Multiuser Diversity *†

Raymond Knopp‡, Pierre A. Humblet§

May 30, 2002

Abstract

This work builds upon earlier results presented in [1, 2], where the authors considered power allocation strategies for multiuser fading channels without processing delay constraints. These optimal strategies make the assumption that perfect channel state information is made available to the users. Our work has since been extended by several authors, most notably the remarkable work of Tse and Hanly [3]. Our results are a special case of their general theory.

Using a general continuous-time wideband *block-fading* channel model we give upper and lower bounds to the probability of joint decoding error for multiuser communication. We then specialize our treatment to *ergodic* or *high diversity* systems where arbitrarily small error probabilities can be achieved. The power spectral allocation strategy for maximizing total system throughput is derived and shown to be an orthogonal dynamic frequency assignment based on the channel responses of *all* the users. The primary quantitative result is that significantly higher data rates are achievable on fading multiuser channels than on non-fading AWGN multiuser channels. Numerical comparisons with traditional multiple-access schemes not benefiting from transmitter channel state side information are made. An interesting by-product is that the allocation strategy is equivalent to a transmit antenna selection diversity scheme, which we appropriately term *multiuser diversity*. We extend the analysis to a worst-case cellular interference channel model and conclude that similar improvements over traditional multiple-access schemes can be expected.

Keywords:Multiple-Access, Fading Channels, Diversity, Dynamic Channel Allocation, Power Control

---

*Eurecom's research is partially supported by its industrial partners: Ascom, Cegetel, Hitachi, IBM France, Motorola, Swisscom, Texas Instruments, and Thomson CSF

†R. Knopp's research was funded in part by a FCAR (Quebec, Canada) doctoral grant

‡R. Knopp (raymond.knopp@epfl.ch) is with Mobile Communications Laboratory (http://lcmwww.epfl.ch) at the Swiss Federal Institute of Technology, Lausanne(EPFL, http://www.epfl.ch)

§P.A. Humblet(humblet@eurecom.fr) is with Institut Eurécom, Sophia Antipolis France. E-mail:humblet@eurecom.fr

1 Introduction

Recent advances in communication electronics have paved the way for the use of sophisticated signal design and processing techniques in wireless infrastructure. The coding and multiple-access techniques used in systems like IS-95 [4] and GSM [5] are examples of how communication theory has seeped its way into wireless communications practice. This will surely continue into the third generation of wireless communications as UMTS/IMT-2000 [6] becomes a reality.

In parallel to these technological advances, recent theoretical work in the area of wireless communications continues to open up many new and promising avenues for systems design. The goal of this paper is to investigate the effect of using channel state (side) information at the transmission end of a wireless $M-$user system perturbed by slowly-fading multipath propagation. One may question the feasibility of having complete channel state information at the transmission end. We maintain, however, that for *time-division duplex (TDD)* systems, it is definitely a possibility, and furthermore it is already used to some extent in systems like DECT [7] and PHS [8]. Since both links are duplexed in time and can therefore share the same frequency band, measurements of the frequency response across the *entire system bandwidth* can be made at the terminal end since the basestation broadcasts on all frequency channels. This is clearly critical for performing optimal spectral allocation, which as we will see results in the per-user use of only portions of the available bandwidth.

In *frequency-division duplex (FDD)* systems it is more difficult to obtain channel state information at the transmitters for two reasons. Firstly, a feedback channel is required (like the power control scheme of IS-95 [4]). Secondly, and more importantly, the basestation can only estimate the response of the users in the frequency band(s) to which they are assigned, and can thus not perform optimal allocation. A possible method for alleviate the second drawback would be to use very low-power spread-spectrum measurement signals superimposed on top of each user's information signal. The measurement signals would occupy the entire signaling bandwidth, whereas the users signals power would be allocated in an optimal fashion.

Shamai and Wyner [9, 10] build upon earlier works of Wyner [11] and Hanly [12] dealing with cellular multiple-access channels. They treat the effect of fading and with and without distributed processing in such systems.

Our work was originally motivated by the single-user studies of Goldsmith and Varaiya[13, 14]. Some of the results presented here appeared in [1, 2]. Recently, other authors have extended these results, and many new results are summarized in [15]. Specifically the remarkable work by Tse and Hanly [3] completely characterize the capacity region of a discrete-time multiple-access channel with optimal power allocation, both with and without processing delay constraints. Our main result is a special case of their general theory, in the sense that we are concerned only with the power allocation strategy which maximizes the total throughput of the system, and not those which achieve any possible set of information rates.

Our main conclusions are primarily of a quantitative nature and we have tried to focus mainly on systems issues. We will show that power controlled multiple-access techniques can provide significant performance improvement over conventional signaling schemes, such as spread-spectrum and orthogonal multiplexing without spectral optimization (CDMA,FDMA,TDMA,F-TDMA). Quantitatively, we may expect close to a twofold increase in spectral efficiency compared to conventional spread-spectrum and orthogonal multiplexing single-cell systems for practical signal-to-noise ratios (SNR). Furthermore, these gains are possible only by *orthogonally multiplexing* users' information signals, which we will show is a sort of *dynamic channel allocation*. This results in a simple system requiring single-user codes without the need for successive interference cancellation as is the case on non-fading AWGN channel [16]. We will also show that similar gains are possible in interference-limited cellular systems which do not benefit from distributed processing (for results on distributed processing systems see [11, 12, 17, 9, 10]).

1.1 Multiuser Diversity

In wireless systems, one normally attempts to provide some mechanism for combating the effects of fading due to multipath propagation. Typically this is an antenna diversity [18] scheme which was traditionally limited to the receiver. This has recently resurfaced with a promising coded-modulation technique known as *space-time coding* which allows for similar performance gains with multiple transmit antennae [19, 20, 21]. Alternately as in the GSM system [5], one may frequency-diversity with error control codes [22, 23, 24, 25, 26]. For quickly time-varying systems, similar techniques are achieved by interleaving coded signals [27, 28, 29]. We will show that a similar diversity effect, which we term *multiuser diversity*, can be achieved without multiple-antennae and results in the aforementioned performance enhancement. The optimal power allocation strategy is essentially a transmitter diversity scheme akin to selection diversity with multiple transmit antennae. One astonishing result is that significantly more information can be transmitted across a multiuser fading AWGN channel than a *non-fading* AWGN channel for the same average signal power at the receiver. It stems from the fact that the channel gain at a particular time and frequency is random and can be significantly higher than its average level. Using proper dynamic time-frequency allocation one can take advantage of this in a multiuser system since resources must be shared by the users. The fading channel gives a non-negligible power boost which can be as high as 5 or 6 dB.

2 Wideband Block-Fading Multiuser Channels

Our signal and channel model consists of block (or packet) based M−user systems with slow multipath fading and additive white Gaussian noise. It is a model studied by several authors for single-user systems [30, 31, 32, 22, 24, 25]. Caire *et al* have used a similar model to study F-TDMA multiuser cellular systems. In [33], Caire *et al* consider the use of channel state information at the transmitter for single-user block-fading channels under processing delay constraint.

We assume that data is coded and split into signal blocks of duration T seconds, and that T is chosen to be shorter than the coherence time of the channel. This means that the channel is stationary for the duration of a block. A guard–time $T_G \ll T$ is inserted between blocks so that the signals from each block do not overlap at the receiver due to time-dispersion. If the delay–spread of the channel is at most T_s seconds we take $T_G > T_s$. Note that since this is a multiuser system, to use a guard-time as short as possible, some loose synchronization between users is required, but by no means are users signals synchronous. We may call them block or frame synchronous. T_G should also be chosen, therefore, to take into account any residual asynchronism between the users. A diagram of this system is shown in Figure 1.

The n^{th} signal block for user m is denoted by the real signal $x_{m,n}(t)$ which is non-zero only in the interval $(n-1)(T+T_G) \leq t \leq (n-1)(T+T_G)+T$. Assuming that the information is coded (and decoded) across N such blocks, the signal processed by the receiver is given by

$$\begin{aligned} y(t) &= \sum_{n=1}^{N} y_n(t) \\ &= \sum_{m=1}^{M}\sum_{n=1}^{N} \int_{(n-1)(T+T_G)}^{(n-1)(T+T_G)+T} h_{m,n}(t-\tau)x_{m,n}(\tau)d\tau + z_n(t) \\ &= \sum_{m=1}^{M}\sum_{n=1}^{N} u_{m,n}(t) + z_n(t) \\ &= \sum_{n=1}^{N}\left[\sum_{m\in S} u_{m,n}(t) + \sum_{m\in S^c} u_{m,n}(t)\right], \end{aligned} \tag{1}$$

where S denotes an arbitrary subset of $\{1, 2, \cdots, M\}$. The transmit power of each user is assumed to satisfy the constraint

$$\sum_{n=1}^{N} \mathrm{E}\; x_{m,n}^2(t)dt \leq N(T+T_G)P_m, \tag{2}$$

given that $x_{m,n}(t)$ is stationary during each block. The information rate of each user, denoted R_i bits/sec, is agreed upon beforehand by the transmitters and receiver and is not changed during the course of communication. The noise processes, $z_n(t)$, are Gaussian and white with power spectral densities $N_0/2$. The input signals are taken to be virtually bandlimited to a frequency band $[-W, W]$ Hz, in the sense that an arbitrarily small amount of signal energy is present outside of this band. We assume wideband systems so that W is considerably larger than the coherence bandwidth of the channel processes.

We will always assume that the channels are known perfectly to the receiver but not necessarily to the transmitter. This implies that some form of training information is transmitted along with the data. In order not to reduce information rates, the amount of training information must be small compared to T which, in turn, implies that the channel cannot vary significantly during each block. This assumption is therefore valid in our case. The receiver also knows the second-order statistics of the input and noise signals.

The channel-corrupted information signal sum in block n of the users in $\mathcal{S}$ may be written as

$$u_{\mathcal{S},n}(t) = \sum_{m \in \mathcal{S}} u_{m,n}(t) = \sum_{i=1}^{\infty} u_{\mathcal{S},n,i} \phi_{\mathcal{S},n,i}(t : T, \mathbf{h}), \tag{3}$$

where $\{\phi_{\mathcal{S},n,i}(t : T, \mathbf{h}), (n-1)(T+T_G) \le t \le (n-1)(T+T_G)+T, i = 1, \cdots, \infty\}$ is an arbitrary orthonormal basis over $L^2((n-1)(T+T_G), (n-1)(T+T_G)+T)$. We have explicitly indicated the dependence of the eigenfunctions on the signal block duration, T, and the set of channel realizations which we denote $\mathbf{h}$. It can be a random vector (known to the receiver) which contains all the information concerning the channel responses (e.g. path gains and delays). Note that the receiver can, at least in principle, compute the $\{\phi_{\mathcal{S},n,i}(t : T, \mathbf{h})\}$ since the channel realization and the input statistics are known.

2.1 Achievable Rates

We now choose the basis functions for each block, $\{\phi_{\mathcal{S},n,i}(t : T, \mathbf{h})\}$, to be the eigenfunctions of the covariance function of $u_{\mathcal{S},n}(t)$ (i.e. a Karhunen-Loève expansion)

$$K_{\mathcal{S},n}(t, t' : T, \mathbf{h}) = \sum_{m \in \mathcal{S}} \int_{(n-1)(T+T_G)}^{(n-1)(T+T_G)+T} \int_{(n-1)(T+T_G)}^{(n-1)(T+T_G)+T} K_{m,n}(\tau - \tau' : T) h(t-\tau) h(t'-\tau') d\tau d\tau' \tag{4}$$

where $K_{m,n}(\tau - \tau' : T)$ is the covariance function of the $x_{m,n}(t)$, which are assumed to be stationary during the block, in the sense that

$$K_{m,n}(t, t' : T) = \begin{cases} K_{m,n}(t - t' : T), & (n-1)(T+T_G) \le t, t' \le (n-1)(T+T_G)+T \\ 0, & \text{otherwise.} \end{cases}$$

Note that the input autocorrelation functions are asymptotically stationary (in each block) as are the $K_{\mathcal{S},n}(t, t' : T, \mathbf{h})$. We take for granted that users transmit (statistically) independently of one another. The $\{u_{\mathcal{S},n,i}\}$ are independent, have zero–mean and variances $\{\lambda_{\mathcal{S},n,i}(T, \mathbf{h})\}$ which, along with $\{\phi_{\mathcal{S},n,i}(t : T, \mathbf{h})\}$, are the solution to

$$\lambda_{\mathcal{S},n,i}(T, \mathbf{h}) \phi_{\mathcal{S},n,i}(t : T, \mathbf{h}) = \int_{(n-1)(T+T_G)}^{(n-1)(T+T_G)+T} K_{\mathcal{S},n}(t, t') \phi_{\mathcal{S},n,i}(t' : T, \mathbf{h}) dt'. \tag{5}$$

We may now write the received signal in terms of its expansion coefficients as

$$y_{\mathcal{S},n,i} = u_{\mathcal{S},n,i} + u_{\mathcal{S}^c,n,i} + z_{n,i} \tag{6}$$

where $u_{\mathcal{S}^c,n,i}$ and $z_{n,i}$ are the projections of the signals of users in $\mathcal{S}^c$ and the noise on $\{\phi_{\mathcal{S},n,i}(t : T, \mathbf{h})\}$. The discrete representation in (6) allows for the following theorem bounding the probability of decoding error for joint decoding of the signals, $P_e(\mathbf{R}, N, T, T_G)$.

