{dB2}. For $\mu_{\mathrm{dB2}} < \mu_{\mathrm{dB1}} = 0$ dB, user 1 contributes more to the sum capacity than user 2. As μ_{dB2} increases, the overall sum capacity increases. Hoewever, the contribution of user 1 decreases, while the contribution of user 2 increases, and for $\mu_{\mathrm{dB2}} >> \mu_{\mathrm{dB1}}$, most of the contribution is due to user 2.

We can also evaluate the probability with which user i is the 'best', ie. $\mathrm{Pr}(\frac{G_i}{\lambda_i} > \frac{G_j}{\lambda_j}, \forall j \neq i)$, as

$$\prod_{j \neq i} \mathrm{Pr} \left(\frac{G_j}{\lambda_j} < \frac{G_i}{\lambda_i} \right) = \int_0^{\infty} p_{G_i}(g_i) \prod_{j \neq i} P_{G_j} \left(\frac{\lambda_j}{\lambda_i} g_i \right) dg_i \tag{13}$$

In Figure 3 we plot the probability that user 1 is 'better' than user 2, for fixed $\mu_{\mathrm{dB1}} = 0$ dB. For $\mu_{\mathrm{dB2}} = \mu_{\mathrm{dB1}} = 0$ dB, each user will transmit approximately half of the time. But

5

when one user is much stronger than another, eg. $\mu_{\mathrm{dB}2} \gg \mu_{\mathrm{dB}1}$, then the stronger user (user 2) will end up transmitting most of the time, while the weaker user (user 1) will not transmit very often. Fairness is therefore an issue in the asymmetric case, and the sum rate may not necessarily the most appropriate quantity to optimize, c.f. [10]. Also see [14] for other related practical issues, such as the variable delay between successive transmissions. In the following sections, we will mainly concentrate on the symmetric case.

4 Sub-optimal power control schemes

In the case of perfect channel side information at both the transmitter and receiver [9]-[10], the power allocation law takes on the relatively simple form $\mu(h) = \frac{1}{\lambda} - \frac{1}{h}$. However, in the case of the transmitters having only slow fading information, the power allocation $e_{\boldsymbol{g}}^{-1}(\lambda)$ defined by (6) does not take on a closed form in general and must be found numerically, eg. the transcendental equation (8). This thus motivates the idea of using simpler power control strategies which hopefully achieves close to the optimal capacities under 'good' channel conditions, whilst still satisfying the average power constraints. Such favorable conditions might include cases where there are either a large number of users in the system (to take advantage of multiuser diversity effects), or if the average received SNR is high. In this section, we consider two simple sub-optimal schemes which particularly suitable for the symmetric users case.

4.1 Statement of the sub-optimal schemes

Suppose there are K symmetric users in the system and an average power constraint $\mathcal{P}$ for each user. Then we have the following two sub-optimal power control schemes.

1. Assume that the fast fading takes on its mean and solve the following problem instead

$$\max_{\boldsymbol{\mu}(\boldsymbol{G})} \mathbb{E}_{\boldsymbol{G}} \left[\ln \left(1 + \frac{\sum_{i=1}^{K} \mu_i(\boldsymbol{G}) G_i \mathbb{E}[F_i]}{N_0} \right) \right]$$

subject to $\mathbb{E}_{\boldsymbol{G}}[\mu_i(\boldsymbol{G})] = \mathcal{P}, i = 1, \ldots, K$ and $\mu_i(\boldsymbol{G}) \geq 0, i = 1, \ldots, K$.

The solution to this problem is given in [10]. We then use the obtained $\boldsymbol{\mu}(\boldsymbol{G})$ to evaluate the sum rate $\mathbb{E}_{\boldsymbol{G}} \left[\mathbb{E}_{\boldsymbol{F}|\boldsymbol{G}} \left[\frac{1}{2} \log_2 \left(1 + \frac{\sum_{i=1}^{K} \mu_i(\boldsymbol{G}) G_i F_i}{N_0} \right) \right] \right]$.

2. At a particular instant n, only the user with the best slow fading at that time is allowed to transmit, with transmission at the constant power $K\mathcal{P}$. This scheme will satisfy the average power constraint $\mathcal{P}$, since a given user will have the best slow fading conditions out of all users $1/K$-th of the time on average. Notice also that there is no need to determine what the Lagrange multipliers are.

