Linear Multiuser Receivers: Effective Interference, Effective Bandwidth and User Capacity

David N.C. Tse*
Dept. of EECS
University of California
Berkeley, CA 94720, USA
dtse@eecs.berkeley.edu

Stephen V. Hanly†
Dept. of EE
University of Melbourne
Australia
s.hanly@ee.mu.oz.au

Abstract

Multiuser receivers improve the performance of spread-spectrum and antenna-array systems by exploiting the structure of the multiaccess interference when demodulating the signals of a user. Much of the previous work on the performance analysis of multiuser receivers has focused on their ability to reject worst-case interference. Their performance in a power controlled network and the resulting user capacity are less well-understood. In this paper, we show that in a large system with each user using random spreading sequences, the limiting interference effects under several linear multiuser receivers can be *decoupled*, such that each interferer can be ascribed a level of *effective interference* that it provides to the user to be demodulated. Applying these results to the uplink of a single power-controlled cell, we derive an *effective bandwidth* characterization of the user capacity: the signal-to-interference requirements of all the users can be met if and only if the sum of the effective bandwidths of the users is less than the total number of degrees of freedom in the system. The effective bandwidth of a user depends only on its own SIR requirement, and simple expressions are derived for three linear receivers: the conventional matched filter, the decorrelator and the MMSE receiver. The effective bandwidths under the three receivers serve as a basis for performance comparison.

Keywords: multiuser detection, MMSE receiver, decorrelator, power control, effective bandwidth, effective interference, user capacity.

---

*The research of this author is supported by the Air Force Office of Scientific Research under grant F49620-96-1-0199, and by an NSF CAREER Award under grant NCR-9734090.

†S. Hanly gratefully acknowledges funding from an Australian Research Council large grant.

1 Introduction

To meet the growing demand of untethered applications, there have been intense efforts in recent years to develop more sophisticated physical layer communication techniques to increase the spectral efficiency of wireless systems. A significant thrust of work has been on developing *multiuser* receiver structures which mitigate the interference between users in spread spectrum systems. These receivers include the optimum multiuser detector [24], the linear decorrelator [11, 12] and the linear minimum mean-square error (MMSE) receiver [29, 13, 16, 17]. Unlike the conventional matched filter receiver, these techniques take into account the structure of the interference from other users when demodulating a user. Another important line of work is the development of processing techniques in systems with antenna arrays. Both spread-spectrum techniques and antenna arrays provide additional *degrees of freedom* through which communication can take place, and multiuser techniques aim to better exploit those degrees of freedom.

Despite significant work done in the area, there is still much debate about the *user capacity*[1] of the various approaches to deal with multiuser interference. One reason is that the performance of multiuser receivers in conjunction with networking level techniques of power control and resource allocation are less well understood than for more traditional multi-access schemes. Indeed, much of the previous work on performance evaluation of multiuser receivers focuses on their ability to reject worst-case interference (*near-far resistance* [11]) rather than on their performance in a power-controlled system. The main goal of this paper is to make progress towards addressing these issues.

One difficulty in understanding the performance of multiuser receivers in power-controlled environments stems from the intertwining of the effects of *all* of the interferers in the system. For example, the MMSE receiver depends on the signature sequences and received powers of all interferers, and hence at the output of the filter, it is hard to separate out the effect of individual interferers. The main result of this paper shows, somewhat surprisingly, that in a large system with many degrees of freedom and many users, a decoupling of the interfering effects is indeed possible for several important linear receivers: each interferer can be ascribed a level of *effective interference* that it provides to the user to be demodulated. The effective interference of an interferer depends only on the received power of the interferer, the received power of the user being demodulated and the achieved signal-to-interference ratio (SIR) at the output of the receiver.

Applying this notion of effective interference to the uplink of a single power-controlled cell, we derive an *effective bandwidth* characterization of the user capacity under several linear receivers. Assuming that each user's requirement can be expressed in terms of a target SIR at the output of the receiver, we will show that a notion of effective bandwidth can be defined such that the SIR requirements of all the users can be met if and only if the sum of the effective bandwidths of the users is less than the total number of degree of freedom in the system. These degrees of freedom can be provided by the processing

[1] In this paper, we use the term *user capacity* to refer to the number of users that can be supported at the desired quality-of-service requirement. This should be distinguished from the *information theoretic capacity* of a channel.

gain in a spread-spectrum system or the number of antenna elements in a system with an antenna array. These capacity characterizations are simple in that the effective bandwidth of a user depends only on its own SIR requirement and nothing else. While this approach yields an interference-limited characterization of the user capacity, we will also quantify the reduction in user capacity when there are additional power constraints on the users. We observe that the SIR is a reasonable performance measure for the class of linear multiuser receivers we are concerned with.

The effective bandwidth of a user depends on the multiuser receiver employed. Results for three receivers are obtained: the linear MMSE receiver, the decorrelator and the conventional matched filter receiver. We will show that the effective bandwidths are respectively:

$$e_{mmse}(\beta) = \frac{\beta}{1+\beta} \qquad e_{dec}(\beta) = 1 \qquad e_{mf}(\beta) = \beta,$$

where β is the SIR requirement of the user. These effective bandwidth expressions also provide a succinct basis for performance comparison between different receiver structures. In particular, the MMSE receiver occupies a special place as it can be shown to lead to the minimum effective bandwidth amongst all linear receivers. Moreover, its performance is the least understood of the three receivers, and its analysis is the main thrust of this paper.

To obtain these results, we assume that the users' signals arrive from *random directions*. In the context of a spread-spectrum system, this means that each of the users employ random spreading sequences. In the context of an antenna array system, this translates into independent fading from each of the users to each of the receiving antenna elements. We will also restrict our analysis to synchronous systems in this paper. Extensions of these results to symbol-asynchronous spread-spectrum systems can be found in [9].

Related results on the performance of multiuser receivers under random spreading sequences were obtained independently in [26], presented simultaneously as a conference version [19] of this work. They considered exclusively the single class case where every user has the same received power and same rate requirement, and derived Shannon theoretic performance. In the present paper, our main results are for situations where users have different received powers and possibly different SIR requirements.

The outline of the paper is as follows. In Section 2, we will introduce the basic model of a multi-access spread-spectrum system and the structure of the MMSE receiver. In Section 3, we will present our main result, that in a large system with each user using random spreading sequences, the limiting interference effects under the MMSE receiver can be decoupled into a sum of effective interference terms, one from each of the interferers. In sections 5 and 6, we apply this result to study the performance under power control and obtain a notion of *effective bandwidth*. In Section 7, we obtain analogous results for the decorrelating receiver. In Section 8, we show that similar ideas carry through for systems with antenna diversity. Section 9 contains our conclusions.

2 Basic Spread-Spectrum Model and the MMSE Receiver

In a spread-spectrum system, each of the user's information or coded symbols is spread onto a much larger bandwidth via modulation by its own *signature* or *spreading sequence*. The following is a chip-sampled discrete-time model for a symbol-synchronous multi-access spread-spectrum system:

$$\mathbf{Y} = \sum_{i=1}^{K} X_i \mathbf{s}_i + \mathbf{W}, \tag{1}$$

where $X_i \in \Re$ and $\mathbf{s}_i \in \Re^N$ are the transmitted symbol and signature spreading sequence of user i respectively, and $\mathbf{W}$ is $N(0, \sigma^2 I)$ background Gaussian noise. The length of the signature sequences is N, which one can also think of as the number of degrees of freedom. The received vector is $\mathbf{Y} \in \Re^N$. We assume the X_i's are independent and that $E[X_i] = 0$ and $E[X_i^2] = P_i$, where P_i is the received power of user i. There are K users in the system.

Rather than looking at multiuser detection, which involves hard decisions on a symbol by symbol basis, we are more interested in the problem of extracting good estimates of the (coded) symbols of each user as soft decisions to be used by the channel decoder. For this reason, we prefer the term "multiuser receiver" rather than "multiuser detector", although the latter is more common in the literature. In this case, the relevant performance measure is the signal-to-interference ratio (SIR) of the estimates.

We shall now focus on the demodulation of user 1, assuming that the receiver has already acquired the knowledge of the spreading sequences. In this paper, we shall confine ourselves to the study of *linear* demodulators, such that the estimates are linear functions of the received vector $\mathbf{Y}$. For user 1, the optimal demodulator $\mathbf{c}_1$ that generates a soft decision $\hat{X}_1 \equiv \mathbf{c}_1^t \mathbf{Y}$ maximizing the signal-to-interference ratio (SIR):

$$\beta_1 \equiv \frac{(\mathbf{c}_1^t \mathbf{s}_1)^2 P_1}{(\mathbf{c}_1^t \mathbf{c}_1)\sigma^2 + \sum_{i=2}^{K} (\mathbf{c}_1^t \mathbf{s}_i)^2 P_i}$$

is the MMSE receiver [2] [13, 16, 17].

As a comparison, note that the conventional CDMA approach simply matches the received vector to $\mathbf{s}_1$, the signature sequence of user 1. This is indeed the optimal receiver when the interference from other users is white. However, in general the multi-access interference is not white and has structure as defined by $\mathbf{s}_2, \mathbf{s}_3, \ldots, \mathbf{s}_K$, assumed to be known to the receiver. The MMSE receiver exploits the structure in this interference in maximizing the SIR of user 1.