Theorem 1 *For a block-fading channel with block duration T and guard-time T_G, there exist joint channel codes with probability of error when jointly decoding a set of M users over N blocks satisfying*

$$\begin{aligned} K(\mathbf{R}, N, T, T_G) P_{\text{out}}(\mathbf{R}, N, T, T_G) &\leq P_e(\mathbf{R}, N, T, T_G) \\ &\leq \min \Bigg(1, \sum_{S \in \{1,2,\cdots,M\}} P_{\text{out}}(\mathbf{R}, S, N, T, T_G) + \\ &\int_{\mathbf{h}: I(\mathcal{S},\mathbf{h}) < \sum_{i \in S} R_i} 2^{-N(T+T_G)E_r(\mathbf{R},\mathbf{h},N,T,T_G)} dF_{\mathbf{H}}(\mathbf{h}) \Bigg) \end{aligned} \tag{7}$$

where $K(\mathbf{R}, N, T, T_G)$ is a constant approaching 1 with increasing $N(T + T_G)$,

$$E_r(\{R_i, i \in S\}, S, \mathbf{h}, N, T, T_G) = \max_{0 \leq \rho \leq 1} \frac{1}{N(T+T_G)} \sum_{n=1}^{N} \sum_{i=1}^{\infty} E_0(\rho, \mathbf{h}, U_{\mathcal{S},n,i}, Y_{\mathcal{S},n,i}) - \rho \sum_{i \in \mathcal{S}} R_i \ \text{ bits/s}$$

$$E_0(\rho, \mathbf{h}, U_{\mathcal{S},n,i}, Y_{\mathcal{S},n,i}) = \frac{\rho}{2} \log_2 \left(1 + \frac{2\lambda_{\mathcal{S},n,i}(T, \mathbf{h})}{(1+\rho)N_0} \right) \ \text{ bits/s} \tag{8}$$

$$I(\mathcal{S}, \mathbf{h}, N, T, T_G) = \frac{1}{2N(T+T_G)} \sum_{n=1}^{N} \sum_{i=1}^{\infty} \log_2 \left(1 + \frac{2\lambda_{\mathcal{S},n,i}(T, \mathbf{h})}{N_0} \right) \ \text{ bits/s,} \tag{9}$$

$$P_{\text{out}}(\mathbf{R}, S, N, T, T_G) = \Pr \left(\sum_{i \in S} R_i > I(S, \mathbf{h}) \right) \tag{10}$$

and

$$P_{\text{out}}(\mathbf{R}, N, T, T_G) = \Pr \left(\bigcup_{S \in \{1,2,\cdots,M\}} \left\{ \sum_{i \in S} R_i > I(\mathcal{S}, \mathbf{h}) \right\} \right) \tag{11}$$

Moreover, there exists no channel code with probability of error less than the lower bound.

Proof: (see Appendix 1)

This shows that reliable communication may be impossible [1], unless $P_{\text{out}}(\mathbf{R}, N, T, T_G) = P_{\text{out}}(\mathbf{R}, S, N, T, T_G) = 0$. Furthermore, unlike the single-user case [22, 23], the upper and lower bounds do not converge in the limit of large $T + T_G$.

[1] in the sense that there will always be an irreducible error probability, irrespective of the codeword duration

Corollary 1 (Ergodic Channels) *For ergodic sequence of channels* $\{h_n(t), n = 1, \cdots, \infty\}$ *we have that*

$$\lim_{N\to\infty} P_{\text{out}}(\mathbf{R}, N, T, T_G) = \lim_{N\to\infty} \min\left(1, \sum_{S\subseteq\{1,2,\cdots,M\}} P_{\text{out}}(\mathbf{R}, S, N, T, T_G)\right)$$
$$= \mathcal{I}\left(\mathbf{R} \in \bigcup_{S\in\{1,2,\cdots,M\}} \left\{\sum_{i\in\mathcal{S}} R_i \leq \mathrm{E}_{\mathbf{H}} I(\mathcal{S}, \mathbf{h})\right\}\right),$$

where $\mathcal{I}(\cdot)$ *is the indicator function.*

Proof: (See Appendix 1)

The practical interpretation of Corollary 1 is that in such cases, which correspond to *high diversity* systems, arbitrarily small error probabilities are possible if the set of information rates satisfies the $2^M - 1$ inequalities $\sum_{i\in\mathcal{S}} R_i \leq \mathrm{E}_{\mathbf{H}} I(\mathcal{S}, \mathbf{h}), \forall \mathcal{S} \subseteq \{1, 2, \cdots, M\}$. We do not require the time-invariant assumption on the channel responses to arrive at these results. For the following corollary, however, this is required in order to obtain useful expressions for computational purposes.

Corollary 2 (Limiting Expressions) *For* $T \to \infty$ *we have that*

$$I(\mathcal{S}, \mathbf{h}) = \frac{1}{2N} \sum_{n=1}^{N} \int_{-W}^{W} \log_2\left(1 + \frac{2}{N_0} \sum_{i\in\mathcal{S}} S_{i,n}(f)|H_{i,n}(f)|^2\right) df \tag{12}$$

where $S_{i,n}(f)$ *is the power spectral density* $\int_{-\infty}^{\infty} K_{i,n}(t)e^{-j2\pi ft}dt$ *and* $H_{i,n}(f)$ *is the Fourier transform* $\int_{-\infty}^{\infty} h(t)e^{-j2\pi ft}dt$. *The average power constraints become*

$$\sum_{n=1}^{N} \int_{-W}^{W} S_{m,n}(f)df \leq NP_m. \tag{13}$$

Proof: (see Appendix 1)

2.2 Fast and Slow Fading

It is typical of wireless radio channels to exhibit both fast and slow fading phenomena. Fast fading arises from multipath propagation which causes constructive and destructive interference of the multiple paths. The rate of change of the signal strength is a function of the carrier wavelength and the relative velocity between the receiver and transmitter, as well as changes in the environment (i.e. moving vehicles, doors, people, etc.). Significant fluctuations occur typically when the receiver/transmitter changes position by a few centimeters. Slow fading is a result of the electromagnetic path loss due to the separation between the receiver and the transmitter which fluctuates when positions change on the order of tens of meters. Another slow fading phenomena, called *shadowing*, arises from a varying portion of the multipath being blocked by large objects such as trees or buildings. The rate of variation of this component is a subject of great debate. Position changes anywhere from meters to tens of meters can yield signal strength variation in the shadowing component, depending on the nature of the environment. Fast fading is frequency-selective whereas slow-fading is virtually frequency-independent. The two fading time-scales are illustrated in Figure 2 where we plot the magnitude of the channel response to a signal $e^{j2\pi ft}$ as a function of t.

In the remainder of this work we will assume that the slow component remains constant during the transmission of the N blocks and that the average fast fading strength at a given frequency is normalized to unity.

2.3 Input Spectra without Channel State Information at the Transmitters

Before considering the optimal input spectra when the transmitters have perfect knowledge of the channels (or can be guided by the receiver), we consider the case when no knowledge is available. Intuition tells us that we should assign energy equally in the spectrum since all frequencies are identically distributed, which is a ideal spread-spectrum system. This is indeed the case, but it is somewhat delicate to show, and to our knowledge this has never been proven in the literature[2]. For the sake of completeness, we have the following theorem:

Theorem 2 *For an ergodic block-fading multiuser system without channel state information and time-invariant slow-fading, the optimal input spectra are of the form*

$$S_m(f) = \frac{P_m}{2W}\mathcal{I}(f \in [-W, W]) \tag{14}$$

Proof:

Suppose we choose an arbitrary $S_m(f)$ for each user and that we partition it into L equally spaced intervals in both the positive and negative frequencies. We can consider the $L!$ permutations of these spectra $S_{m,\boldsymbol{\pi}_l}(f), l = 1, 2, \cdots, L!$ formed by swapping the intervals (similarly in positive and negative frequencies to preserve symmetry) according to the permutation $\boldsymbol{\pi}$. It is clear that the ergodic rate sums are identical for each $\boldsymbol{\pi}$, since the channel responses at each frequency are identically distributed. Now, because of the convexity of the logarithm we have

$$\begin{aligned}\sum_{m\in\mathcal{S}} R_m &\leq \frac{1}{L!}\sum_{\boldsymbol{\pi}} \mathrm{E} \int_{-W}^{W} \log_2\left(1 + \frac{2}{N_0}\sum_{m\in\mathcal{S}}|H_m(f)|^2 S_{m,\boldsymbol{\pi}}(f)\right) df \\ &\leq \mathrm{E} \int_{-W}^{W} \log_2\left(1 + \frac{2}{N_0}\sum_{m\in\mathcal{S}}|H_m(f)|^2 T_{m,L}(f)\right) df \end{aligned} \tag{15}$$

[2]It was however shown by Gallager that spread-spectrum outperforms frequency-division multiple-access [34] for an idealized fading channel model

where

$$
\begin{aligned}
T_{m,L}(f) &= \frac{1}{L!}\sum_{\boldsymbol{\pi}} S_{m,\boldsymbol{\pi}}(f) \\
&= \frac{1}{L}\sum_{k=0}^{L-1} S_m\left(f + \frac{(k-l)W}{L}\right), \quad f \in \left[\frac{lW}{L}, \frac{(l+1)W}{L}\right]
\end{aligned} \tag{16}
$$

which is periodic with period W/L. Defining $\Delta\theta = \frac{W}{L}$ we have

$$
\begin{aligned}
\lim_{L\to\infty} T_{m,L}(f) &= \frac{1}{W}\lim_{\Delta\theta\to 0}\sum_{k=0}^{L-1} S_m\left(f + (k-l)\Delta\theta\right)\Delta\theta, \quad f \in [l\Delta\theta, (l+1)\Delta\theta] \\
&= \frac{1}{2W}\int_{-W}^{W} S_m(\theta)d\theta, \ \ \forall f.
\end{aligned} \tag{17}
$$

Equality in (15) is achieved when $S_m(f)$ is a constant satisfying the power constraint as in (14).

2.4 Optimal Power Allocation for Ergodic Channels

As discussed in the introduction, we are primarily interested in examining the effectiveness of exploiting channel state information at the transmission end for achieving high information rates. In our case, this amounts to choosing the input spectra to maximize the information rates using this *a priori* knowledge. This problem was treated for non-fading two-user, discrete-time, finite-memory ISI channels by Cheng and Verdú in [35]. Recently Tse and Hanly [3] have applied similar ideas to a discrete-time fading channel model with an arbitrary number of users. The techniques are essentially the same in our case, so we summarize these results in the following theorem which is stated without proof:

Theorem 3 (Cheng & Verdú[35], Tse & Hanly[3]) *For an ergodic block-fading multiuser system with channel state information at the transmission end and time-invariant slow-fading, the set of maximal rates and corresponding power spectra are parametrically described by the solution to the maximization problem*

$$
\{R_i(\boldsymbol{\alpha}), i = 1, \cdots, M\} = \underset{\{S_{i,n}(f), i=1,\cdots,M\}}{\arg\max} \sum_{i=1}^{M} \alpha_i R_i - \lambda_i \mathrm{E}\ S_{i,n}(f) \tag{18}
$$

where α_i are parameters defining the maximal rates and λ_i are Lagrange multipliers *satisfying the power constraints.*

This theorem stems from the fact that the capacity region is convex so that each point on its boundary touches the plane $\sum_{i=1}^{M}\alpha_i R_i$ at least once for each set $\{\alpha_i\}$. Furthermore, for each set $\{\alpha_i\}$ the $R_i(\boldsymbol{\alpha})$ can be computed by successive decoding according to the $\{\alpha_i\}$ sorted in increasing order. The general solution to this maximization problem is solved elegantly in [3] and makes use of the *polymatroidal* structure of the achievable rate region of the multiple-access channel.

An important point of the capacity region is the one corresponding to the total rate sum (i.e. when $\alpha_i = 1, \forall i$). We have the following corollary to Theorem 3.

Corollary 3 (Total Rate Sum) *The point on the capacity region corresponding to the maximum total ergodic rate sum $\sum_{i=1}^{M} R_i$ is achievable with input spectra*

$$S_{m,n}(f) = \begin{cases} \left[\frac{1}{\lambda_m} - \frac{N_0}{2}\frac{1}{|H_{m,n}(f)|^2}\right]^+ & |H_{m,n}(f)|^2 \geq \frac{\lambda_m}{\lambda_{m'}}|H_{m',n}(f)|^2, \forall m' \neq m \\ 0 & \text{otherwise} \end{cases} \tag{19}$$

yielding maximal information rates

$$R_m = \frac{1}{2}\int_{-W}^{W} \mathrm{E}\left\{\mathcal{I}\left(|H_{m,n}(f)|^2 = \max\left[0, \left\{\frac{\lambda_m}{\lambda_{m'}}|H_{m',n}(f)|^2, m' = 1, \cdots, M\right\}\right]\right)\right. \\ \left.\log_2\left(\frac{2}{N_0\lambda_m}|H_{m,n}(f)|^2\right)\right\} df \ \ \text{bits/s} \tag{20}$$

if the events $|H_{m,n}(f)|^2 = \frac{\lambda_m}{\lambda_{m'}}|H_{m',n}(f)|^2$ occur with probability zero for $m \neq m'$.

Proof:

Since $\log(\cdot)$ is a convex $\cap$ function, the Kuhn-Tucker conditions [36] applied to the maximization problem in (18) for $\alpha_i = 1, \forall i$ are

$$\frac{\partial \log_2\left(1 + \frac{2}{N_0}\sum_{m=1}^{M} S_{m,n}(f)|H_{m,n}(f)|^2\right)}{\partial S_{m,n}(f)} = \frac{|H_{m,n}(f)|^2}{\frac{N_0}{2} + \sum_{m=1}^{M} S_{m,n}(f)|H_{m,n}(f)|^2} \leq \lambda_m, \quad m = 1, \cdots, M \tag{21}$$

with equality if $S_{m,n}(f) > 0$. If we define the function $K_n(f, \mathcal{E})$ for some set $\mathcal{E}$ of users such that

$$K_n(f, \mathcal{E}) = \frac{|H_{m,n}(f)|^2}{\lambda_m}, \ \forall m \in \mathcal{E} \subseteq \{1, 2, \cdots, M\} \tag{22}$$

and

$$K_n(f, \mathcal{E}) > \frac{|H_{m,n}(f)|^2}{\lambda_m}, \ \forall m \in \mathcal{E}^c \tag{23}$$

then the power spectra of the users in $\mathcal{E}$ must satisfy

$$\sum_{m\in\mathcal{E}} \lambda_m S_{m,n}(f) = \left[K_n(f, \mathcal{E}) - \frac{N_0}{2}\right]^+ \tag{24}$$

and $S_{m,n}(f) = 0$, $\forall m \in \mathcal{E}^c$.