Both policies can be extended to the asymmetric case. For the first scheme, [10] also provides the solution. For the second scheme, the best user will again transmit at constant power, though the powers for each user will be different, since the proportion of time that each of the users is 'best' (13) are different. To determine what the different powers should be, one would need to know the Lagrange multipliers λ_i, which in turn requires us to solve eg. (8). So this scheme is perhaps not as simple/useful in the asymmetric case.

4.2 Simulation results

The sum rate for the Rayleigh-Lognormal example of Section 3, using the above sub-optimal schemes, may be written as

$$R_{subopt1} = \frac{K}{2 \ln 2} \int_0^\infty \exp \left(\frac{1}{\mu_i(\boldsymbol{g}) g_i} \right) E_1 \left(\frac{1}{\mu_i(\boldsymbol{g}) g_i} \right) \times p_G(g_i) P_G(g_i)^{K-1} dg_i$$

with

$$\mu_i(\boldsymbol{g}) = \begin{cases} \frac{1}{\eta} - \frac{1}{g_i} & , \quad g_i > g_j, g_i > \eta, \forall j \neq i \\ 0 & , \quad \text{otherwise} \end{cases}$$

$$\int_\eta^\infty \left(\frac{1}{\eta} - \frac{1}{g_i} \right) p_G(g_i) P_G(g_i)^{K-1} dg_i = 1$$

and

$$R_{subopt2} = \frac{K}{2 \ln 2} \int_0^\infty \exp \left(\frac{1}{K g_i} \right) E_1 \left(\frac{1}{K g_i} \right) \times p_G(g_i) P_G(g_i)^{K-1} dg_i.$$

Table 1 lists numerical figures of the capacity C_{tts} and sub-optimal sum rates for fixed $\mu_{\mathrm{dB}} = 0$ dB and various values of K. Table 2 lists numerical figures for fixed $K = 2$ and various values of μ_{dB}. We can see that C_{tts}, $R_{subopt1}$ and $R_{subopt2}$ are very close to each other for the values considered, and in the cases where the number of users K or the dB mean μ_{dB} gets large, the difference becomes extremely small (smaller than 10^{-5} in magnitude). It seems that unless the channel conditions are very unfavorable, eg. very low average received SNRs, the proposed sub-optimal power control policies will give us achievable rates which are close to optimal. The first sub-optimal scheme also seems to perform better than the second, even though the second scheme is probably simpler. In the following section we will prove that $C_{tts} \to R_{subopt2}$.

5 Some theoretical results

In this section, we prove a number of results for the symmetric case relating to the optimal power allocation scheme

6

Table 1. Sum rates for fixed $\mu_{\mathrm{dB}} = 0$ dB

K	C_{tts}	$R_{subopt1}$	$R_{subopt2}$
2	1.29957	1.29843	1.25836
4	2.10578	2.10476	2.09873
8	3.02345	3.02318	3.02273
16	3.94552	3.94548	3.94545
32	4.83875	4.83874	4.83874
64	5.69960	5.69960	5.69960

Table 2. Sum rates for fixed $K = 2$ users

μ_{dB}(dB)	C_{tts}	$R_{subopt1}$	$R_{subopt2}$
-10	0.506934	0.506378	0.417615
0	1.29957	1.29843	1.25836
10	2.60070	2.59978	2.59484
20	4.17340	4.17326	4.17307
30	5.81884	5.81884	5.81883
40	7.47751	7.47751	7.47751

derived in Section 2.3, which attempts to provide some insight into why the behaviour of the optimal power allocation scheme tends towards that of the second sub-optimal policy of Section 4. By doing so, we will work towards the main convergence theorem, Theorem 9.