While there are well-known formulas for the MMSE receiver and its performance, we will describe a simple derivation, which provides some geometric insights to the operation of this receiver. Let

[2] More precisely, this should be termed the *linear* least square (LLSE) receiver, since it is only optimal within the class of linear receivers if the X_i's are not Gaussian. In deference to the standard practice in the multiuser detection literature, however, we will call this the MMSE receiver.

$$\mathbf{Z} = \sum_{i=2}^{K} X_i \mathbf{s}_i + \mathbf{W}$$

be the total interference for user 1 from other users and background noise. Then

$$\mathbf{Y} = X_1 \mathbf{s}_1 + \mathbf{Z}$$

If $\mathbf{Z}$ were white, then

$$X_{mmse}(\mathbf{Y}) = \frac{\mathbf{s}_1^t \mathbf{Y}}{\mathbf{s}_1^t \mathbf{s}_1},$$

which is a projection onto $\mathbf{s}_1$, i.e. the conventional matched filter. In general, then, we should whiten the interference $\mathbf{Z}$ and then follow that by a projection. The covariance matrix of $\mathbf{Z}$ is

$$K_z = S_1 D_1 S_1^t + \sigma^2 I$$

where S_1 is a N by $K-1$ matrix whose columns are the signature sequences of the other users, and $D_1 = \mathrm{diag}(P_2, \ldots, P_K)$ is the covariance matrix of $(X_2, \ldots, X_K)^t$. K_z is positive definite. Factorize $K_z = Q^t \Lambda Q$, where $\Lambda = \mathrm{diag}\{\lambda_1, \ldots, \lambda_N\}$ is the diagonal matrix of (positive) eigenvalues of K_z, and the columns of Q are the orthonormal eigenvectors of K_z. The whitening filter is simply $\Lambda^{-\frac{1}{2}} Q$. Applying this to $\mathbf{Y}$, we get:

$$\Lambda^{-\frac{1}{2}} Q \mathbf{Y} = X_1 \Lambda^{-\frac{1}{2}} Q \mathbf{s}_1 + \Lambda^{-\frac{1}{2}} Q \mathbf{Z},$$

and we note that the interference is now white. We can then project it along the direction $\Lambda^{-\frac{1}{2}} Q \mathbf{s}_1$ to get a scalar sufficient statistic for the estimation problem:

$$R \equiv \mathbf{s}_1^t K_z^{-1} \mathbf{Y} = \left(\mathbf{s}_1^t K_z^{-1} \mathbf{s}_1 \right) X_1 + \mathbf{s}_1^t K_z^{-1} \mathbf{Z}$$

Thus, the MMSE demodulator is [13]:

$$X_{mmse}(\mathbf{Y}) = \frac{P_1}{1 + P_1 \mathbf{s}_1^t (S_1 D_1 S_1^t + \sigma^2 I)^{-1} \mathbf{s}_1} \mathbf{s}_1^t (S_1 D_1 S_1^t + \sigma^2 I)^{-1} \mathbf{Y} \tag{2}$$

and the signal to interference ratio for user 1 is

$$SIR_1 = P_1 \mathbf{s}_1^t (S_1 D_1 S_1^t + \sigma^2 I)^{-1} \mathbf{s}_1 \tag{3}$$

While the SIR is taken as the basic measure of performance in this paper, we would like to mention some connections to information theoretic quantities:

- If the linear receiver is followed by single-user decoders, one for each user, then the mutual information achieved for each user under an independent Gaussian input distribution is precisely

$$\frac{1}{2}\log(1+\mathrm{SIR}_i)$$

bits per symbol time. There is therefore a one-to-one monotonic relationship between the information theoretic rate and the achieved SIR. In particular, meeting a target SIR is equivalent to meeting a target rate.

- It has been shown [22] that any vertex of the Shannon capacity region of the CDMA channel (1) can be achieved by a combination of successive cancellation and MMSE demodulation. Each vertex corresponds to a particular choice of decoding order. The information theoretic rate achieved for the ith user in a given decoding order is

$$\frac{1}{2}\log\left(1+\tilde{\mathrm{SIR}}_i\right),$$

where $\tilde{\mathrm{SIR}}_i$ is the SIR at the output of the MMSE demodulator for the ith user, with the signals from the first $i-1$ users already cancelled off.

3 Performance Under Random Spreading Sequences

Eqn. (3) is a formula for the performance of the MMSE receiver, which one can compute for specific choices of signature sequences. However, it is not easy to obtain qualitative insights directly from the formula. For example, the effect of an individual interferer on the SIR for user 1 cannot be seen directly from this formula. In practice, it is often reasonable to assume that the spreading sequences are *randomly* and independently chosen. (See eg. [14, 3].) For example, they may be pseudorandom sequences, or the users choose their sequences from a large set of available sequences as they are admitted into the network, or the transmitted sequences may be distorted by random multipath fading channels. In this case, the performance of the optimal demodulator can be modeled as a random variable, since it is a function of the spreading sequences. In this section, we will show that, unlike the deterministic case, there is a great deal of analytical information one can obtain about this random performance in a large network. In the development below, we will assume that although the sequences are randomly chosen, they are known to the receiver once they are picked. In practice, this assumes that the change in the spreading sequences occurs at a much slower timescale than the time required to acquire the sequences. (There are known adaptive algorithms for which acquisition can be done blindly; see [6].) However, the *performance* of the MMSE receiver depends on the initial choice of the sequences and hence is random.

The model for random sequences is as follows : let $\mathbf{s}_i = \frac{1}{\sqrt{N}}(V_{i1},\ldots,V_{iN})^t, i=1,\ldots K$. The random variables V_{ik}'s are i.i.d., zero mean and variance 1. The normalization by $\frac{1}{\sqrt{N}}$ ensures that $E[\|\mathbf{s}_i\|^2]=1$. In practice, it is common that the entries of the spreading sequences are 1 or -1, but we want to keep the model general so that we can later apply our results to problems with other modes of diversity. For technical reasons, we will also make the mild assumption that $E[V_{iN}^4] < \infty$.

Our results are asymptotic in nature, for a large network. Thus, we consider the limiting regime where the number of users is large, i.e. $K \to \infty$. To support a large number of users, it is reasonable to scale up N as well, keeping the number of users per degree of freedom (equivalently, per unit bandwidth), $\alpha \equiv \frac{K}{N}$, fixed. We also assume that as we scale up the system, the empirical distribution of the powers of the users converge to a fixed distribution, say $F(P)$. The following is our main result, giving the asymptotic information about the SIR for user 1.

Theorem 3.1 *Let $\beta_1^{(N)}$ be the (random) SIR of the MMSE receiver for user 1 when the spreading length is N. Then $\beta_1^{(N)}$ converges to β_1^* in probability as $N \to \infty$, where β_1^* is the unique solution to the equation:*

$$\beta_1^* = \frac{P_1}{\sigma^2 + \alpha \mathbb{E}_P\left[I(P, P_1, \beta_1^*)\right]} \tag{4}$$

and

$$I(P, P_1, \beta_1^*) \equiv \frac{PP_1}{P_1 + P\beta_1^*} \tag{5}$$

Here, $\mathbb{E}_P[\cdot]$ denotes taking the expectation with respect to the limiting empirical distribution F of the received powers of the interferers.

Heuristically, this means that in a large system, the SIR β_1 is deterministic and approximately satisfies:

$$\beta_1 \approx \frac{P_1}{\sigma^2 + \frac{1}{N}\sum_{i=2}^{K} I(P_i, P_1, \beta_1)} \tag{6}$$

where as before P_i is the received power of user i. This result yields an interesting interpretation of the effect of each of the interfering users on the SIR of user 1: for a large system, the total interference can be decoupled into a sum of the background noise and an interference term from each of the other users. (The factor $\frac{1}{N}$ results from the processing gain of user 1.) The interference term depends only on the received power of the interfering user, the received power of user 1 and the attained SIR. It does not depend on the other interfering users except through the attained SIR β_1. This decoupling is rather surprising since the effect of an interferer depends on the MMSE receiver $\mathbf{c}_1$, which in turn is a function of the signature sequences and received powers of all the users in the system.

One must be cautioned not to think that this result implies that the interfering effect of the other users on a particular user is additive across users. It is not, since the interference term $I(P_i, P_1, \beta_1)$ from interferer i depends on the attained SIR which in turn is a function of the entire system. Due to the following proposition, on the other hand, one can make a related statement.

Proposition 3.2 *The equation*

$$x = \frac{P_1}{\sigma^2 + \frac{1}{N}\sum_{i=2}^{K} I(P_i, P_1, x)} \tag{7}$$

has a unique fixed point x^*. For any x, $x^* \geq x$ if and only if

$$\frac{P_1}{\sigma^2 + \frac{1}{N}\sum_{i=2}^{K} I(P_i, P_1, x)} \geq x$$

Proof. Define the function

$$\begin{aligned} f(x) &\equiv \frac{1}{P_1}\left(\sigma^2 x + \frac{1}{N}\sum_{i=2}^{K} x I(P_i, P_1, x)\right) \\ &= \frac{1}{P_1}\left(\sigma^2 x + \frac{1}{N}\sum_{i=2}^{K} \frac{P P_1 x}{P_1 + P x}\right) \end{aligned}$$

which we note to be a continuous, strictly increasing function.

To see that a fixed point x^* exists to (7), we note that $f(0) = 0$ and $f(\infty) = \infty$ so it follows that there must exist a value x^* satisfying $f(x^*) = 1$. But this implies that x^* is a unique fixed point of (7). By monotonicity of f,

$$\begin{aligned} x^* \geq x &\Leftrightarrow f(x) \leq 1 \\ &\Leftrightarrow \frac{P_1}{\sigma^2 + \frac{1}{N}\sum_{i=2}^{K} I(P_i, P_1, x)} \geq x \end{aligned}$$

□

It follows then that to check if the target for user 1's SIR, β_T, can be met for a given system of users, it suffices to check the following condition:

$$\frac{P_1}{\sigma^2 + \frac{1}{N}\sum_{i=2}^{K} I(P_i, P_1, \beta_T)} \geq \beta_T$$

Based on this interpretation, it seems justified to term $I(P_i, P_1, \beta_T)$ as the *effective interference* of user i on user 1, at a target SIR of β_T.