In general, an infinite number of points maximize the total rate sum. If, however, the distributions of $|H_{m,n}(f)|^2$ for each f are continuous, then the cardinality of $\mathcal{E}$ is 1 with probability 1 and the power spectra are given by (19). Since only a single user occupies each frequency band in a given block, joint decoding is not required and the ergodic rates are given by (20).

Our first observation is that this point is achievable only with *orthogonal multiplexing.* We note, on one hand, that this is not the case on the non-fading Gaussian channel where there is one point on the capacity region which is achievable both by orthogonal and non-orthogonal multiplexing [37]. On the other hand, when channel state information is unavailable at the transmitter and the system suffers from multipath fading, only a non-orthogonal scheme (i.e. non-orthogonal wideband signals with successive decoding) can achieve achieve the total rate sum.

A second observation is that since only one user occupies each frequency band at any given time, channel estimation errors only affect the user who occupies that band. This is not the case in a system where successive decoding [16] is required since channel estimation errors add up as users signals are stripped from the received signal. Although theoretically this may not pose a problem, it is still unclear at this point whether or not successive decoding is feasible in practical receivers.

Perhaps the most important feature of the optimal power allocation scheme is that it is a form of diversity akin to selection diversity[18] with either multiple transmit or multiple receive antennae. As mentioned in the introduction, we refer to it as *multiuser diversity.* Unlike antenna diversity, however, it does not require any additional resources and may prove to be a simple method for achieving high data rates in a multiuser system.

The optimal power allocation strategy includes standard water-filling in addition to multiuser diversity. We will see that, at least for Ricean/Rayleigh fading statistics, that the water–filling power adjustment plays a rather insignificant role, as is the case for single–user channels [14].

2.5 Two-User Capacity Region

For the purpose of illustration, we have computed the two–user capacity region in Rayleigh fading. The optimal input spectra can be expressed using either the optimization techniques described in Tse and Hanly [3] or Cheng and Verdu [35] which allow the boundary region to be computed numerically.

Let us denote the average channel strength for user m by A_m for a particular transmission and assume unit-variance noise (so that A_m is the average signal-to-noise ratio for user m. In Figure 3(a) we show a symmetric two-user capacity region $A_1 = 10, A_2 = 10$ and in Figure 3(b) a non-symmetric region $A_1 = 15, A_2 = 5$ with unit-mean square Rayleigh statistics for the channel responses (i.e. $f_{|H|^2}(h) = e^{-h}$[28]). We also show the corresponding region for a non-fading channel ($f_{|H|^2}(h) = \delta(h-1)$) with the same strengths for each user. The most important thing to remark is that there is a portion on the fading channel which extends beyond that of a non-fading channel. It is only around the equal rate point, however, when we operate in a symmetric situation. We will see that this performance improvement on fading channels is due to multiuser diversity.

2.6 Ergodic Rates with Slow-Power Control

In the remainder of this work, we focus on the point of the capacity region which maximizes the total sum rate or the throughput. It is also the most likely point at which to operate in a system with *slow-power control* even with unequal rate requirements. By slow-power control we mean that the average received power, $P_{\mathrm{R},m}$ is kept at a constant value achieving the desired rate. This is achieved by amplifying the transmitted signals by the factor $P_{\mathrm{R},m}/A_m$. This type of power control is used in most cellular systems.

With unequal rates, we must assume that higher–rate users are willing to pay more in terms of transmit power. This would also be the case in a slotted system (such as TDMA or FDMA) where high-rate users can occupy more than one slot.

With slow power control, we may optimize the input power spectra now with a constraint on the average received power. Without loss of generality we may write the total rate sum as

$$\sum_{m\in\mathcal{S}} R_m \leq \frac{1}{2}\mathrm{E}\int_{-W}^{W} \log_2\left(1+\frac{2}{N_0}\sum_{m\in\mathcal{S}} P_{\mathrm{R},m}|H_m(f)|^2 S_m(f)\right) df \;\; \text{bits/s} \tag{25}$$

with both $|H_m(f)|^2$ and $S_m(f)$ having unit average power and energy respectively.

2.7 Numerical Results for Different Fast Fading Distributions

Using the normalization $\mu_m = \lambda_m \frac{WN_0}{P_{\mathrm{R},m}}$, it follows that with multiuser diversity the rates of the users are given by

$$R_m^{\mathrm{MD}} = W\int_{\mu_m}^{\infty} \log_2\left(\frac{\gamma}{\mu_m}\right) \prod_{m'\neq m} F_{|H|^2}\left(\frac{\mu_{m'}}{\mu_m}\gamma\right) f_{|H|^2}(\gamma)d\gamma \;\; \text{bits/s} \tag{26}$$

and that the average received signal–to–noise ratio satisfies

$$\frac{P_{R,m}}{WN_0} = \int_{\mu_m}^{\infty}\left(\frac{1}{\mu_m}-\frac{1}{\gamma}\right)\prod_{m'\neq m} F_{|H|^2}\left(\frac{\mu_{m'}}{\mu_m}\gamma\right) f_{|H|^2}(\gamma)d\gamma \tag{27}$$

where $f_{|H|^2}(\gamma)$ and $F_{|H|^2}(\gamma)$ are the p.d.f. and c.d.f of the unit mean fast-fading power. We will consider two types of distributions, namely *Ricean* fading with specular-to-diffuse power ratio K and L–branch diversity combining with uncorrelated equal–mean Rayleigh fading on each branch. The corresponding p.d.f.'s are

$$f_{|H|^2}^{\mathrm{Rice}}(\gamma) = (1+K)\exp(-K(1+(1+1/K)\gamma))\mathrm{I}_0\left(\sqrt{\gamma K(1+K)}\right), \;\; \gamma\geq 0 \tag{28}$$

and

$$f_{|H|^2}^{\text{Diversity}}(\gamma) = \frac{1}{L!}\left(\frac{\gamma}{L}\right)^{L-1} e^{-\gamma/L}, \;\; \gamma \geq 0 \tag{29}$$

We have chosen these two cases to illustrate the effect of the variance of the power distribution on the achievable rates. The diversity scheme would correspond to either a multi-antenna receiver or some form of spread-spectrum signaling with RAKE reception [28]. In the latter case, this would assume that the L multipath components are independent and of equal strength, which is very optimistic. For Ricean fading (i.e. $K > 0$), there are no closed–form expressions for the achievable rates, and they must be computed numerically. This is also true for the Rayleigh fading case with $L > 1$, but for $L = 1$ and $P_{\mathrm{R},m} = P_{\mathrm{R}}, \forall m$ we obtain the simple closed-form expressions (see Appendix B) for the rate of each user and received SNR

$$R^{\mathrm{MD}} = \frac{W}{M \ln 2} \sum_{i=1}^{M} (-1)^{i-1} \binom{M}{i} \mathrm{E}_1(i\mu) \;\; \text{bits/sec} \tag{30}$$

$$\frac{P_{\mathrm{R}}}{WN_0} = \frac{1}{M} \sum_{i=1}^{M} (-1)^{i-1} \binom{M}{i} \left[\frac{1}{\mu} e^{-i\mu} - \mathrm{E}_1(i\mu)\right] \tag{31}$$

where $\mathrm{E}_n(x) = \int_1^\infty e^{-tx}/t^n dt$ is the n^{th} order exponential integral [38], and μ is a non-negative parameter.

We may also consider a simpler sub-optimal scheme where the received SNR is fixed but multiuser diversity is still used. In this case

$$R^{\mathrm{MD2}} = \frac{W}{M \ln 2} \sum_{i=1}^{M} (-1)^{i-1} \binom{M}{i} \exp\left(i \frac{WN_0}{MP_{\mathrm{R}}}\right) \mathrm{E}_1\left(i \frac{WN_0}{MP_{\mathrm{R}}}\right) \;\; \text{bits/sec} \tag{32}$$

We note for $\frac{P_{\mathrm{R}}}{WN_0} \to \infty$ that (31) yields $\mu \to \frac{WN_0}{/} MP_{\mathrm{R}}$ and $R^{\mathrm{MD}} \approxeq R^{\mathrm{MD2}}$, which indicates that the water-filling component of the optimal spectral allocation policy has very little effect on the achievable rates. Moreover, in the single user case (i.e. $M = 1$) this implies that channel state information at the transmission end has very little effect on average information rates.

2.7.1 Comparison with Flat Input Spectra in unit-diversity ($L = 1$) Rayleigh Fading

We first address the issue of how much benefit can be drawn from performing optimal spectral allocation (i.e. exploiting multiuser diversity). To this end, we compare the ergodic rates of an optimal system without channel state feedback, that is one with flat input spectra. Furthermore, we will assume at the moment that the total system bandwidth is a multiple of the number of users in the system so that the bandwidth per user remains constant, say $W = MW_{\mathrm{U}}$. For a system without channel state information at the transmission end and equal average received powers we have that

$$R^{\text{NO-CSF}} = W_{\text{U}} \mathrm{E} \, \log_2 \left(1 + \frac{P_{\text{R}}}{W_{\text{U}} N_0} \frac{1}{M} \sum_{i=1}^{M} \alpha_i \right) \quad \text{bits/sec} \tag{33}$$

It is worthwhile noting that the argument of the logarithm is a central Chi-square random variable with $2M$ degrees of freedom, and that the expectation can be expressed in closed–form involving the incomplete Gamma function (see Appendix C)

$$R^{\text{NO-CSF}} = \frac{W_{\text{U}}}{\ln 2} e^{\frac{W N_0}{P_{\text{R}}}} \sum_{j=0}^{M-1} \sum_{i=0}^{M-1-j} (-1)^i \left(\frac{W N_0}{P_{\text{R}}} \right)^i \frac{\Gamma\left(j, \frac{W N_0}{P_{\text{R}}}\right)}{\Gamma(i+1)\Gamma(j+1)} \quad \text{bits/sec} \tag{34}$$

A similar expression was found in [9]. We see, however, for large M that it tends to the capacity of a non-fading AWGN channel

$$R^{\text{Gauss}} = W_{\text{U}} \log_2 \left(1 + \frac{P_R}{W_{\text{U}} N_0} \right) \quad \text{bits/sec} \tag{35}$$

The performance increase as a function of the number of users for fading channels was first noted by Gallager [34]. It stems from the fact that the total *instantaneous* received energy becomes more deterministic as the number of users increases. This diversity effect can be exploited by a multiuser receiver.

With channel state information, the achievable rate also increases with the number of users, but it is unbounded. Moreover, for $M \geq 2$ it is greater than that of a non-fading channel. For $\frac{P_{\text{R}}}{W_{\text{U}} N_0} \to \infty$ we have that the increase in spectral efficiency with respect to a non-fading channel is

$$W_{\text{U}}^{-1} (R^{\text{MD}} - R^{\text{Gauss}}) = -\frac{\gamma}{\ln 2} + \sum_{i=1}^{M} (-1)^i \binom{M}{i} \log_2 i \quad \text{bits/s/Hz} \tag{36}$$

where $\gamma = .5772$ is Euler's constant. This is found using the expansion [38][p229,5.1.11]

$$\mathrm{E}_1(z) = -\gamma - \ln z - \sum_{n=1}^{\infty} (-1)^n \frac{z^n}{n n!} \quad (|\arg z| < \pi) \tag{37}$$

The gain in spectral efficiency increases logarithmically with the number of users. It arises from the fact that the channel strength can spend a significant portion of time above its mean. Put in other words, the allocation strategy chooses the frequencies/times where the multipath components combine coherently at the receiver. A similar effect can also be obtained with a multiple-antenna transmitter by choosing the transmit antenna gains/phases to obtain a maximum signal-to-noise ratio at the receiver of interest [39]. This *beamforming* technique relies on the spatial variations of the propagation medium whereas in our case it relies on temporal/frequency variations.

The fact that the power gain and information rate are unbounded with an increasing number of users is because theoretical models (i.e. Gaussian fading) have no upper-limit on the received signal energy level. In practice this model will only break down for a very large number of users, since it is common for the strength of a fading channel to rise to as much as 20dB more than its mean. In the following section we will also show that the variance of the received power plays an important role in how much power gain can be expected.

It is worthwhile to mention that for $M = 1$, R^{MD} is nothing but the spectral efficiency of FDMA,TDMA or perfectly orthogonal frequency-hopping systems. This essentially says that the most that can be gained in terms of ergodic rates by using an optimal CDMA system without channel state feedback is .8723 bits/s/Hz. In figure 4 we show the achievable spectral efficiencies as a function of the signal-to-noise ratio per user $\frac{P_{\mathrm{B}}}{W_{\mathrm{U}}N_0}$. It is quite remarkable that at low signal-to-noise ratios, we can expect spectral efficiencies twice as high as classical approaches. This gain will carry through to the interference-limited case considered in Section 3.

2.7.2 The Effect of Received Power Variance

We show in Figures 5(a),(b) the necessary received signal-to-noise ratio to achieve 2 bits/s/Hz for each user with $M = 1, 2, 4, 8$ as the variance of the fading power p.d.f. changes. This is done by varying L in the diversity case, and K in the case of a Ricean channel. Compare this to a non-fading AWGN channel which requires 4.77 dB to achieve 2 bits/s/Hz. What is clear is that systems which have either a strong deterministic component or a high diversity order do not benefit from exploiting channel state information at the transmission end. This is reflected from the fact that the channel has less effective gain when the p.d.f. is more concentrated around its mean.