Let $G_m \equiv G_{m,K} = \max(G_1, \ldots, G_K)$ be the random variable defined to be the maximum slow fading of the K users. g_m will denote a realisation of G_m. Then let $\mu_i(\boldsymbol{g}) \equiv \mu(g_m) > 0$ if $g_i = \max(g_1, \ldots, g_K)$. Firstly we have:

Lemma 4. $\mu(g_m)$ *is a strictly increasing function of g_m*

Proof. From (6)

$$
\begin{aligned}
\lambda &= \mathbb{E}_{F|g_m}\left[\frac{g_m F}{N_0 + \mu(g_m)g_m F}\right] \\
&= \int_0^\infty \frac{g_m f}{N_0 + \mu(g_m)g_m f}p(f)df \\
&= \int_0^\infty \frac{1}{\frac{N_0}{g_m f} + \mu(g_m)}p(f)df \quad (14)
\end{aligned}
$$

which must hold for all $g_m > \frac{N_0}{\mathbb{E}[F]}\lambda$. Consider the expression

$$
\frac{1}{\frac{N_0}{g_m f} + \mu(g_m)}p(f) \quad (15)
$$

Suppose g_m strictly increases. If $\mu(g_m)$ remains constant, then (15) strictly increases (as a function of f), contradicting (14). If $\mu(g_m)$ strictly decreases, then (15) also strictly increases, again contradicting (14). So the only possibility when g_m is strictly increasing is for $\mu(g_m)$ to also strictly increase. $\square$

Lemma 5. $\lim_{g_m \to \infty} \mu(g_m) = \frac{1}{\lambda}$

Proof. A more direct proof in the case of Rayleigh fast fading may be found in [6]. The general proof uses the following, which is a slight variation of Corollary 5.7 in [3].

Proposition 1. *Let f be a function defined on $X \times [c, \infty)$ to $\mathbb{R}$ such that $x \to f(x,t)$ is measurable for each $t \in [c, \infty)$. Suppose that $\lim_{t \to \infty} f(x,t) = f(x)$ for each $x \in X$, and there exists an integrable function g on X such that $|f(x,t)| \leq g(x)m, \forall t \in [c, \infty)$. Then*

$$
\int f(x)d\mu(x) = \lim_{t \to \infty} \int f(x,t)d\mu(x).
$$

From (6), $\lambda = \int_0^\infty \frac{g_m f}{N_0 + \mu(g_m)g_m f}p(f)df$ must hold for all $g_m > \frac{N_0}{\mathbb{E}[F]}\lambda$, with $\lambda > 0$. Let

$$
a(f, g_m) = \frac{g_m f}{N_0 + \mu(g_m)g_m f}p(f), g_m \in [c, \infty)
$$

with c chosen such that $c > \frac{N_0}{\mathbb{E}[F]}\lambda$, which implies $\mu(c) > 0$.

First we show that $\mu(g_m)$ is bounded for all g_m. For a proof by contradiction, assume that $\mu(g_m) \to \infty$ as $g_m \to \infty$. Then

$$
\lim_{g_m \to \infty} a(f, g_m) = 0
$$

Moreover,

$$
|a(f, g_m)| \leq \frac{1}{\mu(g_m)}p(f) \leq \frac{1}{\mu(c)}p(f)
$$

for all $g_m \geq c$ (by Lemma 4), and $\frac{1}{\mu(c)}p(f)$ is integrable. We can then apply Proposition 1, to obtain

$$
\begin{aligned}
\lambda &= \lim_{g_m \to \infty} \int_0^\infty \frac{g_m f}{N_0 + \mu(g_m)g_m f}p(f)df \\
&= \int_0^\infty \lim_{g_m \to \infty} \frac{g_m f}{N_0 + \mu(g_m)g_m f}p(f)df \\
&= 0
\end{aligned}
$$

which contradicts $\lambda > 0$. Thus $\mu(g_m)$ is bounded, and $\lim_{g_m \to \infty} \mu(g_m)$ exists by Lemma 4.

To find $\lim_{g_m \to \infty} \mu(g_m)$, we use a similar argument to the above, to obtain $\lim_{g_m \to \infty} a(f, g_m) = \left(\lim_{g_m \to \infty} \frac{1}{\mu(g_m)}\right)p(f)$ and $\lambda = \lim_{g_m \to \infty} \frac{1}{\mu(g_m)}$ $\square$

Next we have a number of results which are applicable when the number of users $K \to \infty$. Statements and proofs for the high average SNR case are similar, see the end of this section for some comments.