To gain more insights into this concept of effective interference, it is helpful to compare the situation with that when the conventional matched filter $\mathbf{s}_1$ is used for the demodulation. For that case, we have the following proposition, in parallel with Theorem 3.1:

Proposition 3.3 *Let $\beta_{1,MF}^{(N)}$ be the (random) SIR of the conventional matched filter receiver for user 1 when the spreading length is N. Then as $N, K \to \infty$ with $\frac{K}{N} \to \alpha$, $\beta_{1,MF}^{(N)}$ converges in probability to*

$$\beta^*_{1,MF} = \frac{P_1}{\sigma^2 + \mathbb{E}_P[P]}$$

where as before the expectation is taken with respect to the limiting empirical distribution F of the received powers of the interferers.

Proof. See appendix B. □

Hence, for large N, the performance of the matched receiver is approximately:

$$\beta_{1,MF} \approx \frac{P_1}{\sigma^2 + \frac{1}{N}\sum_{i=2}^{K} P_i} \tag{8}$$

Comparing this expression with eqn. (6), we see that the interference due to user i is simply P_i in place of $I(P_i, P_1, \beta_1)$. Since the matched receiver filter is independent of the signature sequences of the other users, it is not surprising that the interference is linear in the received powers of the interferers. In the case of the MMSE receiver, the filter does depend on the signature sequences of the interferers, thus resulting in the interference being a non-linear function of the received power of the interferer. Also, observe that $I(P_i, P_1, \beta_1) < P_i$, which is expected since the MMSE receiver maximizes the SIR amongst all linear receivers. But more importantly, while for the conventional receiver, the interference grows unbounded as the received power of the interferer increases, we see that for the MMSE receiver, the effective interference (5) from user i is bounded and approaches $\frac{P_1}{\beta_1}$ as P_i goes to infinity. Thus, while the SIR of the matched filter receiver goes to zero for large interferers' powers, the SIR of the MMSE receiver does not. This is the well-known *near-far resistance* property of the MMSE receiver [13]. The intuition is that as the power of an interferer grows to infinity, the MMSE receiver will null out its signal. While the near-far resistance property has been reported by previous authors, Theorem 3.1 goes beyond these works in that it not only quantifies the worst-case performance (i.e. large interferer's power) but also the performance for all finite values of the interference. This is useful, for example, in situations when power control is exercised, as we will turn to in the next section.

In general, we have no explicit solution for the SIR β^*_1 in eqn. (4). However, for the special case when the received powers of all users are the same, the equation is quadratic in β^*_1 and a simple solution is obtained (independently obtained in [26]):

$$\beta^*_1 = \frac{(1-\alpha)P}{2\sigma^2} - \frac{1}{2} + \sqrt{\frac{(1-\alpha)^2P^2}{4\sigma^4} + \frac{(1+\alpha)P}{2\sigma^2} + \frac{1}{4}} \tag{9}$$

We see that the β^*_1 is positive for all values of α, and approaches 0 as α, the number of users per degree of freedom, goes to infinity.

To get a sense of the convergence of the random SIR to the asymptotic limit in the equal received power case, Fig. 1 compares the actually realized SIR's from randomly generated spreading sequences to the asymptotic limit (9). For different spreading lengths and for each value of α, 100 samples of realized SIR's for user 1 are obtained from randomly generated $+1$ and -1 spreading sequences. One sees that as the processing gain increases, the spread around the asymptotic limit becomes more narrow, to about 1 or 2 dB when $N = 128$. Note however that for a fixed processing gain, the spread does not get smaller as the number of users increases, which means that the *relative* spread is large when the SIR is low. Fig. 2 plots the SIR's attained across users for a single realization of the random spreading sequences. The processing gain $N = 128$ and the number of users is 80. Again, there is a spread of about 1 dB around the asymptotic limit.

Theorem 3.1 gives only the asymptotic limit but does not describe the fluctuation of the SIR around this limit for finite-sized system. A sequel [20] to this paper is devoted to the analysis of such fluctuations, via Central-Limit theorems. It turns out that even the computation of the variance of the fluctuations is non-trivial. See also [7, 8] for a related study.

Two performance measures commonly used in the literature for multiuser receivers are their *efficiency* and their *asymptotic efficiency* [25]. In the context of linear receivers, the efficiency for user 1 is defined to be the ratio of the achieved SIR to the SIR when there is no interferer and only background noise. For the MMSE receiver with random spreading sequences and equal received power for all users, this is given by:

$$\frac{\beta_1^* \sigma^2}{P}$$

where β_1^* is given by (9). The asymptotic efficiency η_1 is the limiting efficiency as the background noise goes to zero. If $\alpha \leq 1$, this is given by:

$$\eta_1 := \lim_{\sigma \to 0} \frac{\beta_1^* \sigma^2}{P} = 1 - \alpha$$

For $\alpha > 1$, the limiting SIR is positive but bounded:

$$\lim_{\sigma \to 0} \beta_1^* = \frac{1}{\alpha - 1} \tag{10}$$

and so the asymptotic efficiency is 0.

4 Proof of Main Theorem

We will now prove our main result, Theorem 3.1. It hinges on a result about the limiting eigenvalue distribution of large matrices whose elements are random variables. Let X_{ij} be an infinite array of i.i.d. complex-valued random variables with variances 1, and U_i be a sequence of real-valued random variables. Let $A_{n,m}$ be a n by m matrix, whose (i, j) th entry is $\frac{X_{ij}}{\sqrt{n}}$. Let T_m be an m by m diagonal matrix whose diagonal entries are $U_1, \ldots U_m$; we assume that as $m \to \infty$, the empirical distribution of these entries converges almost surely to a non-random limit F. Moreover, T_m is independent of $A_{m,n}$.

The matrix $A_{n,m} T_m A_{n,m}^H$ (A^H is the complex conjugate transpose of A) is n by n Hermitian and has real non-negative eigenvalues $\lambda_1^{(n)}, \ldots, \lambda_n^{(n)}$. Let $G_n(\lambda)$ be the empirical distribution of the eigenvalues; since the eigenvalues are random, so is G_n. (The empirical distribution of the eigenvalues depends on the realization of the random entries of $A_{n,m}$ and T_m.) The following theorem due to Silverstein and Bai [18] , which is a strengthening of an earlier result by Marcenko and Pastur [15], gives the asymptotic behavior of G_n as n and m grows. The solution is in terms of the Stieltjes transform, which for any distribution G is defined as:

$$m_G(z) = \int \frac{1}{\lambda - z} dG(\lambda)$$

for $z \in \mathcal{C}^+ \equiv \{z \in \mathcal{C} : \mathrm{Im} z > 0\}$.

Theorem 4.1 *As $n, m \to \infty$ such that $\frac{m}{n} \to \alpha > 0$, then almost surely G_n converges in distribution to a non-random limit G^*. The Stieltjes transform $m(z)$ of the limit G^* satisfies the following equation:*

$$m(z) = \frac{1}{-z + \alpha \int \frac{\tau dF(\tau)}{1 + \tau m(z)}} \tag{11}$$

for all $z \in \mathcal{C}^+$.

The above theorem says that the empirical distribution of the eigenvalues for large random matrices looks the same for almost all realizations of the entries. Eqn. (11) gives a functional equation for the Stieltjes transform of the limit; in general, it cannot be solved explicitly.

We can apply this result to the covariance matrix $K_z = S_1 D_1 S_1^t + \sigma^2 I$ of the interference; note that in this case F is the distribution function of the received power. It follows that in a large system with random signature sequences, the spectrum of the interference is essentially deterministic, since it converges to the nonrandom limiting distribution described in the theorem. Moreover, since the limiting eigenvalue distribution is not degenerate, it follows that the deterministic spectrum is colored and not white. This is perhaps a little counter-intuitive. For, if the number of interferers were fixed and the number of degrees of freedom increased, then each interferer would be more or less orthogonal to user 1 and the overall interference would be white. On the other hand, if the number of degrees of freedom were fixed and the number of interferers increased, the aggregate interference would also become increasing white because of averaging. Theorem 4.1 tells us, however, that when there are many interferers *and* many degrees of freedom, neither intuition is correct and the aggregate interference has a colored spectrum in the limit. As a consequence, the MMSE receiver outperforms the conventional matched filter, even in the limit.

Theorem 4.1 gives the asymptotic distribution of the eigenvalues of the covariance matrix K_z. However, this is in general not enough for characterizing the SIR performance for user 1, as that depends on the position of $\mathbf{s}_1$ relative to the *eigenvectors* of K_z. This can be seen by writing $K_z = U^t \Lambda U$ where Λ is diagonal and U is orthogonal, so that the SIR for user 1 is given by

$$\beta_1 = P_1 \mathbf{s}_1 K_z^{-1} \mathbf{s}_1 = P_1 (U\mathbf{s}_1)^t \Lambda^{-1} (U\mathbf{s}_1).$$

However, the following lemma shows that the distribution of the eigenvectors is asymptotically irrelevant since for large spreading length, $\mathbf{s}_1$ looks "white" in any coordinate system, in the sense of containing about the same amount of energy in each direction.