The optimal power allocation has a rather negligible effect, which allows us to conclude that the major factor which allows for higher rates/lower power consumption is multiuser diversity.

3 Rates for Interference-Limited Cellular Systems

In cellular systems, signals from adjacent cells utilizing the same carrier frequency interfere with the signals received at a particular basestation. In current systems, joint processing between basestations is not performed, so these interferers are treated as noise. We will assume that this is still the case. Studies of multiple-access schemes with joint processing using several receivers can be found in [11, 12, 9, 10].

If we consider a system where multiuser diversity is used, we remark that the allocation of the best user for a particular frequency band depends only on the channel states of each user and not the interference levels. This is because of the interference level in a particular band is same for each user. As a result, under the assumption that users transmit Gaussian signals, the optimal spectral allocation schemes remain the same. We note, however, that Gaussian signals do not necessarily maximize mutual information on such interference channels.

To simplify the analysis, we will focus on a worst case scenario. This occurs when all interferers have maximum received power at the receiver in the cell of the desired user. This is illustrated for 2-cell system without frequency-reuse in figure 6.

In general we may consider systems with I interfering cells yielding in the worst case MI equal strength interferers with attenuation A with respect to the desired user. With interference and equal received powers (25) generalizes to

$$\sum_{m\in\mathcal{S}} R_m \leq \frac{1}{2}\mathrm{E}\ \int_{-W}^{W} \log_2\left(1+\frac{P_{\mathrm{R}}}{N_0/2+AP_{\mathrm{R}}\sum_{i=1}^{I}\sum_{m=1}^{M}|H_{m,i}(f)|^2 S_{m,i}(f)}\right.$$
$$\left.\sum_{m\in\mathcal{S}}|H_{m,0}(f)|^2 S_{m,0}(f)\right) df\ \text{ bits/s} \tag{38}$$

where $H_{m,0}(f)$,$S_{m,0}(f)$ are the channel response and power spectrum of user m in the reference cell and $H_{m,i}(f)$, $S_{m,i}(f)$ are those in the $i=1,\cdots,I$ interfering cells.

We now compare the maximal symmetric rate again for different scenarios. First let us examine a wideband system without fading where $|H_{m,i}(f)|=1$ and $S_{m,i}(f)=\frac{1}{2W}\mathcal{I}(f\in[-W,W])$. With multiuser detection for users inside the cell we obtain

$$R^{\text{Gauss1}} = \frac{W}{M}\log_2\left(1+\frac{MP_{\mathrm{R}}}{WN_0+IAMP_{\mathrm{R}}}\right)\ \text{ bits/s} \tag{39}$$

and with single-user detection

$$R^{\text{Gauss2}} = \frac{W}{M}\log_2\left(1+\frac{P_{\mathrm{R}}}{WN_0+((IA+1)M-1)P_{\mathrm{R}}}\right)\ \text{ bits/s} \tag{40}$$

For an optimal spread-spectrum system with multiuser detection in Rayleigh fading we obtain

$$\begin{aligned} R^{\text{NO}-\text{CSF1}} &= \frac{W}{M\ln 2}\mathrm{E}\ \log_2\left(1+\frac{P_{\mathrm{R}}}{WN_0+IAMP_{\mathrm{R}}\beta}\sum_{i=1}^{M}h_i\right)\ \text{ bits/s} \\ &= \frac{W}{M\ln 2}\int_0^{\infty}\sum_{m=0}^{M-1}\sum_{i=0}^{M-1-j}(-1)^i\left(\frac{WN_0}{P_{\mathrm{R}}}+IAM\beta\right)^i \\ &\qquad \exp\left(\frac{WN_0}{P_{\mathrm{R}}}+(IAM-1)\beta\right)\frac{\Gamma\left(m,\frac{WN_0}{P_{\mathrm{R}}}+IAM\beta\right)}{\Gamma(i+1)\Gamma(j+1)}d\beta\ \text{ bits/s} \end{aligned} \tag{41}$$

where β and h_i are unit mean exponential random variables. For $P/WN_0 \to \infty$ we may use [40][p.663,6.455] yielding

$$R^{\mathrm{NO-CSF1}} = \frac{W}{M \ln 2} \sum_{j=0}^{M-1} \sum_{i=1}^{M-j} (-1)^i (IAM)^{i+j} \frac{\Gamma(i+j)}{\Gamma(i+2)\Gamma(j+1)} F(1, i+j+1; i+2; 1-IAM) \text{ bits/sec} \tag{42}$$

where $F(\alpha, \beta; \gamma; z)$ is the *hypergeometric series* [40][p. 1039,9.100]. With single-user detection we obtain

$$\begin{aligned} R^{\mathrm{NO-CSF2}} &= \frac{W}{M \ln 2} \mathrm{E} \log_2 \left(1 + \frac{P_{\mathrm{R}}}{WN_0 + (M(IA+1)-1)P_{\mathrm{R}}\beta} h\right) \text{ bits/s} \\ &= \frac{W}{M \ln 2} \int_0^\infty \exp\left(\frac{WN_0}{P_{\mathrm{R}}} + (M(IA+1)-1)\beta\right) e^{-\beta} \\ &\qquad \mathrm{E}_1 \left(\frac{WN_0}{P_{\mathrm{R}}} + (M(IA+1)-1)\beta\right) d\beta \text{ bits/sec} \end{aligned} \tag{43}$$

where h and β are unit-mean exponential random variables. For $P/WN_0 \to \infty$ we may again use [38][p.230,5.1.34] yielding

$$R^{\mathrm{NO-CSF2}} = \frac{W}{M} \frac{1}{M(IA+1)-2} \log_2 (M(IA+1)-1) \quad \text{bits/s.} \tag{44}$$

In the sub-optimal multiuser diversity case where a constant transmit power is used (32) generalizes to

$$R^{\mathrm{MD2}} = \frac{W}{M \ln 2} \sum_{i=1}^{M} (-1)^{i-1} \binom{M}{i} \int_0^\infty \exp\left\{i \frac{WN_0}{MP_{\mathrm{R}}} + (IAi-1)\beta\right\} \mathrm{E}_1 \left(i\left[\frac{WN_0}{MP_{\mathrm{R}}} + IA\beta\right]\right) d\beta \text{ bits/sec.} \tag{45}$$

Again for $P_{\mathrm{R}}/WN_0 \to \infty$ we may use [38][p. 230,5.1.34] yielding

$$R^{\mathrm{MD2}} = \frac{W}{M} \sum_{i=1}^{M} (-1)^{i-1} \binom{M}{i} \frac{\log_2(iIA)}{iIA-1} \quad \text{bits/s} \tag{46}$$

We have not considered the optimal water-filling power control since this will cause a tail biting effect. This is due to the fact that the power at each frequency depends on the interference level which, in turn, depends on the signal power.

3.1 Discussion

We plot the worst-case spectral efficiencies for noise free interference channels in Figure 7 for $IA = 1, 2, 4, 6$. Notice that we have nearly a twofold increase in spectral efficiency over classical approaches for a reasonable number of users, which would double system capacity of interference-limited systems. As pointed out in [9, 10] higher data rates can be expected with fading interference channels, since the interference levels have a high probability of being weak. This is not the case for non-fading channels. This "interference-diversity" effect also carries through in our case, resulting in the significant spectral efficiency increase with the added gain afforded by multiuser diversity. We note also that although optimal spread-spectrum signals ($R^{\mathrm{NO-CSI}}$) have a higher throughput capability than orthogonal signals ($M = 1, R^{\mathrm{NO-CSI}}$), the difference is rather small.

Let us now examine the effect of using frequency-reuse strategies. We will assume that groups of 16 users are multiplexed using multiuser diversity. Assuming a signal–to–noise ratio of $P^{\mathrm{R}}/W_U N_0 = 10\mathrm{dB}$, we could expect at most a spectral efficiency of 5 bits/s/Hz (see Figure 4) under the assumption that interference is negligible with frequency reuse. With 2 cells this becomes 2.5 bits/s/Hz, which is comparable to the rates without frequency reuse ($IA = 1$). Examining the case with 6 interfering cells (i.e. hexagonal coverage), we require at least a reuse factor of 3, yielding 1.67 bits/s/Hz, which is noticeably higher than the case for $IA = 6$, 1.2 bits/s/Hz. These results are somewhat different from the case where channel state information is unavailable at the transmitter and delay constraints are imposed [29].

4 Practical Considerations

The optimal spectral allocation would surely be hard to achieve in practice. Nevertheless, consider a system with many narrowband frequency slots, something along the lines of the DECT[7] or PHS systems[8], or even multiuser OFDM. Here, *narrowband* means that the channel responses in each subband are virtually flat (i.e. the bandwidth of the subbands is much less than the coherence bandwidth of the channel.) We may then interpret the optimal spectral shaping as *gaining or losing narrowband channels in time* as the strengths in each subband vary. This system is still wideband since the users have the potential of transmitting anywhere within the system bandwidth. Some users generally occupy a larger portion of the frequency band than the others, on a short time-span. On average, however, users will share subbands equally provided enough variation in time and/or frequency is present in the system.

This frequency-flat subband approximation will not incur a significant performance loss, if the subbands are sufficiently narrow. Moreover, if this is not the case, the subband frequency response will be characterized by more than one significant degree of freedom which will result in a smaller dynamic range of the total received energy per block. We will clearly suffer a similar reduction in the amplification effect discussed previously.

One of the main drawbacks of the optimal system is that we must wait for the channels to change significantly in time for the rates to converge to their average value in (25). This assumes, of course, that the bandwidth of the system is not large enough for this to occur without the need for time variation. A sub-optimal scheme where the processing delay is reduced is considered in

[41].

Although not implied by our results, a practical system employing multiuser diversity will be variable-rate, where a user's instantaneous data rate will depend on the number of channels that are allocated to him. Moreover, the amount of time a user is allocated a particular rate will depend solely on the time-variation of the channels. In delay-sensitive systems this will clearly pose a problem.

To get an idea of the rate variation we may consider the probability of transmitting at a given rate. Assume that the system bandwidth is divided into B subbands and that a coherence-bandwidth spans B_c subbands. There are therefore $N = B/B_c$ approximately independent fading strengths for each user, but we will take this to be the case exactly. The probability that any user occupies a set of B_c contiguous subbands is $1/M$. It follows that the number of subbands occupied by any user (i.e. his rate) is N times a Bernulli random variable with success probability $1/M$. This means that a complete outage is unlikely (exponential in N/M) in high data-rate applications (i.e. many coherence bandwidths per user) where the number of channels must be large compared to the number of users.

4.1 Fixed-Rate Coding in Each Subband

The rates considered in (25) are achieved by coding over a long time-frame (i.e. with interleaving) with *single-user codes.* This can be done using either bit interleaving with traditional fixed-rate error-control codes [29] or using variable-rate techniques [14]. Both are completely equivalent provided the block length in the variable-rate scheme is long enough (which is a central assumption in this work) to achieve a chosen code rate. The depth of the interleaver in the fixed-rate scheme will be roughly equal to the amount of time needed for the time-average rate of the variable–rate scheme to converge to (25). An alternate sub-optimal approach would be to vary the transmit power to achieve a desired fixed-rate in each allocated subband. This is desirable since the channel becomes an AWGN channel and stadard coding techniques can be employed. We will consider this option first.

4.1.1 Fixed Received Signal-to-Noise Ratio

We now examine a system which employs multiuser diversity and keeps the *instantaneous* received signal–to–noise ratio constant while transmitting (i.e. no water–filling). If we denote the signal attenuation for each user in a particular subband by h_i, the attenuation for the user currently active is $h = \max\{h_1, h_2, \cdots, h_M\}$. We employ a perfect power controller in the transmitter

$$P_T(h) = P_{\mathrm{R}} \frac{K(M)}{h} \tag{47}$$

where we define the factor $K(M) = \mathrm{E}\ 1/h$ which is the gain/loss in transmit power using multiuser diversity with respect to a non–fading channel with received power P_{R}. For Rayleigh fading with multiuser (selection) diversity we have [40]

$$
\begin{aligned}
K(M) &= \left(\int_0^\infty M e^{-u}(1-e^{-u})^{M-1}\frac{du}{u}\right)^{-1} \\
&= \left(\int_0^1 M(-1)^M (u-1)^{M-1}\frac{du}{\ln u}\right)^{-1} \\
&= \left(\sum_{m=1}^{M}(-1)^m \binom{M}{m} m\ln m\right)^{-1}, \quad M>1.
\end{aligned} \tag{48}
$$

which we plot in figure 8. We see that the gain increases logarithmically with M (as in the optimal scheme) and that the difference from the optimal scheme (fig. 4) for $M > 4$ is small.

The main advantage of using this type of power control is that any AWGN coding scheme can be employed directly with well-defined performance since the received SNR is kept constant. The price to be paid is reflected in the *peak-to-average* power ratio, $P_{\max}$. In inverting the channel response, we run the risk of driving our transmitter power to a level which saturates the transmit amplifier. Typically, a system will have some upper-limit on $P_{\max}$ which must be respected. In a cellular system, this limit will become important for users on the edge of a cell where the transmit amplifiers are running at full power. With such a limitation we can naturally define the outage probability $P_{\mathrm{out}}(M, P_{\max}) = \mathrm{Prob}\left(h \le \frac{K(M)}{P_{\max}}\right)$ which in unit-mean square Rayleigh fading is

$$
P_{\mathrm{out}}(M, P_{\max}) = \left(1 - e^{-\frac{K(M)}{P_{\max}}}\right) \tag{49}
$$

If the probability of error of the system when the transmit power is less than $P_{\max}$ is larger than this quantity then the effect of peak-power outages will be negligible. We will examine this in the following section.