Lemma 6. *Given an $\epsilon > 0$,*

$$
\Pr\left(\frac{1}{\lambda} - \epsilon < \mu(G_m) < \frac{1}{\lambda}\right) \to 1 \text{ as } K \to \infty.
$$

7

Proof. Let $L > 0$ be such that $\frac{1}{\lambda} - \epsilon < \mu(g_m), \forall g_m > L$. The existence of such an L follows from Lemma 5. The inequality $\mu(g_m) < \frac{1}{\lambda}$ always holds by Lemma 4 and Lemma 5. Hence

$$
\begin{aligned}
\Pr\left(\frac{1}{\lambda} - \epsilon < \mu(G_m) < \frac{1}{\lambda}\right) &= \Pr\left(\frac{1}{\lambda} - \epsilon < \mu(G_m)\right) \\
&\geq \Pr(G_m > L) \\
&= 1 - \Pr(G_m \leq L) \\
&= 1 - \prod_{i=1}^{K} \Pr(G_i \leq L)
\end{aligned}
$$

which approaches 1 as K approaches infinity, since each of the terms in the product is strictly less than 1 by the unbounded support assumption. $\square$

Lemma 7. $\frac{1}{\lambda} \to K\mathcal{P}$ as $K \to \infty$

Proof. From the average power constraint, we have

$$
\int_0^\infty \mu(g_m) p(g_m) dg_m = K\mathcal{P}
$$

Let

$$
A = \{g_m : \frac{1}{\lambda} - \epsilon < \mu(g_m) < \frac{1}{\lambda}\}
$$

and $A^c = [0,\infty) \backslash A$. Then $\Pr(A) \to 1$ as $K \to \infty$ by Lemma 6. The inequality $K\mathcal{P} < \frac{1}{\lambda}$ follows easily from the previous results, and so

$$
\begin{aligned}
&\frac{1}{\lambda} - K\mathcal{P} \\
&= \frac{1}{\lambda} - \int_0^\infty \mu(g_m) p(g_m) dg_m \\
&= \frac{1}{\lambda} - \int_A \mu(g_m) p(g_m) dg_m - \int_{A^c} \mu(g_m) p(g_m) dg_m \\
&\leq \frac{1}{\lambda} - \int_A \mu(g_m) p(g_m) dg_m \\
&\leq \frac{1}{\lambda} - \left(\frac{1}{\lambda} - \epsilon\right) \Pr(A)
\end{aligned}
$$

Since ϵ is arbitrary, the result follows. $\square$

Corollary 8. *Given an $\epsilon > 0$,*

$$
\Pr(K\mathcal{P} - \epsilon < \mu(G_m) < K\mathcal{P} + \epsilon) \to 1 \text{ as } K \to \infty
$$

Proof. From Lemma 7, we have for sufficiently large K

$$
\frac{1}{\lambda} - \epsilon < K\mathcal{P} < \frac{1}{\lambda} \tag{16}
$$

Combining (16) with Lemma 6, we obtain

$$
\begin{aligned}
\Pr&\left(\frac{1}{\lambda} - \epsilon < \mu(G_m) < \frac{1}{\lambda}\right) \\
&\leq \Pr(K\mathcal{P} - \epsilon < \mu(G_m) < K\mathcal{P} + \epsilon) \\
&\to 1.
\end{aligned}
$$

$\square$

Corollary 8 says that as $K \to \infty$, the power allocated to the best user will lie within ϵ of the constant value $K\mathcal{P}$ with probability approaching 1, for some arbitrary $\epsilon > 0$. Heuristically leting $\epsilon = 0$ and having each user transmitting at the *constant* power $K\mathcal{P}$ whenever it is the best, then corresponds to the second sub-optimal scheme proposed.

We can now prove the main result of this section.

Theorem 9. $C_{tts} \to R_{subopt2}$ as $K \to \infty$

Proof. First, note that the sum rates may be written as

$$
\begin{aligned}
C_{tts} &= \mathbb{E}_{G_m}\left[\mathbb{E}_{F|G_m}\left[\frac{1}{2}\log_2\left(1 + \frac{\mu(G_m)G_m F}{N_0}\right)\right]\right] \\
&= \int_{\frac{N_0\lambda}{\mathbb{E}[F]}}^\infty \int_0^\infty \frac{1}{2}\log_2\left(1 + \frac{\mu(g_m)g_m f}{N_0}\right) p(f)p(g_m) df dg_m
\end{aligned}
$$

and

$$
\begin{aligned}
R_{subopt2} &= \mathbb{E}_{G_m}\left[\mathbb{E}_{F|G_m}\left[\frac{1}{2}\log_2\left(1 + \frac{K\mathcal{P}G_m F}{N_0}\right)\right]\right] \\
&= \int_0^\infty \int_0^\infty \frac{1}{2}\log_2\left(1 + \frac{K\mathcal{P}g_m f}{N_0}\right) p(f)p(g_m) df dg_m.
\end{aligned}
$$