Lemma 4.2 *Let Q be a random m by n matrix ($m < n$) such that every realization consists of orthonormal rows. Let $\mathbf{X} = (V_1, \ldots, V_n)^t$ where the V_i's are i.i.d. random variables independent of Q, $E[V_i] = 0$, $E[V_i^2] = 1$ and $E[V_i^4] < \infty$. Then for any $\epsilon > 0$,*

$$Pr\left[\left| \frac{\|Q\mathbf{X}\|^2}{n} - \frac{m}{n} \right| > \epsilon \right] < \frac{C}{n}$$

for some constant C which depends only on ϵ and the statistics of V_i.

Proof. See appendix B. □

This lemma allows us to express the limiting SIR in terms of only the eigenvalue distribution of K_z.

Lemma 4.3 *As $N, K \to \infty$, $\frac{K}{N} \to \alpha$, the SIR $\beta_1^{(N)}$ converges in probability to*

$$\beta^* = \int_0^\infty \frac{P_1}{\lambda + \sigma^2} dG^*(\lambda),$$

where G^* is the limiting eigenvalue distribution of the random matrix $S_1 D_1 S_1^t$.

Proof. See appendix B. □

We shall now complete the proof of the Theorem 3.1 by evaluating this limit β^*.

Consider the Stieltjes transform of the limiting spectrum G^* of the matrix $S_1 D_1 S_1^t + \sigma^2 I$:

$$m_{G^*}(z) = \int_0^\infty \frac{1}{\lambda - z} dG^*(\lambda) \qquad z \in \mathcal{C}^+.$$

By Theorem 4.1, this satisfies:

$$m_{G^*}(z) = \frac{1}{-z + \alpha \int \frac{P dF(P)}{1 + P m_{G^*}(z)}} \tag{12}$$

where F is the limiting distribution of the received powers of the users.

Since the support of G^* is on the non-negative real axis, m_{G^*} is continuous in the neighborhood of $z = -\sigma^2$. It follows that

$$\lim_{z \to -\sigma^2} m_G^*(z) = \int_0^\infty \frac{1}{\lambda + \sigma^2} dG^*(\lambda) = \frac{\beta^*}{P_1}$$

By the continuity of the righthand side of eqn. (12) as a function of $m_{G^*}(z)$, it follows from Lemma (4.3) that

$$\frac{\beta^*}{P_1} = \frac{1}{\sigma^2 + \alpha \int_0^\infty \frac{P dF(P)}{1 + \frac{P\beta^*}{P_1}}}.$$

Hence the limiting SIR for user 1 satisfies:

$$\beta_1^* = \frac{P_1}{\sigma^2 + \alpha \int_0^\infty \frac{P_1 P dF(P)}{P_1 + P\beta_1^*}}.$$

which completes the proof of the theorem.

While the above provides a rigorous proof, it provides little intuition as *why* Theorem 3.1 is true. In particular, a better understanding of the decoupling phenomenon between interferers is desired. Based on some new results obtained in [27], we provide a heuristic but more intuitive derivation of formula (4) in Appendix A, bypassing the mysterious Stieltjes transform characterization of the limiting eigenvalue distribution of random matrices in (11) and only basing ourselves on Lemma 4.3.

5 User Capacity under Power Control

We observed in Section 3 that in the conventional receiver case, the interference of a user is proportional to its power, and hence a strong interferer can completely overcome a weaker signal. This is the so-called near-far problem, and a well-known consequence is that the conventional receiver can only avoid this via tight power control. We also observed that the MMSE receiver does not suffer arbitrarily poorly from the near-far problem, and indeed this is one of the key motivations for the original work on multiuser detection [24]. Nevertheless, a MMSE receiver still suffers interference from other users, and it follows that user capacity can be increased and power consumption reduced, if power control is employed.

In the present section we consider the case in which all users require an SIR of exactly β^*, given a processing gain of N degrees of freedom per symbol. For a given number of users we compute the minimum power consumption required to achieve β^* for all users, and then look at the maximum number of users per degree of freedom supportable for a given power constraint under power control. Of particular interest is the maximum number without power constraint, which we refer to as the *user capacity* of the system (in terms of number of users per degree of freedom.) This is the point at which saturation occurs as we put in so many users that we drive the required power level to infinity. We will show that this user capacity is different but finite for both the conventional and the MMSE receivers, showing that both are interference-limited systems. As before, our results are asymptotic as the the processing gain N goes to infinity.

Let us focus first on the conventional receiver. Under the matched filter, Prop. 3.3 tells us that, asymptotically, users receive the same level of interference, and hence must be received at the same power level to get the same SIR β^*. It is easy to compute that with a processing gain of N and $N\alpha$ users, the common received power required, asymptotically as $N \to \infty$, for the conventional receiver is given by

$$P_{mf}(\beta^*) = \frac{\beta^* \sigma^2}{1 - \alpha\beta^*} \tag{13}$$

For a given constraint P on the received power, the maximum number of users supportable is then:

$$\frac{1}{\beta^*} - \frac{\sigma^2}{P} \quad \text{users/degree of freedom}$$

The user capacity of the conventional receiver when $P = \infty$ is then

$$C_{mf}(\beta^*) = \frac{1}{\beta^*} \quad \text{users/degree of freedom} \tag{14}$$

Put it another way, as $\alpha \to \frac{1}{\beta^*}$, the system saturates and the required power level goes to infinity. A similar result is given in [5].

Now let us turn to the MMSE receiver. To satisfy given target SIR requirements for each user, [10, 21, 2] showed that there is an optimal solution for which the received power of *every user* is minimized; moreover, they gave an iterative algorithm to compute it. However, here we can give an explicit solution and characterize the resulting user capacity.

To begin, we fix the number of users per degree of freedom at α. As in the conventional receiver case, it turns out that the system saturates if α is too high, so we first obtain a necessary and sufficient condition for feasibility. The following theorem shows that in the limit of a large number of degrees of freedom, the system is feasible if and only if the SIR can be met with equal received powers for all users.

Theorem 5.1 *If*

$$\alpha \geq \frac{1+\beta^*}{\beta^*}$$

then there is no distribution F of received powers such that the SIR requirements of all users are satisfied, i.e.:

$$\frac{Q}{\sigma^2 + \alpha \int_0^\infty I(P, Q, \beta^*) dF(P)} \geq \beta^* \quad \textit{for all } Q \textit{ in the support of } F \tag{15}$$

On the other hand, if $\alpha < \frac{1+\beta^*}{\beta^*}$, the SIR requirements of all users can be satisfied and the minimum power solution is having the received powers of all users to be

$$P_{mmse}(\beta^*) = \frac{\beta^* \sigma^2}{1 - \alpha \frac{\beta^*}{1+\beta^*}} \tag{16}$$

Proof. Suppose that there is a power distribution F such that all users get β^*, i.e.

$$\frac{Q}{\sigma^2 + \alpha \int_0^\infty I(P, Q, \beta^*) dF(P)} \geq \beta^* \quad \text{for all } Q \text{ in the support of } F$$

Let P^* be the power of the weakest user in this distribution, i.e.

$$P^* \equiv \inf\{P : F(P) > 0\}$$

and note that $\forall P \geq P^*, I(P^*, P^*, \beta^*) \leq I(P, P^*, \beta^*)$. Focusing on the user with received power P^*, since

$$\frac{P^*}{\sigma^2 + \alpha \int_0^\infty I(P, P^*, \beta^*) dF(P)} \geq \beta^*$$

therefore

$$\frac{P^*}{\sigma^2 + \alpha I(P^*, P^*, \beta^*)} \geq \beta^*$$

Using the explicit expression for the effective interference term and rearranging terms, the last statement is equivalent to:

$$P^*(1 - \alpha \frac{\beta^*}{1+\beta^*}) \geq \beta^* \sigma^2$$

Hence,

$$\alpha < \frac{1+\beta^*}{\beta^*}$$

This proves the first part of the proposition.

Conversely, if $\alpha < \frac{1+\beta^*}{\beta^*}$, then it can be easily checked that $P_{mmse}(\beta^*)$ is positive and satisfies

$$\frac{P_{mmse}(\beta^*)}{\sigma^2 + \alpha I(P_{mmse}(\beta^*), P_{mmse}(\beta^*), \beta^*)} = \beta^*$$

By Theorem 3.1, this implies that by assigning all users the same received power $P_{mmse}(\beta^*)$, they will all achieve the SIR requirement β^*. To see that this is the minimal solution, suppose that F is another power distribution such that the SIR requirements of all users are satisfied, and let P^* be the power of the weakest user of this distribution. By exactly the same argument as the proof of the first half of this proposition, we conclude that:

$$P^* \geq \frac{\beta^* \sigma^2}{1 - \alpha \frac{\beta^*}{1+\beta^*}} = P_{mmse}(\beta^*)$$

This shows that indeed the solution with equal received powers at $P_{mmse}(\beta^*)$ is the minimal solution.

$\square$

Hence, the user capacity of the system under MMSE receiver is:

$$C_{mmse}(\beta^*) = 1 + \frac{1}{\beta^*} \quad \text{users/degree of freedom.} \tag{17}$$

Moreover, for a given received power constraint P, the maximum number of users that can be supported is to assign each user the same received power, and that number is given by:

$$(1+\beta^*)(\frac{1}{\beta^*} - \frac{\sigma^2}{P}) \quad \text{users/degree of freedom.}$$

Contrasting (13) and (14) with (16) and (17), we note that if α is feasible for both types of receiver, then the MMSE power consumption is less than the matched filter power consumption, and the MMSE has potentially much greater user capacity. Indeed, if $\alpha < 1$ then we can take β^* arbitrarily high without saturating the MMSE receiver, whereas the the conventional receiver saturates as $\beta^* \uparrow \frac{1}{\alpha}$. For fixed β^*, we also note that the MMSE saturates at a higher value of α, yielding a user capacity of precisely 1 more user per degree of freedom than the conventional receiver. On the other hand, the relative gain of the MMSE is not so large for small values of β^*.