4.1.2 Bit Error Rates with Multiuser Diversity

As an example of using multiuser diversity with common signaling schemes, let us consider uncoded QPSK with and without power control. When no power control is employed, we have that the error probability conditioned on the maximum channel state h is [28]

$$
P_{b|h} = \mathrm{Q}\left(\sqrt{2h\frac{\mathcal{E}_b}{N_0}}\right) \tag{50}
$$

where $\mathcal{E}_b$ is the energy per information bit. Averaging over h yields (see Appendix B)

$$
P_b = \mathrm{E}_h P_{b|h} = \frac{1}{2}\sum_{m=1}^{M}(-1)^{m-1}\binom{M}{i}\left(1 - \frac{1}{\sqrt{1 + i\frac{N_0}{\mathcal{E}_b}}}\right) \tag{51}
$$

Some BER vs. SNR curves are plotted in figure 9 where we see the familiar diversity behaviour. We have performance superior to that of a non–fading channel for $M \geq 8$ at practical SNR.

We stress that care must be taken when interpreting these results. The SNR $\mathcal{E}_b/N_0$ is the average received SNR of a particular user taken over the *entire* system bandwidth and processing time-span. The actual received SNR will be significantly higher, with high probability, in the frequency/time slot occupied by the user. Synchronization and channel estimation will therefore still be feasible even if the system operates at low average SNR.

For the case with power control we have the following average error probability

$$P_b = (1 - P_{\text{out}}(M, P_{\max}))\text{Q}\left(\sqrt{2K\frac{\mathcal{E}_b}{N_0}}\right) + \int_0^{K(M)/P_{\max}} \text{Q}\left(\sqrt{2u\frac{\mathcal{E}_b}{N_0}}\right) f_h(u)du \qquad (52)$$

This expression is plotted in figure 10 for a varying number of users and peak–to–average power ratios $P_{\max}$. Here we see that power control yields improved performance compared to the previous case. Moreover, for $M \geq 4$ we notice that peak–power constraints do not pose a significant problem, since the curves tail-off at very low BER.

The important conclusion to be drawn from this analysis is that channel state feedback has a dramatic effect on bit-error rates even without additional channel coding. Dynamic allocation such as this may be a simpler option than sophisticated coding schemes for achieving acceptable performance on fading multiuser channels.

5 Conclusion

This work considered power allocation schemes for maximizing the total throughput of continuous-time multiuser channels in a multipath fading environment. In order to obtain tractable numerical results, we opted to use a wideband block-fading channel model, which reflects the essential characteristics of typical wireless communication channels. Moreover, our numerical results and derivations of optimal spectral allocation policies are specialized to *ergodic* or high-diversity systems, which assume that either the decoding time-span or system bandwidth is very large.

Our main result is that the maximal system throughput (i.e. the sum of the information rates of all users) on fading multiuser channels is achieved by orthogonal signaling based on dynamic time/frequency allocation. This is made possible under the assumption of *a priori* channel state information at the transmitters. The selection policy is similar to multi-antenna selection diversity and yields a similar diversity effect, in the sense that at any frequency only the user with the strongest response is permitted to transmit. We have termed this signaling scheme *multiuser diversity* and have shown that information rates approaching double that of traditional wideband signaling schemes can be expected. This is the case both on systems with and without interfering signals. Moreover, since multiuser diversity is essentially a narrowband signaling scheme, receiver complexity is significantly reduced compared to wideband systems.

We argued that a practical system employing multiuser diversity would consist of many narrowband channels, along the lines of systems such as DECT and PHS, with added intelligence in the basestations, who would instruct the users to transmit in an optimal fashion. Moreover, the system would be variable rate, in the sense that the information rate would depend on the number of occupied channels at any given time. This would clearly be a random quantity and could therefore pose problems in systems with very strict processing delay requirements. If the number of coherence bandwidths in the system bandwidth is large compared to the number of users, as will be the case in any high data-rate application, this will pose less of a problem.

Although applicable to both slowly time-varying FDD and TDD systems, it is our belief that schemes such as multiuser diversity would be most appropriate in the latter case, since channel estimation *over the entire system bandwidth* can be accomplished at the user terminals from the downlink signal.

We showed that multiuser diversity can have a dramatic effect when combined with traditional power control even without the need for additional channel coding.

A Proof of Theorem 1 and Corollaries

We first consider the upper-bound in Theorem 1. Let $E_s(\mathbf{H})$ be the event that the set $\{x_i(t), i \in S\}$ of users are decoded incorrectly, while those in the set $\{x_i(t), i \in S^c\}$ are decoded correctly, conditioned on the state of the channels. Using the standard random coding construction, the total ensemble average probability of block decoding error conditioned on the channel state $\mathbf{H}$ is

$$
\begin{aligned}
P_{\text{ens}}(\mathbf{h}) &= \Pr\left(\bigcup_{S\subseteq\{1,2,\cdots,M\}, S\neq\phi} E_s(\mathbf{h})\right) \\
&\leq \sum_{S\subseteq\{1,2,\cdots,M\}, S\neq\phi} \Pr\left(E_s(\mathbf{h})\right)
\end{aligned}
\tag{53}
$$

This is independent of the actual codeword being sent by virtue of the random coding construction. Since the decomposition in (6) is a memoryless channel for each S, we have from Gallager [36, p.136-138] that

$$
\Pr\left(E_s(\mathbf{h})\right) \leq \begin{cases} 2^{-N(T+T_G)E_r(\{R_i, i\in S\}, S, \mathbf{h}, N, T, T_G)} & I(S,\mathbf{h}) > \sum_{i\in S} R_i \\ 1 & I(S,\mathbf{h}) \leq \sum_{i\in S} R_i \end{cases}
$$

Choosing a Gaussian *a priori* distribution for $x_i(t), \forall i$ maximizes $I(S,\mathbf{h})$ for each $\mathbf{h}$ and yields (8) and (9). We obtain the upper-bound by averaging over the distribution of $\mathbf{H}$.

For the lower-bound, we write the probability of error for any code (i.e. not an ensemble average) as

$$
P_e(N,T,T_G) = \int_{\{\mathbf{h}|\mathbf{R}\notin C(\mathbf{h})\}} P_e(\mathbf{h},N,T,T_G)dF_{\mathbf{H}}(\mathbf{h}) + \int_{\{\mathbf{h}|\mathbf{R}\in C(\mathbf{h})\}} P_e(\mathbf{h},N,T,T_G)dF_{\mathbf{H}}(\mathbf{h}) \tag{54}
$$

$$
\geq \int_{\{\mathbf{h}|\mathbf{R}\notin C(\mathbf{h})\}} \left(1 - P_c(\mathbf{h},N,T,T_G)\right) dF_{\mathbf{H}}(\mathbf{h}) \tag{55}
$$

$$
= P_{\text{out}}(\mathbf{R},N,T,T_G) - \int_{\{\mathbf{h}|\mathbf{R}\notin C(\mathbf{h})\}} P_c(\mathbf{h},N,T,T_G)dF_{\mathbf{H}}(\mathbf{h}) \tag{56}
$$

$$
\geq P_{\text{out}}(\mathbf{R},N,T,T_G) - \sum_{S\subseteq\{1,2,\cdots,M\}} \int_{\{\mathbf{h}|\sum_{i\in S} R_i > C_S\}} P_c(\mathbf{h},N,T,T_G)dF_{\mathbf{H}}(\mathbf{h}) \tag{57}
$$

$$
= P_{\text{out}}(\mathbf{R},N,T,T_G)\Bigg[1 - \sum_{S\subseteq\{1,2,\cdots,M\}} \int_{\{\mathbf{h}|\sum_{i\in S} R_i > I(S,\mathbf{h})\}} P_c(\mathbf{h},N,T,T_G)dF_{\mathbf{H}|\mathbf{R}\notin C(\mathbf{h})}(\mathbf{h}|\mathbf{R}\notin C(\mathbf{h}))\Bigg] \tag{58}
$$

where $P_e(\mathbf{h},N,T,T_G)$ and $P_c(\mathbf{h},N,T,T_G)$ are the probabilities of incorrectly and correctly decoding all the users signals conditioned on the channel state $\mathbf{H}$ and $C(\mathbf{h})$ is the region defined by

$\{\mathbf{R} | \sum_{i\in S} R_i \leq I(S,\mathbf{h}), \forall S \subseteq \{1,2,\cdots,M\}\}$. Clearly $P_c(\mathbf{h},N,T,T_G)$ is upper-bounded by the probability of correctly decoding the signals in S conditioned on the signals in S^c being known by the receiver. We now prove a strong converse similar to the one given in [36, p.173-176] for decoding the users in S. We denote the decision regions for an arbitrary decoding rule by $Y_{l,S}, l = 1,\cdots,L_S$ where $L_S = 2^{N(T+T_G)\sum_{i\in S} R_i}$. The resulting probability of correct decoding is upper-bounded as

$$P_c(\mathbf{h},N,T,T_G) \leq \frac{1}{L_S}\sum_{l=1}^{L_S}\int_{\mathbf{y}\in Y_{l,S}} f_{\mathbf{Y}_S|\mathbf{U}_{l,S},\mathbf{U}_{l,S^c}}(\mathbf{y}_S|\mathbf{u}_{l,S},\mathbf{u}_{l,S^c})d\mathbf{y} \tag{59}$$

We now introduce the regions

$$B_{l,S} = \left\{\mathbf{y}_S \middle| \log_2 \frac{f_{\mathbf{Y}_S|\mathbf{U}_{l,S},\mathbf{U}_{l,S^c}}(\mathbf{y}_S|\mathbf{u}_{l,S},\mathbf{u}_{l,S^c})}{f_{\mathbf{Y}_S}(\mathbf{y}_S)} > N(T+T_G)(I(S,\mathbf{h})+\epsilon)\right\}, l = 1,\cdots,L_S \tag{60}$$

where $\epsilon > 0$ is an arbitrary constant and write

$$\begin{aligned} P_c(\mathbf{h},N,T,T_G) \leq \frac{1}{L_S}\sum_{l=1}^{L_S}\int_{\mathbf{y}\in Y_{l,S}\bigcap B_{l,S}} & f_{\mathbf{Y}_S|\mathbf{U}_{l,S},\mathbf{U}_{l,S^c}}(\mathbf{y}_S|\mathbf{u}_{l,S},\mathbf{u}_{l,S^c})d\mathbf{y} + \\ \int_{\mathbf{y}\in Y_{l,S}\bigcap B_{l,S}^c} & f_{\mathbf{Y}_S|\mathbf{U}_{l,S},\mathbf{U}_{l,S^c}}(\mathbf{y}_S|\mathbf{u}_{l,S},\mathbf{u}_{l,S^c})d\mathbf{y} \end{aligned} \tag{61}$$

We first note that

$$\int_{\mathbf{y}\in Y_{l,S}\bigcap B_{l,S}^c} f_{\mathbf{Y}_S|\mathbf{U}_{l,S},\mathbf{U}_{l,S^c}}(\mathbf{y}_S|\mathbf{u}_{l,S},\mathbf{u}_{l,S^c})d\mathbf{y} \leq 2^{-N(T+T_G)[\sum_{i\in S} R_i - I(S,\mathbf{h})-\epsilon]} \tag{62}$$

and second that

$$\begin{aligned} &\int_{\mathbf{y}\in Y_{l,S}\bigcap B_{l,S}} f_{\mathbf{Y}_S|\mathbf{U}_{l,S},\mathbf{U}_{l,S^c}}(\mathbf{y}_S|\mathbf{u}_{l,S},\mathbf{u}_{l,S^c})d\mathbf{y} \\ &\qquad \leq \int_{\mathbf{y}\in B_{l,S}} f_{\mathbf{Y}_S|\mathbf{U}_{l,S},\mathbf{U}_{l,S^c}}(\mathbf{y}_S|\mathbf{u}_{l,S},\mathbf{u}_{l,S^c})d\mathbf{y} \\ &\qquad = \Pr\left(\log_2 \frac{f_{\mathbf{Y}_S|\mathbf{U}_{l,S},\mathbf{U}_{l,S^c}}(\mathbf{y}_S|\mathbf{u}_{l,S},\mathbf{u}_{l,S^c})}{f_{\mathbf{Y}_S}(\mathbf{y}_S)} > N(T+T_G)(I(S,\mathbf{h})+\epsilon)\middle| \mathbf{u}_{S,m},\mathbf{u}_{S^c}\right) \end{aligned} \tag{63}$$

Denoting $z_{S,l} = \sum_{n=1}^{N} \sum_{i=1}^{\infty} \frac{(y_{n,i,S} - u_{n,i,S^c})^2}{N_0 + 2\lambda_{S,n,i}} - \frac{(y_{S,n,i} - u_{S,l,n,i} - u_{S^c,n,i})^2}{N_0}$ we have

$$
\begin{aligned}
P_c(\mathbf{h}, N, T, T_G) &\leq \frac{1}{L_S} \sum_{l=1}^{L_S} \Pr\left(z_{S,l} > N(T+T_G)\epsilon \,|\mathbf{u}_{S,m}, \mathbf{u}_{S^c}\right) \\
&\leq \frac{1}{N(T+T_G)\epsilon L_S} \sum_{l=1}^{L_S} \overline{z_{S,l}} \qquad (64) \\
&= \frac{1}{N(T+T_G)\epsilon} \sum_{n=1}^{N} \sum_{i=1}^{\infty} \frac{\mu_{n,i,S}}{N_0 + 2\lambda_{n,i,S}} \qquad (65)
\end{aligned}
$$

where $\mu_{n,i,S} = \frac{1}{L_S} \sum_{l=1}^{L_S} u_{l,n,i,S}^2 - \lambda_{n,i,S}$, which tends to zero with increasing $N(T+T_G)$. (64) follows from the Markov inequality. Letting $\epsilon = .5(\sum_{i \in S} R_i - I(S, \mathbf{h}))$ and replacing (62) and (65) in (58) yields the lower bound in (7) with

$$
\begin{aligned}
K(\mathbf{R}, N, T, T_G) = 1 - \sum_{S \subseteq \{1,2,\cdots,M\}} \int_{\{\mathbf{h} | \sum_{i \in S} R_i > I(S,\mathbf{h})\}} \Bigg[2^{-\frac{N(T+T_G)}{2}(\sum_{i \in S} R_i - I(S,\mathbf{h}))} \\
+ \frac{2}{N(T+T_G)(\sum_{i \in S} R_i - I(S,\mathbf{h}))} \sum_{n=1}^{N} \sum_{i=1}^{\infty} \frac{\mu_{n,i,S}}{N_0 + 2\lambda_{n,i,S}} \Bigg] dF_{\mathbf{H}|\mathbf{R} \notin C(\mathbf{h})}(\mathbf{h}|\mathbf{R} \notin C(\mathbf{h})). \qquad (66)
\end{aligned}
$$