The main idea in the proof is that as $K \to \infty$, we can use Corollary 8 to restrict ourselves to the cases when $K\mathcal{P} - \epsilon < \mu(g_m) < K\mathcal{P} + \epsilon$ in the expression for C_{tts}, since the probability of this occurring tends to 1. To see this, let

$$
B = \{g_m : K\mathcal{P} - \epsilon < \mu(g_m) < K\mathcal{P} + \epsilon\}
$$

and $B^c = [\frac{N_0\lambda}{\mathbb{E}[F]}, \infty) \backslash B$. Then

$$
\begin{aligned}
C_{tts} &= \int_B \int_0^\infty \frac{1}{2}\log_2\left(1 + \frac{\mu(g_m)g_m f}{N_0}\right) p(f)p(g_m) df dg_m \\
&+ \int_{B^c} \int_0^\infty \frac{1}{2}\log_2\left(1 + \frac{\mu(g_m)g_m f}{N_0}\right) p(f)p(g_m) df dg_m.
\end{aligned}
$$

Now by the inequality $\mu(g_m) < \frac{1}{\lambda}$ and Lemma 7, for K sufficiently large, we will have $\mu(g_m) < K\mathcal{P} + \epsilon$ for all $g_m \in [\frac{N_0\lambda}{\mathbb{E}[F]}, \infty)$. Thus $\mu(g_m) < K\mathcal{P} - \epsilon, \forall g_m \in B^c$, and it may then be seen that the contribution

$$
\int_{B^c} \int_0^\infty \frac{1}{2}\log_2\left(1 + \frac{\mu(g_m)g_m f}{N_0}\right) p(f)p(g_m) df dg_m
$$

to C_{tts} will be negligible as $K \to \infty$. That is, as $K \to \infty$,

$$
C_{tts} \to \int_B \int_0^\infty \frac{1}{2}\log_2\left(1 + \frac{\mu(g_m)g_m f}{N_0}\right) p(f)p(g_m) df dg_m \tag{17}
$$

Next we define a function $\mu'(g_m) : [0,\infty) \to \mathbb{R}$ as

$$
\mu'(g_m) = \begin{cases} \mu(g_m) &, \quad g_m \in B \\ K\mathcal{P} &, \quad \text{otherwise} \end{cases}
$$

8

Then

$$\int_B \int_0^\infty \tfrac{1}{2}\log_2\left(1+\tfrac{\mu(g_m)g_m f}{N_0}\right)p(f)p(g_m)df\,dg_m$$

$$= \int_B \int_0^\infty \tfrac{1}{2}\log_2\left(1+\tfrac{\mu'(g_m)g_m f}{N_0}\right)p(f)p(g_m)df\,dg_m$$

$$< \int_0^\infty \int_0^\infty \tfrac{1}{2}\log_2\left(1+\tfrac{\mu'(g_m)g_m f}{N_0}\right)p(f)p(g_m)df\,dg_m$$

and by (17), we have for large enough K that

$$C_{tts} \leq \int_0^\infty \int_0^\infty \tfrac{1}{2}\log_2\left(1+\tfrac{\mu'(g_m)g_m f}{N_0}\right)p(f)p(g_m)df\,dg_m \tag{18}$$

Using (18) and the fact that $C_{tts} \geq R_{subopt2}$ (by definition), we thus have that as $K \to \infty$

$$|C_{tts} - R_{subopt2}|$$

$$\leq \left|\int_0^\infty \int_0^\infty \tfrac{1}{2}\log_2\left(1+\tfrac{\mu'(g_m)g_m f}{N_0}\right)p(f)p(g_m)df\,dg_m\right.$$
$$\left.- \int_0^\infty \int_0^\infty \tfrac{1}{2}\log_2\left(1+\tfrac{K\mathcal{P}g_m f}{N_0}\right)p(f)p(g_m)df\,dg_m\right|$$

$$\leq \int_0^\infty \int_0^\infty \left|\left(\tfrac{1}{2}\log_2\left(1+\tfrac{\mu'(g_m)g_m f}{N_0}\right)\right.\right.$$
$$\left.\left.-\tfrac{1}{2}\log_2\left(1+\tfrac{K\mathcal{P}g_m f}{N_0}\right)\right)p(f)p(g_m)\right|df\,dg_m$$

$$= \int_0^\infty \int_0^\infty \left|\tfrac{1}{2}\log_2\left(\tfrac{N_0+\mu'(g_m)g_m f}{N_0+K\mathcal{P}g_m f}\right)\right|p(f)p(g_m)df\,dg_m$$