The above user capacity results are derived in the context of random spreading sequences. A natural question to ask is whether one can get performance gain if we one optimizes the choice of the sequences. In [27], it is shown that even with the optimal choice of sequences, the user capacity (without power constraint) under the MMSE receiver is still $1 + \frac{1}{\beta^*}$ users per degree of freedom. However, somewhat surprisingly, the capacity gap between the MMSE and conventional receiver disappears under optimal sequences.

6 Multiple Classes and Effective Bandwidths

It is straightforward to generalize our results to the case in which we have J classes, with class j users requiring a SIR of β_j. We denote the number of users of class j by $\alpha_j N$, and again consider the limiting regime $N \uparrow \infty$.

The conventional matched filter results generalize very easily to

$$P_{mf}(j) = \frac{\beta_j \sigma^2}{1 - \sum_{j=1}^{J} \alpha_j \beta_j}$$

where $P_{mf}(j)$ denotes the common received power level of all users of class j (see [5]). Thus, the user capacity constraint on feasible values of $(\alpha_1, \ldots, \alpha_J)$ is the linear constraint $\sum_{j=1}^{J} \alpha_j \beta_j < 1$. Furthermore, if class j users have a maximum power constraint that $P_{mf}(j) \leq \bar{P}_j$, for each j, then the tighter user capacity constraint:

$$\sum_{j=1}^{J} \alpha_j \beta_j \leq \min_{1 \leq i \leq J} \left[1 - \frac{\beta_i \sigma^2}{\bar{P}_i} \right]$$

emerges ([4]). It seems very reasonable to call β_j the bandwidth of class j users, in degrees of freedom per class j user. Let us denote this bandwidth by

$$e_{mf}(\beta_j) \equiv \beta_j \quad \text{degrees of freedom per class j user.}$$

We now show that the MMSE filter results generalize in a similar manner. It is clear in this case also that the minimal power solution consists of the same received power for each class; let all users in class j be received at power P_j. Then the power control equations become

$$\frac{P_j}{\sigma^2 + \sum_{i=1}^{J} \alpha_j I(P_i, P_j, \beta_j)} = \beta_j \quad j = 1, 2, \ldots, J \tag{18}$$

where, as in Theorem 3.1, $I(P_i, P_j, \beta_j) \equiv \frac{P_i P_j}{P_j + P_i \beta_j}$. But (18) implies that $\frac{\beta_j}{P_j}$ is a constant, which allows us to simplify (18) down to

$$P_{mmse}(i) = \frac{\beta_i \sigma^2}{1 - \sum_{j=1}^{J} \alpha_j \frac{\beta_j}{1+\beta_j}} \quad \text{i= 1, 2, \ldots, J.} \tag{19}$$

The user capacity constraint for the MMSE receiver with J classes is therefore given by

$$\sum_{j=1}^{J} \alpha_j \frac{\beta_j}{1 + \beta_j} < 1 \tag{20}$$

which is linear in $\alpha_1, \ldots, \alpha_J$.

As above, maximum power constraints provide tighter capacity constraints, and in this context we note that (19) implies that

$$\sum_{j=1}^{J} \alpha_j \frac{\beta_j}{1 + \beta_j} = 1 - \frac{\beta_i \sigma^2}{P_{mmse}(i)} \quad i = 1, 2, \ldots, J.$$

Thus if $P_{mmse}(i) \leq \bar{P}_i$ is a maximum power constraint on class i, then the linear constraint

$$\sum_{j=1}^{J} \alpha_j \frac{\beta_j}{1 + \beta_j} \leq \min_{1 \leq i \leq J} \left[1 - \frac{\beta_i \sigma^2}{\bar{P}_i} \right]$$

defines the restricted user capacity region of the system. It seems very reasonable to define the effective bandwidth of class j users to be $e_{mmse}(\beta_j)$ degrees of freedom per user, where

$$e_{mmse}(\beta_j) \equiv \frac{\beta_j}{1+\beta_j}.$$

Linearity in the matched filter case is a straightforward consequence of the fact that powers of interferers add. However, our MMSE effective bandwidth results are rather surprising, and it is a consequence of the asymptotic decoupling of the interference due to other users. For more discussions about the linearity of the user capacity region under MMSE, please consult Appendix A.

Fig. 3 gives an example of a user capacity region for two classes of users, one with SIR requirement 1dB and the other 10 dB. The upper line gives the asymptotic limit for the boundary of the region, under the MMSE receiver. The simulation curve gives the average number of class 2 users admissible as a function of the number of class 1 users in the system, for a spreading length of 64. The average number is obtained by averaging over 100 realizations of the spreading sequences. The actual number of class 2 users depend on the realization of the spreading sequences, and will fluctuate around this average, as was seen in Fig. 1.

One interesting observation is that no matter how high β is, the MMSE effective bandwidth of a user is upper bounded by unity. We will gain further insight into why this is so in the next section.

7 The Decorrelator

To this point we have contrasted the performance of the MMSE receiver with that of the conventional matched filter receiver. It is also illuminating to compare its performance with that of the decorrelator.

The decorrelator was in fact the first linear multiuser detector, introduced by Lupas and Verdu [11]. This receiver is known to be optimal in the worst case scenario in which interferers' powers tend to infinity; its near-far resistance is optimal [12]. Its main shortcoming, as we will see, is that each user has an effective bandwidth of 1 degree of freedom, which can be wasteful when the SIR of the user is small. On the other hand, it is hardly wasteful when the SIR is large.

We can write the channel equation (1) in matrix form:

$$\mathbf{Y} = S\mathbf{X} + \mathbf{W}$$

where $\mathbf{X} = (X_1, \ldots, X_K)^t$, and $S = [\mathbf{s}_1, \ldots, \mathbf{s}_K]$ is the matrix of signature sequences. It is well known ([11]) that the matched filter outputs

$$\mathbf{R} = S^t S\mathbf{X} + S^t\mathbf{W}$$

are sufficient statistics to recover the inputs X.

Consider now a further linear transformation applied to the matched filter outputs, to obtain

$$\mathbf{U} \equiv (S^tS)^{-1}\mathbf{R} = \mathbf{X} + (S^tS)^{-1}S^t\mathbf{W}$$

The overall filter $(S^tS)^{-1}S^t$ is called the decorrelating receiver. If the inverse does not exist, then the pseudo-inverse is used in its place. Observe that in the absence of external noise the decorrelator output would be the vector $\mathbf{X}$, and as such it represents the optimal zero-forcing linear filter. At this point, it is useful to provide an expression for the covariance matrix Σ of the "noise" $(S^tS)^{-1}S^t\mathbf{W}$, namely

$$\Sigma = (S^tS)^{-1}\sigma^2$$

The decorrelator for the user i returns U_i as an estimate of X_i. Thus, the channel for user i is given by

$$X_i \to X_i + N_i$$

where N_i is a zero-mean, Gaussian random variable of variance Σ_{ii}. The SIR for user i is given by P_i/Σ_{ii}. An important point about the decorrelator detector is that the correlation between the noise variables $(N_i)_{i=1}^N$ is not exploited, which explains why it is suboptimal to the MMSE receiver.

We can think of each row of $(S^tS)^{-1}S^t$ as a separate decorrelating filter for each user. For example, if we denote the first row of $(S^tS)^{-1}S^t$ by $\mathbf{r}_1$, then the decorrelating filter for user 1 is to apply $\mathbf{r}_1^t$ to the received signal $\mathbf{Y}$ to obtain the decorrelator estimate of user 1's symbol. It is insightful to look for a geometric interpretation. It turns out that each row of $(S^tS)^{-1}S^t$ is the orthogonal projection of the corresponding $\mathbf{s}_k$ onto the subspace $(\text{span}\{(\mathbf{s}_j)_{j\neq k}\})^\perp$. For example, $\mathbf{r}_1$ is the orthogonal projection of $\mathbf{s}_1$ onto $(\text{span}\{(\mathbf{s}_j)_{j\neq 1}\})^\perp$. To see why this is so, we prove the following proposition:

Proposition 7.1 *The vector $\mathbf{r}_1$ is the orthogonal projection of $\mathbf{s}_1$ onto the subspace V, defined by $V \equiv (span(\{\mathbf{s}_2, \mathbf{s}_3, \ldots, \mathbf{s}_K\}))^\perp$, and the SIR_1 for user 1 is given by*

$$SIR_1 = \frac{P_1}{\sigma^2}\mathbf{r}^t\mathbf{r}$$

Proof. Let us begin by denoting the orthogonal projection of $\mathbf{s}_1$ onto V by $\mathbf{v}_1$. Since $\mathbf{v}_1$ lies in V, the effect of applying $\mathbf{v}_1$ to $\mathbf{Y}$ nulls out the interference of users $2, \ldots, K$, and the SIR under $\mathbf{v}_1$, which we denote by $\text{SIR}_1^{(\mathbf{v}_1)}$, satisfies:

$$\text{SIR}_1^{(\mathbf{v}_1)} = \frac{P_1}{\sigma^2}\frac{(\mathbf{v}_1^t\mathbf{s}_1)^2}{\mathbf{v}_1^t\mathbf{v}_1}$$

The same applies to the decorrelator $\mathbf{r}_1$, since it also lies in V, which has corresponding SIR given by

$$\text{SIR}_1^{(\mathbf{r}_1)} = \frac{P_1}{\sigma^2}\frac{(\mathbf{r}_1^t\mathbf{s}_1)^2}{\mathbf{r}_1^t\mathbf{r}_1}$$