Corollary 1 follows from the weak law of large numbers. First assume that $\sum_{i \in S} R_i - \mathrm{E} I(S, \mathbf{h}) = \epsilon > 0$ for some $S \in \{1, 2, \cdots, M\}$. We have, therefore, that

$$
\begin{aligned}
P_{\text{out}}(\mathbf{R}, S, N, T, T_G) &= \Pr\left(I(S,\mathbf{h}) - \mathrm{E}\, I(S,\mathbf{h}) \leq \epsilon\right) \\
&= 1 - \Pr\left(I(S,\mathbf{h}) - \mathrm{E}\, I(S,\mathbf{h}) > \epsilon\right) \\
&\geq 1 - \Pr\left(|I(S,\mathbf{h}) - \mathrm{E}\, I(S,\mathbf{h})| > \epsilon\right) = 1 - \delta(N)
\end{aligned}
$$

Similarly when $\sum_{i \in S} R_i - \mathrm{E} I(S, \mathbf{h}) = \epsilon < 0$ we have

$$
\begin{aligned}
P_{\text{out}}(\mathbf{R}, S, N, T, T_G) &= \Pr\left(I(S,\mathbf{h}) - \mathrm{E}\, I(S,\mathbf{h}) < -\epsilon\right) \\
&\leq \Pr\left(|I(S,\mathbf{h}) - \mathrm{E}\, I(S,\mathbf{h}| > \epsilon\right) = \delta(N) \qquad (67)
\end{aligned}
$$

In the limit of large N, the weak law gives us $\delta(N) \to 0$ and $P_{\text{out}}(\mathbf{R}, S, N, T, T_G) \to \mathcal{I}\left(\sum_{i \in S} R_i < \mathrm{E}\, I(S, \mathbf{h})\right)$. The first limit of corollary 1 follows from the fact that $P_{\text{out}}(\mathbf{R}, N, T, T_G) \geq \max_{S \subseteq \{1,2,\cdots,M\}} P_{\text{out}}(\mathbf{R}, S, N, T, T_G)$, and the second because of the minimization.

Corollary 3 follows by directly applying the Szegö eigenvalue distribution theorem (c.f. [36, p.416]).

B Derivation of $R^{\mathrm{MD}}, R^{\mathrm{MD2}}$

In unit-diversity ($L = 1$) Rayleigh fading, $f_{|H|^2}(u) = e^{-u}\mathcal{I}(u > 0)$ and $F_{|H|^2}(u) = 1 - f_{|H|^2}(u)$. For $P_{\mathrm{R},m} = P_{\mathrm{R}}, \forall m$, we have that $\mu_m = \mu, \forall m$. Inserting these in (26) yields

$$\begin{aligned}
R^{\mathrm{MD}} &= W \int_{\mu}^{\infty} \log_2\left(\frac{\gamma}{\mu}\right)(1 - e^{-\gamma})^{M-1} e^{-\gamma} d\gamma \ \text{ bits/s} \\
&= \frac{W}{\ln 2} \sum_{i=0}^{M-1} (-1)^i \binom{M-1}{i} \int_{\mu}^{\infty} \ln\left(\frac{\gamma}{\mu}\right) e^{-(i+1)\gamma} d\gamma \ \text{ bits/s} \\
&= \frac{W}{\ln 2} \sum_{i=0}^{M-1} (-1)^i \binom{M-1}{i} \mathrm{E}_1\left((i+1)\mu\right) \ \text{ bits/s} \\
&= \frac{W}{M \ln 2} \sum_{i=1}^{M} (-1)^{i-1} \binom{M}{i} \mathrm{E}_1\left(i\mu\right) \ \text{ bits/s}
\end{aligned} \tag{68}$$

Similarly for the case when constant power is allocated to each user with the maximum channel strength we obtain

$$\begin{aligned}
R^{\mathrm{MD2}} &= W \int_{0}^{\infty} \log_2\left(1 + \frac{P_R}{W N_0}\gamma\right)(1 - e^{-\gamma})^{M-1} e^{-\gamma} d\gamma \ \text{ bits/s} \\
&= \frac{W}{M \ln 2} \sum_{m=1}^{M} (-1)^{i-1} \binom{M}{i} e^{i\frac{W N_0}{M P_{\mathrm{R}}}} \mathrm{E}_1\left(i \frac{W N_0}{M P_{\mathrm{R}}}\right) \ \text{ bits/s}
\end{aligned} \tag{69}$$

Note that since each user occupies each frequency a fraction $1/M$ of time, it must use power M times as much power.

C Derivation of $R^{\mathrm{NO-CSF}}$

Defining $x = \sum_{i=1}^{M} \alpha_i$ in (33) we have that its p.d.f. for Rayleigh fading is

$$f_X(x) = \frac{1}{\Gamma(M)} x^{M-1} e^{-x} \mathcal{I}(x > 0) \tag{70}$$

It follows that

$$\begin{aligned}
R^{\text{NO-CSF}} &= \frac{W_U}{\Gamma(M)\ln 2}\int_0^\infty \ln\left(1+\frac{P_{\text{R}}}{WN_0}x\right)x^{M-1}e^{-x}dx \ \ \text{bits/s} \\
&= \frac{W_U \exp\left(\frac{WN_0}{P_{\text{R}}}\right)}{\Gamma(M)\ln 2}\int_{WN_0/P_{\text{R}}}^\infty \ln\left(\frac{P_{\text{R}}}{WN_0}u\right)\left(u-\frac{WN_0}{P_{\text{R}}}\right)^{M-1}e^{-u}du \ \ \text{bits/s} \\
&= \frac{W_U \exp\left(\frac{WN_0}{P_{\text{R}}}\right)}{\Gamma(M)\ln 2}\sum_{i=0}^{M-1}(-1)^{M-1-i}\left(\frac{WN_0}{P_{\text{R}}}\right)^{M-1-i}\binom{M-1}{i}\underbrace{\int_{WN_0/P_{\text{R}}}^\infty \ln\left(\frac{P_{\text{R}}}{WN_0}u\right)u^i e^{-u}du}_{I_i} \ \ \text{bits/s}
\end{aligned}$$

Using integration by parts we have that

$$\begin{aligned}
I_i &= \int_{WN_0/P_{\text{R}}}^\infty e^{-u}u^{i-1}du + i\int_{WN_0/P_{\text{R}}}^\infty \ln\left(\frac{P_{\text{R}}}{WN_0}u\right)u^{i-1}e^{-u}du \\
&= \Gamma\left(i,\frac{WN_0}{P_{\text{R}}}\right)+iI_{i-1} \\
&= \sum_{j=0}^{i}\frac{\Gamma(i+1)}{\Gamma(j+1)}\Gamma\left(j,\frac{WN_0}{P_{\text{R}}}\right) \qquad (71)
\end{aligned}$$

which by interchanging the order of summation yields (34).

D Uncoded Bit-Error-Rates

When a user transmits, the channel gain is the maximum of M exponential random variables so that its p.d.f. is $f_{|h|^2}b(u) = Me^{-u}(1-e^{-u})^{M-1}$. The bit error-rate with QPSK is therefore

$$\begin{aligned}
P_b &= \int_0^\infty M\mathrm{Q}\left(\sqrt{2h\frac{\mathcal{E}_b}{N_0}}\right)(1-e^{-h})^{M-1}e^{-h}dh \\
&= M\sum_{i=0}^{M-1}(-1)^i\binom{M-1}{i}\int_0^\infty e^{-(i+1)h}\mathrm{Q}\left(\sqrt{2h\frac{\mathcal{E}_b}{N_0}}\right)dh \\
&= \sum_{i=1}^{M}(-1)^{i-1}\binom{M}{i}\int_0^\infty ie^{-ih}\mathrm{Q}\left(\sqrt{2h\frac{\mathcal{E}_b}{N_0}}\right)dh \qquad (72)
\end{aligned}$$

Using the fact [38]

$$\int_0^\infty e^{kx}\mathrm{Q}\left(\sqrt{ax}\right)dx = \frac{1}{2}\int_0^\infty e^{-kx}\left\{1-\mathrm{erf}\sqrt{\frac{ax}{2}}\right\}dx = \frac{1}{2k}\left(1-\sqrt{\frac{1}{1+2k/a}}\right) \qquad (73)$$

yields (51)

References

[1] R. Knopp and P. Humblet, "Information Capacity and Power Control in Single–Cell Multiuser Communications," in *Proc. IEEE ICC'95, Seattle, Wa.*, June 1995.

[2] R. Knopp and P. Humblet, "Multiple–Accessing over Frequency–Selective Fading Channels," in *Proc. IEEE PIMRC'95, Toronto, Ont.*, Sept. 1995.

[3] Tse, D.N. and Hanly, S.V., "Multi–Access Fading Channels: Part I: Polymatroidal Structure, Optimal Resource Allocation and Throughput Capacities," *IEEE Transactions on Information Theory*, vol. 44, pp. 2796–2815, November 1998.

[4] Proposed EIA/TIA Interim Standard, *Wideband Spread Spectrum Digital Cellular System Dual–Mode Mobile Station–Base Station Compatibility Standard*, 1992.

[5] European Telecommunications Standards Institute, *European Digital Cellular Telecommunications System : Physical Layer on the Radio Path (GSM 05.02)*, 1990.

[6] "IEEE Personal Communications Magazine," April 1998.

[7] *DECT*.

[8] ARIB, *Personal Handy Phone System ARIB Standard, version 2*, 1995.

[9] S. Shamai (Shitz) and A. Wyner, "Information-theoretic considerations for symmetric,cellular, multiple-access fading channels - part I," *IEEE Transactions on Information Theory*, vol. 43, pp. 1877–1894, Nov. 1997.

[10] S. Shamai (Shitz) and A. Wyner, "Information-theoretic considerations for symmetric,cellular, multiple-access fading channels - part II," *IEEE Transactions on Information Theory*, vol. 43, pp. 1895–1911, Nov. 1997.

[11] A. Wyner, "Shannon–theoretic approach to a Gaussian cellular multiple–access channel," *IEEE Transactions on Information Theory*, vol. 40, Nov. 1994.

[12] S. Hanly, *Information Capacity of Radio Networks*. PhD thesis, University of Cambridge, 1994.

[13] A. Goldsmith and P. Varaiya, "Increasing Spectral Efficiency Through Power Control," in *Proceedings of the IEEE International Conference on Communications*, (Geneva), pp. 600–604, IEEE, 1993.

[14] A. Goldsmith and P. Varaiya, "Capacity of Fading Channels with Channel Side Information," *IEEE Transactions on Information Theory*, vol. 43, pp. 1986–1992, Nov. 1997.

[15] E. Biglieri, J. Proakis, and S. Shamai (Shitz), "Fading Channels: Information-Theoretic and Communications Aspects," *IEEE Transactions on Information Theory*, vol. 44, pp. 2619–2692, oct 1996.

[16] B. Rimoldi and R. Urbanke, "A rate–splitting approach to the Gaussian multiple–access channel," *IEEE Transactions on Information Theory*, vol. 42, pp. 364–375, Mar. 1996.

[17] K. Kerpez, "Radio access system with distributed antennas," *IEEE Transactions on Vehicular Technology*, vol. 45, pp. 265–275, May 1996.

[18] M. Schwartz, W. Bennett, and S. Stein, *Communication Systems and Techniques*. New York: McGraw Hill, 1966.

[19] V. Tarrokh, N. Seshadri, and A. Calderbank, "Space-Time Coding for High Data Rate Wireless Communication: Performance Analysis and Code Construction," *IEEE Transactions on Information Theory*, vol. 44, pp. 744–765, March 1998.

[20] G. Raleigh and J. Cioffi, "Spatio-Temporal Coding for Wireless Communication," *IEEE. Trans. on Communications*, vol. 46, pp. 357–366, March 1998.

[21] E. Telatar, "Capacity of Multi-Antenna Gaussian Channels," tech. rep., AT&T Bell Laboratories, 1995.

[22] R. Knopp and P. Humblet, "On Coding for Block-Fading Channels," *IEEE Trans. on Information Theory (to appear)*, 1999.

[23] R. Knopp, *Coding and Multiple-Access over Fading Channels*. PhD thesis, Swiss Federal Institute of Technology, Communications Systems Dept., EL-ECUBLENS, 1015 Lausanne, Switzerland, 1997. URL:`http://lcmwww.epfl.ch/~knopp/PUBLIC/thesis97.ps.gz`.

[24] E. Malkamäki and H. Leib, "Rate $1/n$ Convolutional Codes with Interleaving Depth of n over a Block-Fading Ricean Channel," in *Proc. IEEE Vehicular Technology Conference*, pp. 2002–2006, May 1997.

[25] E. Malkamäki and H. Leib, "Coded Diversity Schemes over Block-Fading Rayleigh Channels," in *Proc.* 6$^{\text{th}}$ *IEEE International Conference on Universal Personal Communications*, pp. 494–499, Nov. 1997.

[26] G. Caire, R. Knopp, and P. Humblet, "System Capacity of F-TDMA Cellular Systems," *IEEE Trans. on Comm.*, vol. 46, pp. 1649–1661, December 1998.

[27] E. Biglieri, D. Divsalar, P. McLane, and M. Simon, *Introduction to Trellis–Coded Modulation with Applications*. New York: MacMillan, 1991.