Now notice that $\log_2\left(\tfrac{N_0+\mu'(g_m)g_m f}{N_0+K\mathcal{P}g_m f}\right)$ may be either positive or negative, depending on whether $K\mathcal{P} \leq \mu'(g_m) < K\mathcal{P}+\epsilon$ or $K\mathcal{P}-\epsilon < \mu'(g_m) < K\mathcal{P}$. In the case $K\mathcal{P} \leq \mu'(g_m) < K\mathcal{P}+\epsilon$, we have

$$\left|\log_2\left(\tfrac{N_0+\mu'(g_m)g_m f}{N_0+K\mathcal{P}g_m f}\right)\right| \leq \log_2\left(\tfrac{N_0+(K\mathcal{P}+\epsilon)g_m f}{N_0+K\mathcal{P}g_m f}\right)$$
$$= \log_2\left(1+\tfrac{\epsilon g_m f}{N_0+K\mathcal{P}g_m f}\right)$$

In the case $K\mathcal{P}-\epsilon < \mu'(g_m) < K\mathcal{P}$, we have

$$\left|\log_2\left(\tfrac{N_0+\mu'(g_m)g_m f}{N_0+K\mathcal{P}g_m f}\right)\right| \leq -\log_2\left(\tfrac{N_0+(K\mathcal{P}-\epsilon)g_m f}{N_0+K\mathcal{P}g_m f}\right)$$
$$= -\log_2\left(1-\tfrac{\epsilon g_m f}{N_0+K\mathcal{P}g_m f}\right)$$

Using the inequality $\ln(1+x) \leq -\ln(1-x), 0 \leq x < 1$ [1], we thus have

$$|C_{tts} - R_{subopt2}|$$

$$\leq \int_0^\infty\int_0^\infty -\tfrac{1}{2\ln 2}\ln\left(1-\tfrac{\epsilon g_m f}{N_0+K\mathcal{P}g_m f}\right)p(f)p(g_m)df\,dg_m$$

$$\leq \tfrac{1}{2\ln 2}\int_0^\infty\int_0^\infty \tfrac{\epsilon g_m f}{N_0+(K\mathcal{P}-\epsilon)g_m f}p(f)p(g_m)df\,dg_m$$

$$\leq \tfrac{1}{2\ln 2}\int_0^\infty\int_0^\infty \tfrac{\epsilon}{K\mathcal{P}-\epsilon}p(f)p(g_m)df\,dg_m$$

$$= \tfrac{\epsilon}{2(K\mathcal{P}-\epsilon)\ln 2}$$

where we have now used the inequality $-\ln(1-x) \leq \tfrac{x}{1-x}$ for $0 \leq x < 1$ [1].

Since $|C_{tts} - R_{subopt2}| \to 0$ as $K \to \infty$ (ϵ is also arbitrary), we are done. $\square$

High average SNR case

For the high average SNR ($\overline{\text{SNR}}$) case, we replace the statement $K \to \infty$ with the statement $\overline{\text{SNR}} \to \infty$ in the previous results, the proofs will be mostly very similar. The proof of Lemma 6 will hold provided the term $1 - \prod_{i=1}^K \Pr(G_i \leq L)$ goes to 1 as $\overline{\text{SNR}}$ approaches infinity. In other words, we need the (slow) fading distribution to satisfy the condition that for any $L > 0$, $\Pr(G_i \leq L)$ decreases to 0 as $\overline{\text{SNR}}$ increases. This holds for many common fading distributions, eg. lognormal, Rayleigh. However, if this condition is not satisfied, then the proof techniques of Corollary 8 and Theorem 9 will not apply, so there may not necessarily be convergence between the optimal and suboptimal policies. See [2] for such an example in the single user, single time-scale case.