Now let $\mathrm{SIR}_1^{\mathrm{mmse}}(\sigma^2)$ be the SIR of user 1 under the MMSE linear receiver. Since $\mathbf{v}_1$ is the projection of $\mathbf{s}_1$ onto V, it can be seen that $\mathbf{v}_1$ achieves the best SIR among all linear receivers constrained to be in V. On the other hand, the MMSE is the optimal linear receiver overall. Hence, the following inequalities must hold:

$$\mathrm{SIR}_1^{(\mathbf{r}_1)} \leq \mathrm{SIR}_1^{(\mathbf{v}_1)} \leq \mathrm{SIR}_1^{\mathrm{mmse}}$$

But it is shown in [11] that the decorrelator has the optimal asymptotic efficiency in the class of linear receivers, i.e.

$$\lim_{\sigma^2 \to 0} \mathrm{SIR}_1^{(\mathbf{r}_1)} \sigma^2 = \lim_{\sigma^2 \to 0} \mathrm{SIR}_1^{\mathrm{mmse}} \sigma^2,$$

and hence

$$P_1 \frac{(\mathbf{r}_1^t \mathbf{s}_1)^2}{\mathbf{r}_1^t \mathbf{r}_1} = P_1 \frac{(\mathbf{v}_1^t \mathbf{s}_1)^2}{\mathbf{v}_1^t \mathbf{v}_1}$$

It follows that $\mathbf{r}_1$ must lie in the direction $\mathbf{v}_1$. Further, since $\mathbf{r}_1 - \mathbf{s}_1$ is then orthogonal to $\mathbf{r}_1$, we have that

$$\begin{aligned} \mathbf{r}_1^t \mathbf{s}_1 &= \mathbf{r}_1^t (\mathbf{s}_1 + \mathbf{r}_1 - \mathbf{s}_1) \\ &= \mathbf{r}_1^t \mathbf{r}_1 \end{aligned}$$

and hence that $\mathrm{SIR}_1^{(\mathbf{r}_1)} = \frac{P_1}{\sigma^2} \mathbf{r}_1^t \mathbf{r}_1$. $\Box$

We can therefore think of the decorrelator receiver for user 1 as the orthogonal projection of the received signal onto the orthogonal complement to the interferers' signals. In this way, interfering signals are effectively "nulled out" in an optimal way. We can think of the overall matrix $(S^t S)^{-1} S^t$ as a bank of decorrelating receivers, one given by each row of the matrix.

We now study the performance of the decorrelator in the asymptotic regime in which the processing gain N tends to infinity, the number of users is αN. The following result was also obtained independently in [26].

Theorem 7.2 *Let $\beta_1^{(N)}$ be the (random) SIR of the decorrelating receiver for user 1 when the spreading length is N. Then $\beta_1^{(N)}$ converges to β_1^* in probability as $N \to \infty$, where β_1^* is given by*

$$\beta_1^* = \begin{cases} \frac{P_1(1-\alpha)}{\sigma^2} & \alpha < 1 \\ 0 & \alpha \geq 1 \end{cases}$$

Proof. As in Proposition 7.1, let us denote the subspace orthogonal to the span of $\{\mathbf{s}_2, \mathbf{s}_3 \ldots, \mathbf{s}_K\}$ by V. We note that V has dimension equal to $\max\{N - \mathrm{rank}(S_1), 0\}$, where S_1 is the matrix with columns consisting of the signature sequences of users $2, 3 \ldots, K$. We also note that since the signature sequences are selected randomly, V is a random subspace, independent of the choice of $\mathbf{s}_1$. Finally, as in Proposition 7.1, we denote the decorrelating vector for user 1 by $\mathbf{r}_1$.

The simplest case is $\alpha \geq 1$. In this case, $\frac{\dim(V)}{N} \to 0$, and since $\mathbf{r}_1$ is the orthogonal projection of $\mathbf{s}_1$ onto V (Proposition 7.1), it follows that $\frac{P_1}{\sigma^2}\mathbf{r}_1^t\mathbf{r}_1 \to 0$. Proposition 7.1 implies that the SIR of user 1 tends to zero in this case.

Now consider the case $\alpha \leq 1$. In this case, Bai and Yin [1] show that the smallest eigenvalue of the random matrix $S_1^t S_1$ converges almost surely to a strictly positive number; hence S_1 is almost surely of full rank $K-1$ when L is large. Thus,

$$\frac{\dim(V)}{N} \to 1-\alpha$$

as $N \uparrow \infty$. But by Proposition 7.1, $\mathbf{r}_1$ is the orthogonal projection of $\mathbf{s}_1$ onto V, and we note that $\mathbf{s}_1$ is independent of V. It follows from Lemma 4.2 that $\mathbf{r}_1^t\mathbf{r}_1 \to 1-\alpha$. The theorem then follows from Proposition 7.1. $\square$

We observe that as $\alpha \to 1$, i.e. the number of users per degree of freedom approach 1, the SIR goes to zero. Geometrically, as the dimensionality of the orthogonal complement to the span of the interference decreases to zero, the length of the projection of the desired signal onto this orthogonal complement tends to zero, and so in the limit the projected signal is lost in the background noise. This is the high price paid for ignoring the background noise. In contrast, the MMSE receiver can support more users than the number of degrees of freedom as it takes both the interference and the background noise into account.

By comparing Theorem 7.2 and Theorem 3.1, it can be seen that the effective interference for an interferer on user 1 under the decorrelator is $\frac{P_1}{\beta_1}$, which does not depend on the power of the interferer. The theorem states that the user capacity constraint on the system is $\alpha < 1$.

We also observe that if all users require an SIR of β and employ power control then it is sufficient for each user to be received with power at least $\dfrac{\beta\sigma^2}{1-\alpha}$. Thus, for a given received power constraint $\bar{P}$, the maximum number of users with SIR requirement β supportable is $1-\frac{\beta\sigma^2}{\bar{P}}$. Similarly, for multiple classes of users with SIR requirement β_j and power constraint $\bar{P}_j$ for each class, then the system can support α_j users (per degree of freedom) from each class if

$$\sum_{j=1}^{J} \alpha_j \leq \min_{1\leq j\leq J}\left[1-\frac{\beta_j\sigma^2}{\bar{P}_j}\right]$$

Thus, the user capacity region under the decorrelator is given by:

$$\sum_{j=1}^{J} \alpha_j \leq 1 \tag{21}$$

when there are no power constraints, or equivalently, when the background noise power σ^2 goes to zero. Thus, each user occupies an effective bandwidth of 1 degree of freedom, independent of the value of β.

¿From Theorem 7.2, it can be immediately inferred that the efficiency of a decorrelator in a large system with random spreading sequences is $1-\alpha$ if α, the number of users per degree of freedom, is less than 1 and zero otherwise. Since this does not depend on the background noise power σ^2, this is also the asymptotic efficiency.

It is well known [13] that the MMSE receiver has the same asymptotic efficiency as the decorrelator, and hence the decorrelator is optimal in this sense among all linear receivers. However, comparing eqn. (20) and (21), it can be seen that the user capacity region under the MMSE receiver is strictly larger than that under the decorrelator, even as the background noise goes to zero. In particular, the MMSE receiver can in general accommodate more users than the available degrees of freedom, while the decorrelator cannot. This apparent paradox can be resolved by noting that when $\alpha > 1$, the attained SIR by the decorrelator is zero (Theorem 7.2) while the attained SIR by the MMSE receiver is strictly positive but bounded, as the noise power σ^2 goes to zero. Since the asymptotic efficiency only measures the *rate* at which the SIR goes to infinity as σ^2 goes to zero, they are the same (zero) for both receivers. On the other hand, the user capacity region quantifies the number of users with *fixed* SIR requirements a receiver can accommodate; hence the difference between the decorrelator and the MMSE receiver is captured when we compare their user capacity regions. In practice, users have target SIR requirements and hence the user capacity region characterization seems to be a more natural performance measure than the asymptotic efficiency. In this context, the decorrelator remains sub-optimal even as $\sigma^2 \to 0$, when $\alpha > 1$.

8 Antenna Diversity

In spread-spectrum systems, diversity gain is obtained by spreading over a wider bandwidth. However, there are other ways to obtain diversity benefits in a wireless system. A technique, particularly effective for combating multipath fading, is the use of an *adaptive antenna array* at the receiver. Multipath fading can be very detrimental as the received signal power can drop dramatically due to destructive interference between different paths of the transmitted signal. By placing the antenna elements greater than half the carrier wavelength apart, one can ensure that the received signal fades more or less independently at the different antenna elements. By appropriately weighting, delaying and combining the received signals at the different antenna elements, one can obtain a much more reliable estimate of the transmitted signal than with a single antenna. Such antenna arrays are said to be *adaptive* as the combining depends on the strengths of the received signals at the various antenna elements. This in turn depends on the locations of the users. Moreover, the combining weights will be different for different users, allowing the array to focus on specific users while mitigating the interference from other users. This is so-called *beam-forming.* Using our previous results, it turns out that the user capacity of such antenna array systems can again be characterized by effective bandwidths.

The following is a model for a synchronous multi-access antenna-array system:

$$\mathbf{Y} = \sum_{m=1}^{K} X_m \mathbf{h}_m + \mathbf{W},$$

Here, X_m is the transmitted symbol of the mth user, and $\mathbf{Y}$ is a N-dimensional vector of received symbols at the N antenna elements of the array. The vector $\mathbf{h}_m$ represents the fading of the mth user at each of the antenna array. The entries are complex to incorporate both phase and magnitude information. The vector $\mathbf{W}$ is complex-valued, background Gaussian noise.