[28] J. Proakis, *Digital Communications*. New York: McGraw Hill, 3rd ed., 1995.

[29] G. Caire, G. Taricco, and E. Biglieri, "Bit–Interleaved Coded Modulation," *IEEE Transactions on Information Theory*, vol. 44, pp. 927–947, May 1998.

[30] R. McEliece and W. Stark, "Channels with Block Interference," *IEEE Transactions on Information Theory*, vol. 30, pp. 44–53, Jan. 1984.

[31] L. Ozarow, S. Shamai (Shitz), and A. Wyner, "Information Theoretic Considerations for Cellular Mobile Radio," *IEEE Transactions on Vehicular Technology*, vol. 43, pp. 359–378, May 1994.

[32] G. Kaplan and S. Shamai (Shitz), "Error Probabilities for the Block–Fading Gaussian Channel," *Archiv für Elektronik und Übertragungstechnik*, vol. 49, no. 4, pp. 192–205, 1995.

[33] G. Caire, G. Taricco, and E. Biglieri, "Optimum Power Control over Fading Channels," *IEEE Transactions on Information Theory*, vol. 45, pp. 1468–1489, July 1999.

[34] R. Gallager, "An Inequality on the Capacity Region of Multiaccess Multipath Channels," in *Communications and Cryptography – Two Sides of One Tapestry–* (U. Maurer, T. Mittelholzer, G. Forney, and R. Blahut, eds.), pp. 129–139, Kluwer Academic Publishers, 1994.

[35] R. Cheng and S. Verdú, "Gaussian Multiaccess Channels with Capacity Region and Multiuser Water–Filling," *IEEE Transactions on Information Theory*, vol. 39, may 1993.

[36] R. Gallager, *Information Theory and Reliable Communication.* John Wiley and Sons, 1968.

[37] T. Cover and J. Thomas, *Elements of Information Theory.* New York: John Wiley and Sons, 1991.

[38] M. Abramovitz and I. Stegun, *Handbook of Mathematical Functions.* Dover, 1965.

[39] P. Humblet, *Digital Communications and Radio Transmission.* Institut Eurecom Class Notes, 1998.

[40] I. Gradsteyn and I. Ryzhik, *Tables of Integrals, series and products.* New York: Academic Press, 1980.

[41] I. Bettesh and S. Shamai(Shitz), "A low delay algorithm for the multiple-access channel with Rayleigh fading," in *9th IEEE Int. Symp. on Personal, Indoor, and Mobile Radio Communications (PIMRC'98) (Boston, MA)*, Sept. 8-11 1998.

List of Figures

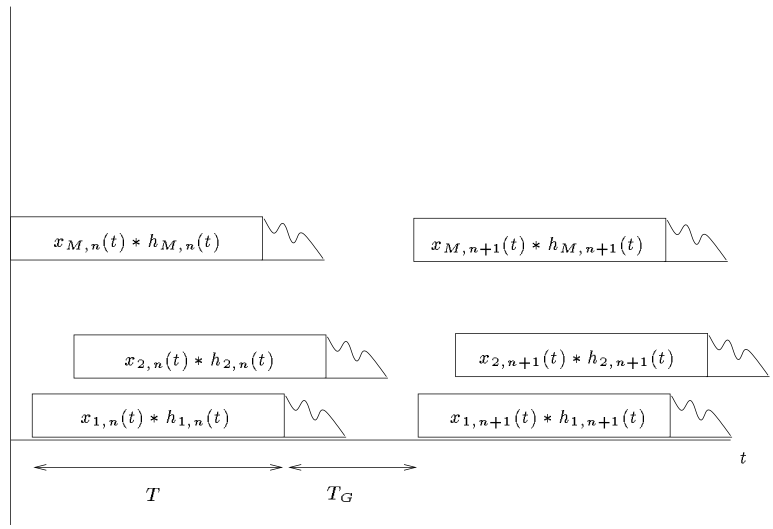


Figure 1: System Model

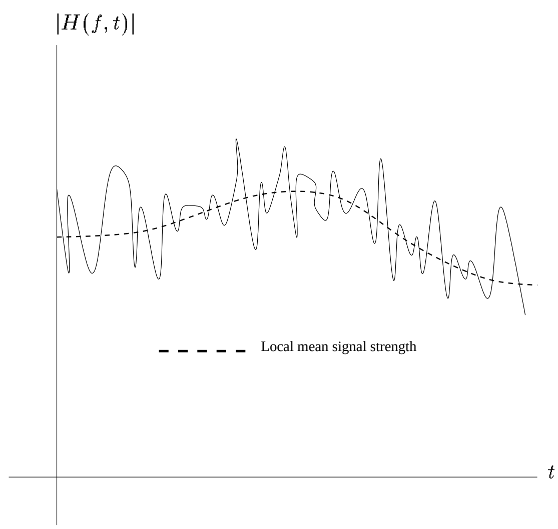


Figure 2: Slow and Fast Fading

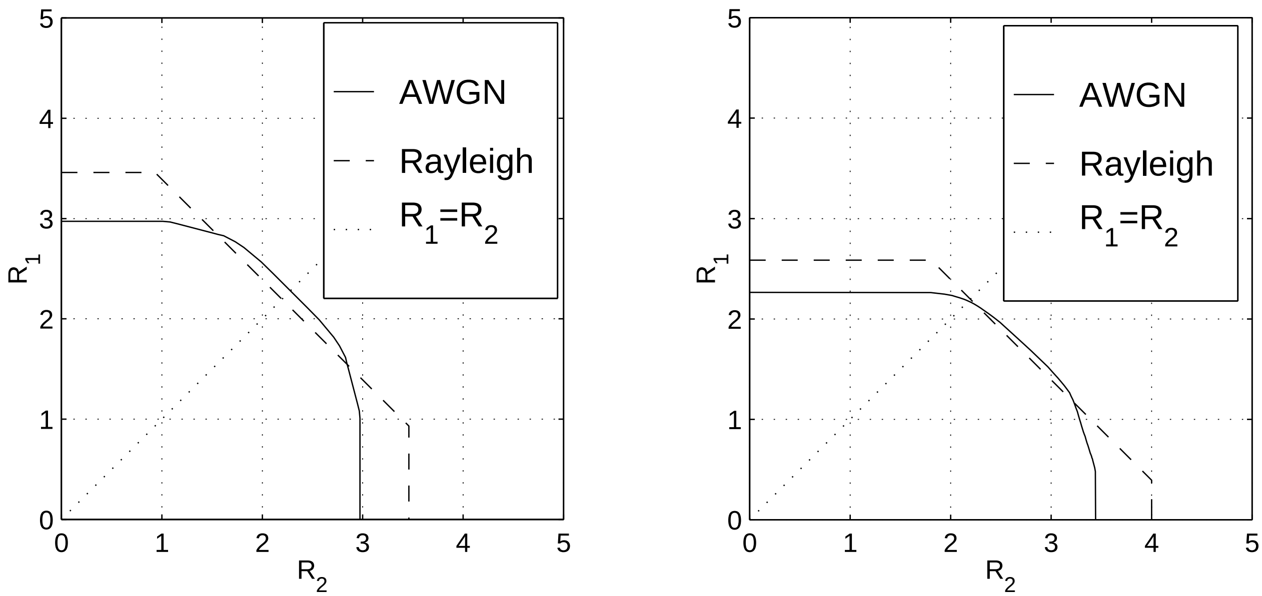


Figure 3: Two user capacity Regions in Rayleigh Fading

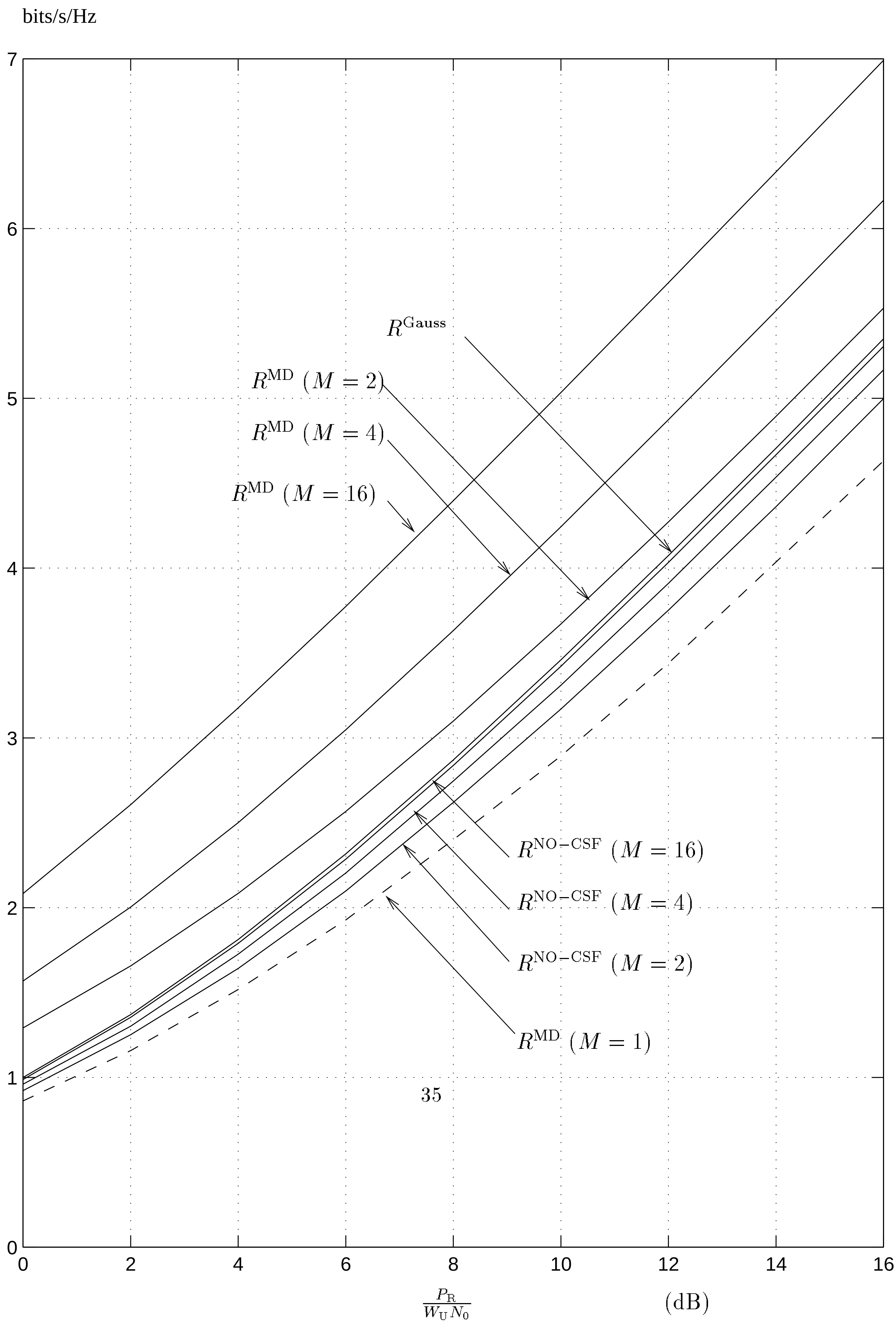
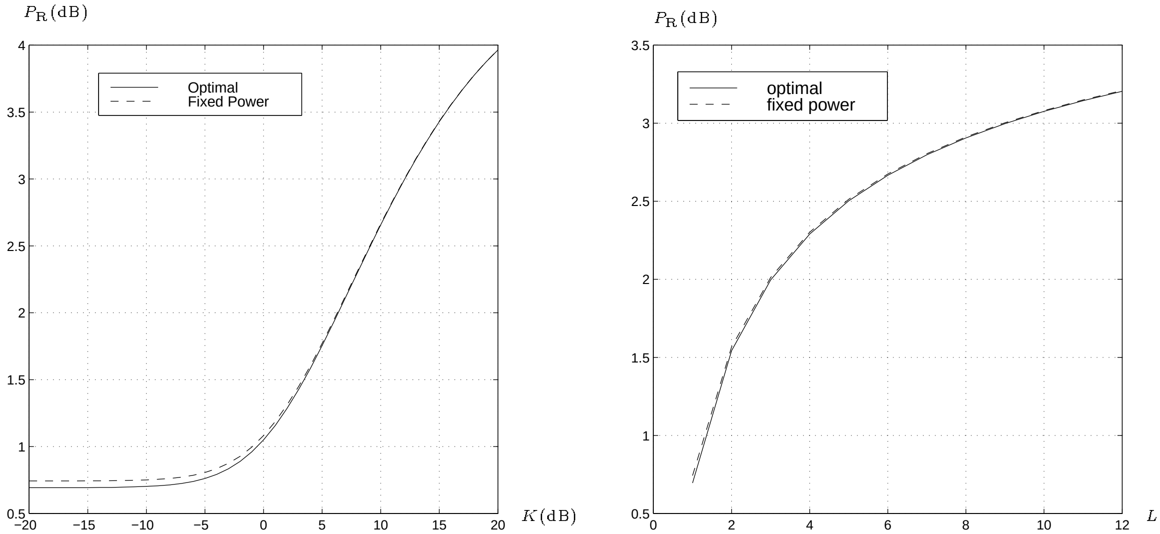


Figure 4: Comparison of Maximal Multiuser Information Rates (Rayleigh Fading)



(a)

(b)