6 Asymptotic analysis of the sum capacity

From a plot of the figures in Table 1, it seems that C_{tts} and $R_{subopt2}$ both scale logarithmically with the number of users K. In this section we will attempt to study the asymptotic behavior of C_{tts} as $K \to \infty$. Specifically, we show that $R_{subopt2}$ is upper-bounded by an expression which will be $O(\log K)$ provided a certain condition is satisfied, and we demonstrate that this condition holds for lognormal slow fading. By definition, this then implies that $R_{subopt2}$ is $O(\log K)$. This also implies that C_{tts} is $O(\log K)$, as $C_{tts} \to R_{subopt2}$ by Theorem 9. Note that this result does not say that $R_{subopt2}/C_{tts}$ grows exactly like $\log K$, but only that $R_{subopt2}/C_{tts}$ cannot grow any faster than $\log K$.

Lemma 10. *Let $G_m \equiv G_{m,K} = \max(G_1,\ldots,G_K)$. As $K \to \infty$, if*

$$\mathbb{E}[G_m] = O(K^n)$$

for some $n > 0$, then $R_{subopt2} = O(\log K)$.

Proof. Recall that using the sub-optimal scheme of Section 4, the sum rate may be written as

$$R_{subopt2} = \mathbb{E}_{G_m}\left[\mathbb{E}_{F|G_m}\left[\tfrac{1}{2}\log_2\left(1+\tfrac{K\mathcal{P}G_m F}{N_0}\right)\right]\right].$$

As $K \to \infty$, we have

$$R_{subopt2} \leq \tfrac{1}{2}\log_2\left(1+\tfrac{K\mathcal{P}\mathbb{E}[G_m]\mathbb{E}[F]}{N_0}\right)$$

$$\sim \tfrac{1}{2}\log_2\left(\tfrac{K\mathcal{P}\mathbb{E}[G_m]\mathbb{E}[F]}{N_0}\right)$$

$$= \tfrac{1}{2}\left(\log_2 K + \log_2 \mathbb{E}[G_m] + \log_2\tfrac{\mathcal{P}\mathbb{E}[F]}{N_0}\right)$$

9

where the inequality comes from applying Jensen's inequality twice. Noticing that the first term of the expression

$$\log_2 K + \log_2 \mathbb{E}[G_m] + \log_2 \frac{\mathcal{P}\mathbb{E}[F]}{N_0}$$

is $O(\log K)$ and the last term is a constant, we only need to ensure that $\log_2 \mathbb{E}[G_m]$ will be $O(\log K)$. This occurs provided that $\mathbb{E}[G_m]$ does not grow too quickly with K. $\mathbb{E}[G_m] = O(K^n)$ for some $n > 0$ will certainly suffice. $\square$

We will now show that this condition is satisfied for log-normal slow fading. Without loss of generality, let each of the G_i's have the 'standard' lognormal distribution with pdf and cdf given by

$$p_G(g) = \frac{1}{\sqrt{2\pi}g} \exp\left[\frac{-(\ln g)^2}{2}\right], g > 0$$

$$P_G(g) = \frac{1}{2} + \frac{1}{2}\mathrm{erf}\left[\frac{\ln g}{\sqrt{2}}\right], g > 0$$

Following the idea of [14], we will look at the limiting distribution of G_m as $K \to \infty$. Note however that Lemma 2 as stated in [14] will not actually apply to the lognormal distribution, as one can show that $\frac{1-P(g)}{p(g)} \sim \frac{g}{\ln g}$, which diverges as $g \to \infty$. Fortunately, more general statements of this result as given in eg. [7] does still apply. From [7], we have the following.

Proposition 2. *Let* $X_1, X_2, \ldots, X_K$ *be i.i.d. random variables with common distribution function* $P(x)$. *Let* $Z_K = \max(X_1, X_2, \ldots, X_K)$ *and* $H_K(x) = \Pr(Z_K \le x)$. *Then as* $K \to \infty$, $H_K(a_K + b_K x)$ *converges in distribution to one of the three distribution functions* $H_{1,\gamma}(x), H_{2,\gamma}(x)$ *or* $H_{3,0}(x)$, *where*

$$H_{1,\gamma}(x) = \begin{cases} \exp(-x^{-\gamma}) & , \quad x > 0 \\ 0 & , \quad otherwise \end{cases}$$

$$H_{2,\gamma}(x) = \begin{cases} 1 & , \quad x \ge 0 \\ \exp(-(-x)^{-\gamma}) & , \quad x < 0 \end{cases}$$

$$H_{3,0}(x) = \exp(-e^{-x}), x \in \mathbb{R}$$

and a_K *and* $b_K > 0$ *are suitable constants.*

Technical details, such as the definition of γ, and how to determine the constants a_K and b_K, may be found in [7]. For the standard lognormal distribution, it has been shown that the limiting distribution is of the third type, ie.