The fading is time-varying, as the mobile users move. However, this is usually at a much slower time-scale than the symbol rate of the system. Assuming then that the channel fading of the users can be measured and tracked perfectly at the receiver, we would like to combine the vector of received symbols appropriately to maximize the SIR of the estimates of the transmitted symbols of the users. The optimal linear receiver is clearly the MMSE. Assuming that the fading of each user at each antenna element is independent and identically distributed, we are essentially in the same set-up as for spread-spectrum systems. Thus, for a system with a large number of antenna elements and large number of users, we can treat each of the interfering users as contributing an additive *effective interference*. Under perfect power control, the user capacity is characterized by sharing the N degree of freedom among the users according to their *effective bandwidths* given by the previous expressions for the different receivers. The only difference here is that the N degrees of freedom is obtained by spatial rather than frequency diversity.

These results should be compared with that of Winters et. al. [28], which showed that for a flat Rayleigh fading channel, a combiner which attempts to null out all the interferers will cost one degree of freedom per interferer. This combiner is of course the sub-optimal decorrelator, which we have shown earlier to be very wasteful of degrees of freedom if interferers are weak. It should be noted that while Winters' result holds for the Rayleigh model and any number of antennas, our results hold for *any* fading distribution, but are asymptotic in the number of antennas.

Fig. 4 illustrates the performance of MMSE receiver under a Rayleigh fading environment. It compares the asymptotic limit of the SIR for user 1 given by eqn. 9, as a function of the number of users per antenna element, with actual SIR achieved depending on realizations of the Rayleigh fading. The number of antenna elements is 128. The similarity between Fig. 4 and Fig. 1 further emphasizes the fact that the asymptotic limit does not depend on the interpretation of the $\mathbf{s}_i$'s as spreading sequences or as channel fading.

9 Summary of Results and Conclusions

It is illuminating to compare the effective interference and effective bandwidths of the users in the three cases: the conventional matched filter, the MMSE filter, and the decorrelating filter (Fig. 5 and 6.). The effective interference under MMSE is non-linear, and depends on the received power P of the user to be demodulated as well as the achieved SIR β. The effective interference under the conventional matched filter is simply P_i, the received power of the interferer. Under the decorrelator, the effective interference is $\frac{P}{\beta}$, independent of the actual power of the interferer. The intuition here is that the decorrelator completely nulls out the interferer, no matter how strong or weak it is. The MMSE receiver, on the other hand, is sensitive to the received power of the interferer.

Assuming perfect power control, we can define effective bandwidths which characterize the amount of network resource a user consumes for a given target SIR. The effective bandwidths under the conventional, MMSE and decorrelating receivers are β, $\frac{\beta}{1+\beta}$ and 1 respectively. We note that the conventional receiver is more efficient than the decorrelator when β is small, and far less efficient when β is large. Intuitively, at high SIR requirements, a user has to transmit at high power, thus causing a lot of interference to other users under the conventional receiver. Not surprisingly, since it is by definition optimal, the MMSE filter is the most efficient in all cases. When β is small, it operates more like the conventional receiver, allowing many users per degree of freedom, but when β is large, each user is decorrelated from the rest, much as in the decorrelator receiver, and therefore the interferers can still occupy no more than 1 degree of freedom per interferer. The performance gain afforded by the MMSE receiver over the conventional receiver depends on the SIR at which the system is to be operated, and this in turn depends on the data rate, amount of coding and symbol constellation size. However, due to the superior performance of the MMSE receiver over a wide range of SIR's, it can be seen that it is particularly suitable in a heterogeneous network with multiple traffic types.

In the present paper, we have focused our attention on the simplest possible multi-access CDMA model. There are many possible extensions of this work to study various physical layer and networking layer issues. At the physical layer, an important problem is to understand the performance of multiuser receivers in more realistic scenarios with asynchrony, multipath fading and channel uncertainty. Under these channel imperfections, one can expect an even larger performance gap between the MMSE receiver and the decorrelator. This is because while signals of an interferer arrive from a *single* direction in a synchronous system with perfect channel knowledge, the channel imperfections typically spread the interferer's energy into multiple directions, so that nulling out all the directions would be very wasteful, if at all possible. It is hoped that the extension of the notions of effective interference and effective bandwidth can give insights to the performance gain of the MMSE receiver over both the decorrelator and conventional receiver in these situations. Some results along these lines have been obtained in [9] for asynchronous systems.

At the networking layer, important issues to study include multiple cells and the effect of traffic burstiness such as voice activity. In fact, we believe it is already possible to directly draw some insights into these issues from some of our present results. For example, despite the fact that we have not addressed voice activity in an explicit way, it is clear from Theorem 3.1 and the notion of effective interference that the "averaging of voice activity" property of the conventional receiver will carry over to the MMSE receiver, in contrast to claims made in [23]. Furthermore, we have demonstrated that simple power control mechanisms can be used for resource allocation in almost exactly the same way that they are used in the IS95 standard, and this will clearly also hold in the multiple cell scenario (indeed, see [21]). It is important to note that the effective bandwidth concept we have developed for the MMSE receiver is only valid in the perfectly power-controlled single cell case. However, the concept of effective interference applies with or without perfect power control, and may prove more useful in the multi-cell context.

In a TDMA or FDMA system, the network resource is shared amongst users via disjoint frequency and time slots, and these models provide a simple abstraction of the resource consumed by a user at the physical layer. Such an abstraction allows a clean separation between the physical layer and networking layer resource allocation problems, such as call admissions control, cell handoffs and resource allocation for bursty traffic. It is hoped that the effective bandwidth results in the present paper will be a first step in providing such an abstraction for systems with multiuser receivers. It must be emphasized however that the results reported here are asymptotic in the system size. Thus, a better understanding of the performance fluctuations in finite-size systems is needed before they can be directly applied to real-time control problems such as admission control [7]. We note that a recent paper, [20], provides Central Limit theorems to characterize the performance fluctuations around the asymptotic limits.

Acknowledgments

We would like to thank Kiran for his kind help in obtaining the simulation results, and Professor Jack Silverstein for discussions on random matrix theory.

References

[1] Bai, Z.D. and Y.Q. Yin, " Limit of the smallest eigenvalue of a large dimensional sample covariance matrix", *Annals of Probability*, vol. 21, July, 1993, pp. 1275-1294.

[2] Rashid-Farrokhi, F., L. Tassiulas, and K.J.R. Liu, " Joint Power Control and Beamforming for Capacity Improvement in Wireless Networks with Antenna Arrays," *Proc. IEEE GLOBECOM*, pp. I-555-559, London, Nov. 1996.

[3] Grant, A. and P. Alexander, "Randomly selected spreading sequences for coded CDMA", *Proc. of the 5th Int. Spread Spectrum Techniques and Applications*, Sept., 1996, pp. 54-57.

[4] Hanly S.V. "Information Capacity of Radio Networks" *PhD Thesis, Cambridge University*, Aug. 1993.

[5] Hanly, S.V. "Capacity and power control in spread spectrum macrodiversity radio networks" *IEEE Trans. on Communications, Vol. 44, No. 2* Feb., 1996.

[6] Honig, M., U. Madhow and S. Verdu, "Blind adaptive multiuser detection", *IEEE Trans. on Information Theory*, July, 1995, pp.944-960.

[7] Honig, M. and W. Veerakachen, " Performance Variability of Linear Multiuser Detection for DS-CDMA", *Proc. of Vehicular Technology Conference*, 1996.

[8] Kim, J.B. and M. Honig, "Outage probability of multi-code DS-CDMA with linear interference suppression", to be presented at MILCOM 1998.

[9] Kiran and D. Tse, "Effective Bandwidths and Effective Interference for Linear Multiuser Receivers in Asynchronous Channels", *Proceedings of IEEE Information Theory Workshop, Killarney*, Ireland, June 1998, p. 141-142; full paper in preparation and latest version can be found at http://www.eecs.berkeley.edu/~dtse.

[10] Kumar, P. and J. Holtzman, "Power control for a spread-spectrum system with multiuser receivers", *Proc. 6th IEEE PIMRC*, Sept. 1995, pp.955-959.

[11] Lupas, R. and S. Verdu, "Linear multiuser detectors for synchronous code-division multiple access", *IEEE Trans. on Information Theory*,IT-35, Jan., 1989, pp.123-136.

[12] Lupas, R. and S. Verdu, "Near-far resistance of multiuser detectors in asynchronous channels," *IEEE Trans. on Communications*, COM-38, Apr. 1990, pp. 496-508.

[13] Madhow, U. and M. Honig, "MMSE interference suppression for direct-sequence spread-spectrum CDMA, *IEEE Trans. on Communications*, Dec., 1994,pp.3178-3188.

[14] Madhow,U. and M. Honig, "MMSE Detection of Direct-Sequence CDMA Signals: Performance Analysis for Random Signature Sequences", submitted to *IEEE Trans. on Information Theory.*

[15] Marcenko, V.A., Pastur L.A. (1967) "Distribution of eigenvalues for some sets of random matrices" *Math. USSR-Sb* 1. pp457–483.

[16] Rapajic, P. and B. Vucetic, "Adaptive receiver structures for asynchronous CDMA systems, *IEEE JSAC*, May, 1994, pp. 685-697.

[17] Rupf, M., F. Tarkoy and J. Massey, "User-separating demodulation for code-division multiple access systems",*IEEE JSAC*, June,1994, pp.786-795.

[18] Silverstein J.W., Bai Z.D. "On the empirical distribution of eigenvalues of a class of large dimensional random matrices", *Journal of Multivariate Analysis* 54(2) (1995), pp. 175-192.