Figure 5: Effects of Received Power Variance

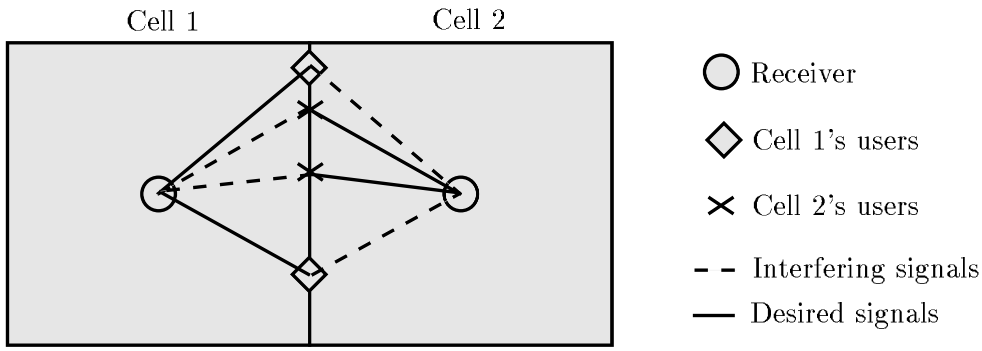


Figure 6: Cellular Interference Model

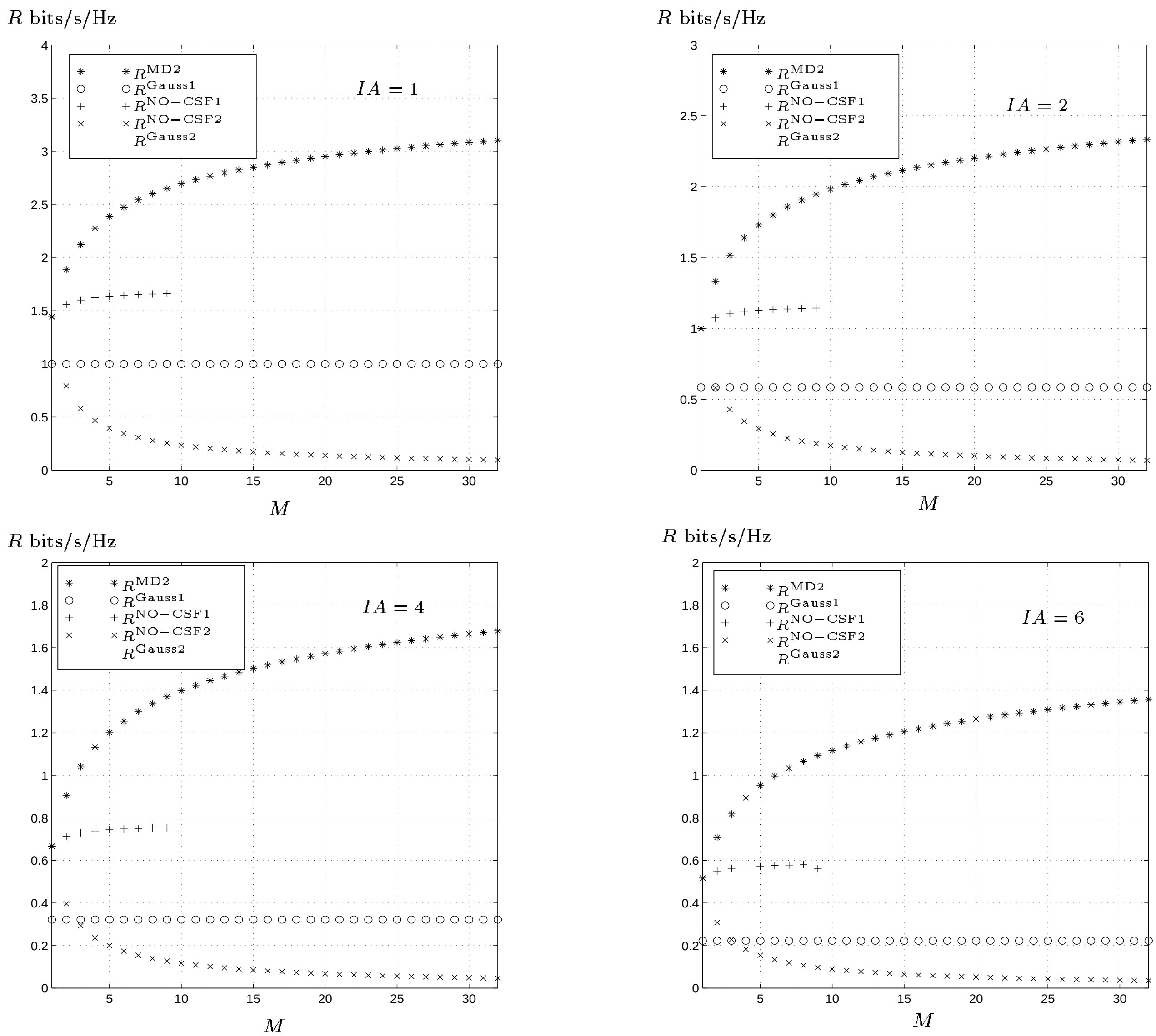


Figure 7: Achievable Rates on Interference-Limited Channels

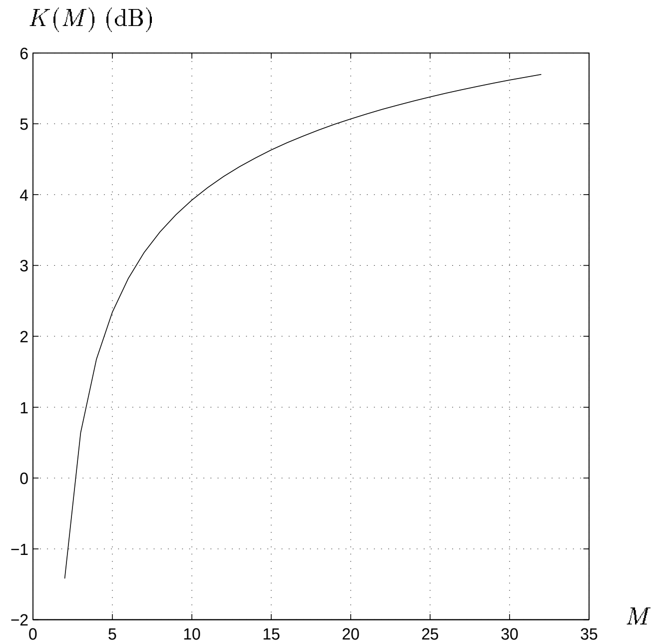


Figure 8: Multiuser Diversity Gain over Non–Fading Channel

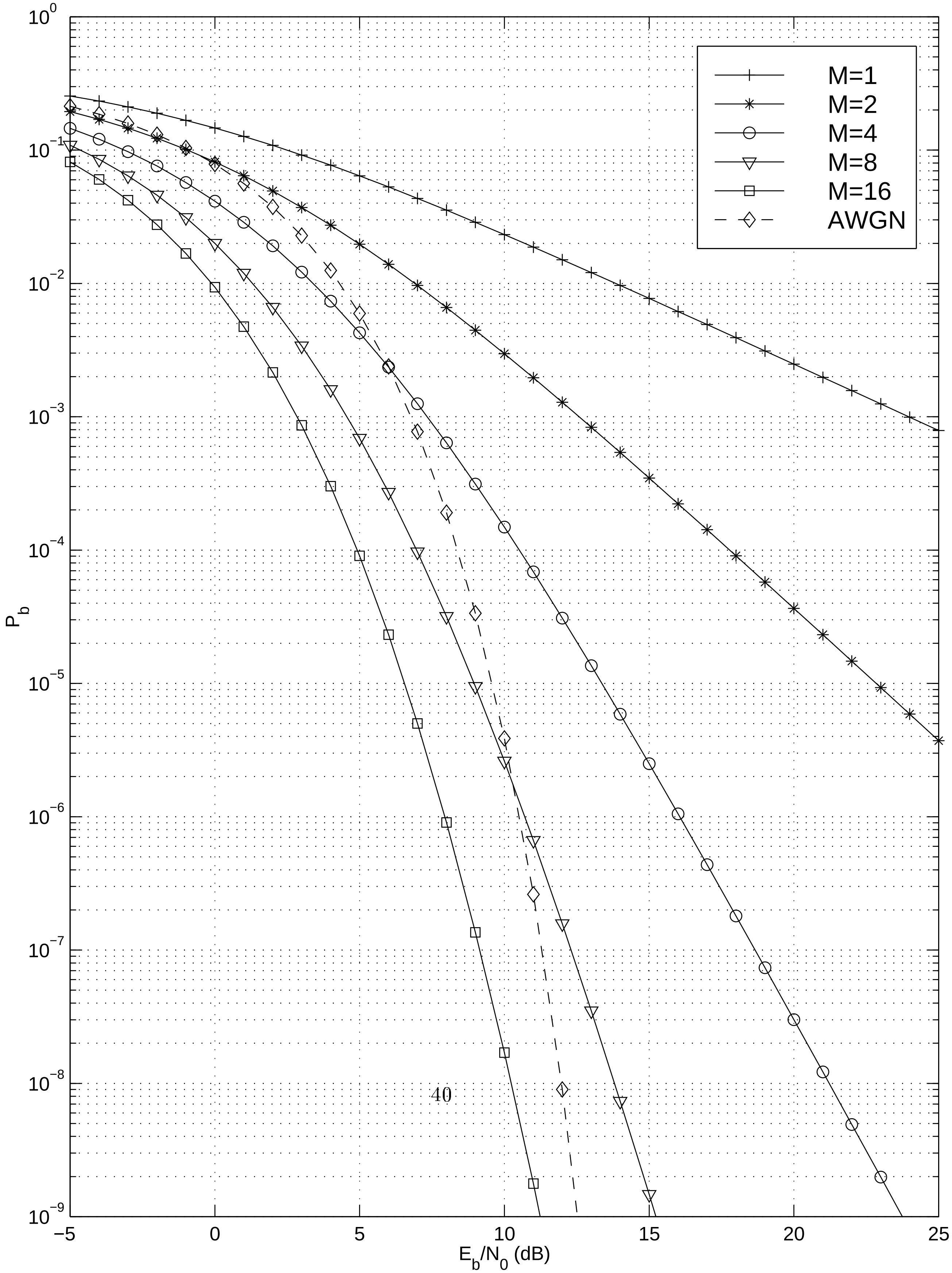


Figure 9: BER for Different Multiuser Diversity Systems without Power Control

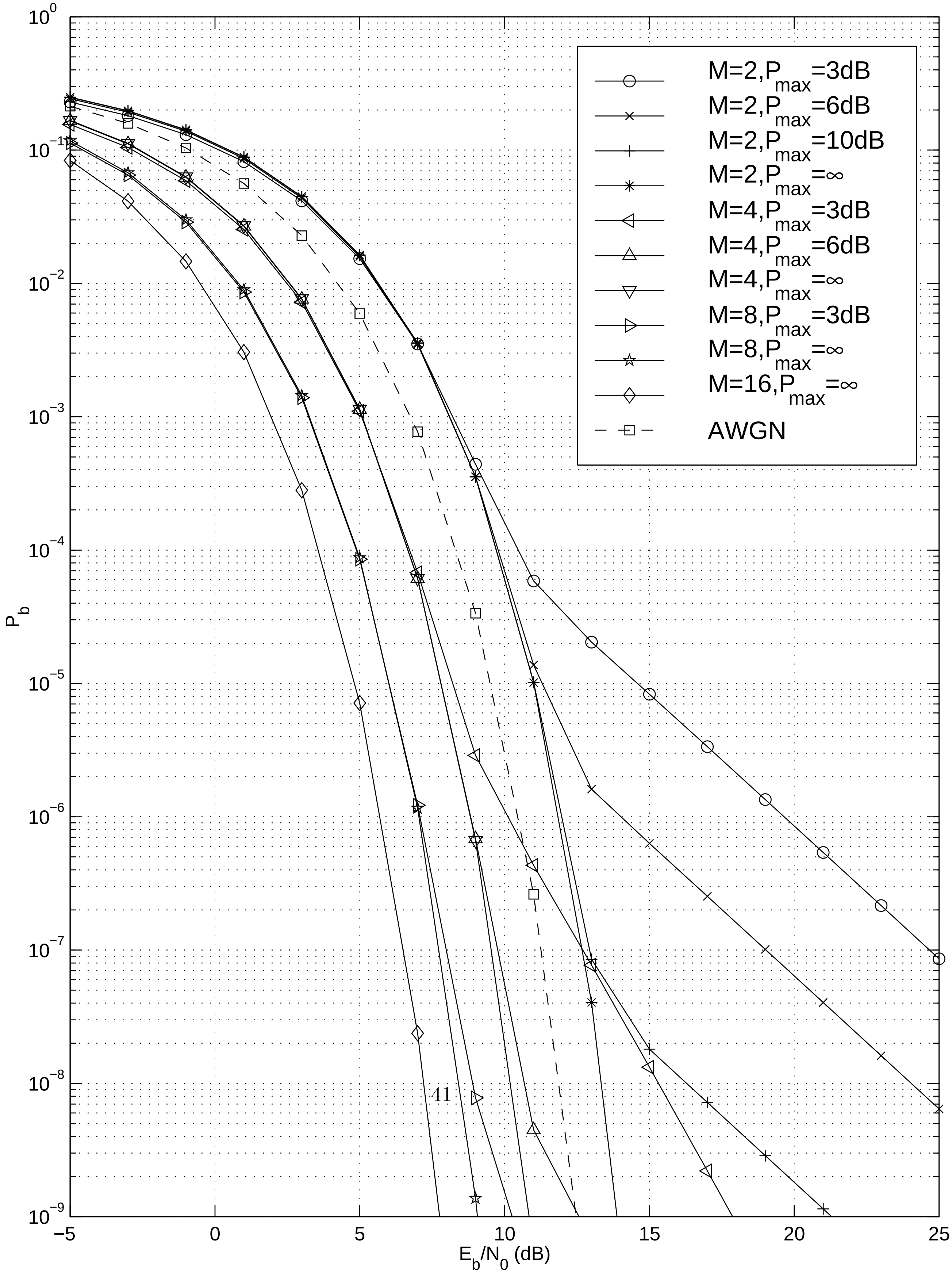


Figure 10: BER for a Multiuser Diversity Systems with Power Control