$$\Pr(G_m < a_K + b_K x) \to \exp(-e^{-x}) \text{ as } K \to \infty \quad (19)$$

where the normalising constants a_K and b_K are

$$a_K = \exp\left[(2\ln K)^{1/2} - \tfrac{1}{2}(2\ln K)^{-1/2}(\ln\ln K + \ln 4\pi)\right]$$

$$b_K = (2\ln K)^{-1/2}a_K$$

Let Λ be the random variable with cdf $P_\Lambda(x) = \exp(-e^{-x})$. Then (19) implies that

$$\mathbb{E}[G_m] \sim a_K + b_K\mathbb{E}[\Lambda]$$
$$= a_K + b_K\gamma^*$$

where γ^* is Euler's constant. It is easy to see that $\mathbb{E}[G_m] = O(a_K)$ and very roughly that $a_K = O(K^2)$, so that $\mathbb{E}[G_m] = O(K^2)$. A graph of a_K indicates that much sharper bounds are possible, but our rough bound on $\mathbb{E}[G_m]$ is nevertheless sufficient to show from Lemma 10 that both $R_{subopt2}$ and C_{tts} are $O(\log K)$.

References

[1] M. Abramowitz and I. A. Stegun. *Handbook of Mathematical Functions*. Dover Publications, New York, 1965.

[2] M.-S. Alouini and A. J. Goldsmith. Comparison of fading channel capacity under different CSI assumptions. In *Proc. IEEE Veh. Technol. Conf.*, pages 1844–1849, Boston, MA, Sep. 2000.

[3] R. G. Bartle. *The Elements of Integration and Lebesgue Measure*. John Wiley & Sons, New York, Wiley Classics Library edition, 1995.

[4] E. Biglieri, J. Proakis, and S. Shamai (Shitz). Fading channels: Information-theoretic and communications aspects. *IEEE Trans. Inform. Theory*, 44(6):2619–2692, Oct. 1998.

[5] G. Caire and S. Shamai (Shitz). On the capacity of some channels with channel state information. *IEEE Trans. Inform. Theory*, 45(6):2007–2019, Sep. 1999.

[6] S. Dey and J. S. Evans. Optimal power control over multiple time-scale fading channels with service outage constraints. In *Proc. WiOpt'04*, pages 254–263, Cambridge, U.K., Mar. 2004. Also *IEEE Trans. Commun.*, to appear.

[7] J. Galambos. *The Asymptotic Theory of Extreme Order Statistics*. John Wiley & Sons, New York, 1978.

[8] A. J. Goldsmith. *Wireless Communications*. Cambridge University Press, 2005. To be published.

[9] A. J. Goldsmith and P. P. Varaiya. Capacity of fading channels with channel side information. *IEEE Trans. Inform. Theory*, 43(6):1986–1992, Nov. 1997.

[10] R. Knopp and P. A. Humblet. Information capacity and power control in single-cell multiuser communications. In *Proc. IEEE. Int. Conf. Commun.*, pages 331–335, Seattle, WA, Jun. 1995.

[11] D. G. Luenberger. *Optimization by Vector Space Methods*. John Wiley & Sons, New York, 1969.

[12] D. N. C. Tse and S. V. Hanly. Multiaccess fading channels - Part I: Polymatroid structure, optimal resource allocation and throughput capacities. *IEEE Trans. Inform. Theory*, 44(7):2796–2815, Nov. 1998.

[13] D. N. C. Tse and P. Viswanath. *Fundamentals of Wireless Communication*. Cambridge University Press, 2005. To be published.

[14] P. Viswanath, D. N. C. Tse, and R. Laroia. Opportunistic beamforming using dumb antennas. *IEEE Trans. Inform. Theory*, 48(6):1277–1294, Jun. 2002.