[19] Tse, D. and S.V. Hanly, "Multiuser demodulation: effective interference, effective bandwidth and capacity", Proceedings of Allerton Conference, Oct. 1997.

[20] Tse, D. and O. Zeitouni, "Fluctuations of performance of multiuser receivers under random signature sequences", submitted to IEEE Trans. on Information Theory, Sept. 1998.

[21] Ulukus,S. and R.D. Yates, "Adaptive power control and MMSE interference suppression", *Wireless Networks*, accepted for publication.

[22] Varanasi, M.K. and T. Guess, "Optimum decision feedback multiuser equalization and successive decoding achieves the total capacity of the Gaussian multiple-access channel", Proc. Asilomar Conf. on Signals, Systems and Computers, 1997.

[23] Vembu,S. and A.J. Viterbi, "Two Philosophies in CDMA - a comparison", *Proc. IEEE Vehicular Technology Conference*, 1996, pp. 869-873.

[24] Verdu,S., "Minimum probability of error for asynchronous Gaussian channels", *IEEE Trans. on Information Theory*, IT-32, Jan.,1986, pp.85-96.

[25] Verdu, S., "Optimum multiuser asymptotic efficiency ", *IEEE Trans. on Comm.*, COM-34, Sept.1996, pp. 890-897.

[26] Verdu, S. and S. Shamai (Shitz), "Multiuser detection with random spreading and error-correction codes: fundamental limits", presented at Allerton Conference, Oct., 1997.

[27] Viswanath, P., V. Anantharam and D. Tse, "Optimal sequences, power control and capacity of spread-spectrum systems with multiuser receivers", *Proceedings of International Symposium on Information Theory*, Cambridge MA, Aug. 1998, p.121; full paper submitted to *IEEE. Trans. on Information Theory* and latest version can be found at http://www.eecs.berkeley.edu/~dtse.

[28] Winters J.H., Salz J., Gitlin R. "The impact of antenna diversity on the capacity of wireless communications systems" *IEEE Trans. on Comm.*, Vol. 42, No. 2/3/4, 1994, pp. 1740-1751.

[29] Xie, Z., R. Short and C. Rushforth, "A family of suboptimum detectors for coherent multiuser communications", IEEE JSAC, May, 1990, pp.683-690.

Appendices

A A Heuristic Derivation of Theorem 3.1

In this appendix, we gave an alternative and heuristic derivation of expression (4), without invoking the Stieltjes transform characterization of the limiting eigenvalue distribution (11). The goal is to shed more light into the form of the expression and to provide some intuition about the decoupling of the interference from different users and the consequent linearity in the effective bandwidth characterization of the capacity region. The derivation given here makes use of some ideas developed in [27] but is self-contained.

We first give a formula for the MMSE receiver and the associated SIR under the MMSE receiver, alternative but equivalent to (2) and (3). First recall the channel model in matrix form:

$$\mathbf{Y} = S\mathbf{X} + \mathbf{W}$$

where S is the matrix the columns of which are the signature sequences of the users. If $\hat{X}$ is the vector MMSE estimate of $\mathbf{X}$, a direct application of the orthogonality principle $E[(\hat{X} - \mathbf{X})^t\mathbf{Y}] = 0$ yields

$$\hat{X} = DS^t \left[SDS^t + \sigma^2 I\right]^{-1} \mathbf{Y}$$

and the covariance matrix of the error $\epsilon \equiv \hat{X} - \mathbf{X}$ is given by

$$K_\epsilon = D - DS^t \left[SDS^t + \sigma^2 I\right]^{-1} SD \tag{22}$$

where $D \equiv \mathrm{diag}(P_1, \ldots, P_K)$ is the covariance matrix of $\mathbf{X}$. Right multiplying the above equation with D^{-1} and taking the trace of both sides, we get:

$$\begin{aligned}
& \mathrm{trace}(K_\epsilon D^{-1}) && (23)\\
= \;& K - \mathrm{trace}\left(DS^t \left[SDS^t + \sigma^2 I\right]^{-1} S\right) \\
= \;& K - \mathrm{trace}\left(SDS^t \left[SDS^t + \sigma^2 I\right]^{-1}\right) && \text{using the fact } \mathrm{trace}(AB) = \mathrm{trace}(BA)\\
= \;& K - \sum_{i=1}^{N} \frac{\lambda_i}{\lambda_i + \sigma^2} && (24)
\end{aligned}$$

where λ_i's are the eigenvalues of the matrix SDS^t. If we let

$$\mathrm{MMSE}_i \equiv \frac{E[(\hat{X} - X_i)^2]}{P_i}$$

be the (normalized) minimum mean-square error for user i, then eqn. (24) says that

$$\sum_{i=1}^{K} \mathrm{MMSE}_i = K - \sum_{i=1}^{N} \frac{\lambda_i}{\lambda_i + \sigma^2} \tag{25}$$

Now it is well known that the SIR $\beta_i^{(N)}$ and the MMSE error are related as follows (see eg. [13]):

$$\mathrm{MMSE}_i = \frac{1}{1+\beta_i^{(N)}}. \tag{26}$$

Substituting this into eqn. (25) and rearranging terms, we obtain,

$$\frac{1}{N}\sum_{i=1}^{K}\frac{\beta_i^{(N)}}{1+\beta_i^{(N)}} = 1-\sigma^2\frac{1}{N}\sum_{i=1}^{N}\frac{1}{\lambda_i+\sigma^2} \tag{27}$$

So far, we have not introduced any probabilistic model for the spreading sequences, and this equation holds for every choice of the sequences and for every N. Now, let us assume the sequences are randomly chosen, and each component is i.i.d., and consider what happens when $K, N \to \infty$, $\frac{K}{N} \to \alpha$ and the empirical distribution of the received powers converge to F. The right-hand side of the above equation converges to

$$1-\sigma^2\int_0^\infty \frac{1}{\lambda+\sigma^2}dG^*(\lambda)$$

where G^* is the limiting eigenvalue distribution of SDS^t, and by Lemma 4.3, $\beta_i^{(N)}$ converges to

$$\beta_i^* = P_i\int_0^\infty \frac{1}{\lambda+\sigma^2}dG^*(\lambda)$$

Expressing everything in terms of β_1^*, one can expect that the limiting form of eqn. (27) to become [3]:

$$\alpha\int_0^\infty \frac{\frac{P\beta_1^*}{P_1}}{1+\frac{P\beta_1^*}{P_1}}dF(P) = 1-\frac{\sigma^2\beta_1^*}{P_1}$$

Dividing throughout by $\frac{\beta_1^*}{P_1}$ and rearranging terms gives us the desired fixed-point equation (4):

$$\beta_1^* = \frac{P_1}{\sigma^2+\alpha\int_0^\infty \frac{P_1 P dF(P)}{P_1+P\beta_1^*}}.$$

This development allows us to understand the linearity of the effective bandwidth characterization of the capacity region. First, consider the simpler case when $\sigma^2 \to 0$, i.e. no power constraint. Assuming that the spreading sequences span a space of dimension $\min\{K, N\}$. Then precisely $\min\{K, N\}$ of the eigenvalues λ_i's are non-zero. Eqn. (25) becomes:

$$\sum_{i=1}^{K}\mathrm{MMSE}_i = K-\min\{K,N\}$$

Note that the total MMSE of the users is a constant, irrespective of the received powers of the users. Since the SIR of a user is a function of the MMSE error, this is the reason

[3]This is the heuristic step of the derivation.

for the linearity of the capacity region with no power constraint. For the case when there are power constraints (i.e. $\sigma^2 \neq 0$), the situation is more subtle. Asymptotically, the right-hand side of eqn. (25) depends on the received powers of the users only through

$$\int_0^\infty \frac{1}{\lambda + \sigma^2} dG^*(\lambda)$$

which can be interpreted as the SIR achieved by a user with unit received power.

B Proofs

Proof of Proposition 3.3:

Now,

$$\beta_{1,MF}^{(N)} = \frac{P_1(\mathbf{s}_1^t\mathbf{s}_1)^2}{\sigma^2(\mathbf{s}_1^t\mathbf{s}_1)^2 + \sum_{i=2}^K P_i(\mathbf{s}_1^t\mathbf{s}_i)^2}.$$

Clearly $(\mathbf{s}_1^t\mathbf{s}_1)^2$ converges to 1 in probability, by the weak law of large numbers. We now look at the interference from the other users. Consider a scaled version of the cross-correlation between the signature sequences of user 1 and user i:

$$\xi_i \equiv \frac{1}{\sqrt{N}} \sum_{k=1}^N V_{1,k} V_{i,k} \qquad i = 1, 2, \ldots, K$$

where $\mathbf{s}_i = \frac{1}{\sqrt{N}}(V_{i1}, \ldots, V_{iN})^t$. Also, define $\bar{P}^{(K)} = \frac{1}{K}\sum_{i=1}^K P_i$. Let us first condition on a random realization of powers $P_1, P_2, \ldots$. Then

$$\text{Var}\,(\sum_{i=2}^K P_i\xi_i^2 | P_1, P_2, \ldots) = \sum_{i=2}^K \sum_{j=2}^K \mathbb{E}\left[\left(\frac{P_i}{N}(\sum_{k_1} V_{1,k_1}V_{i,k_1})^2 - \bar{P}^{(K)}\right)\left(\frac{P_j}{N}(\sum_{k_2} V_{1,k_2}V_{j,k_2})^2 - \bar{P}^{(K)}\right) | P\right] \tag{28}$$

By expanding out the product, we obtain that for $i \neq j$, the term

$$\mathbb{E}\left[\left(\frac{P_i}{N}(\sum_{k_1} V_{1,k_1}V_{i,k_1})^2 - \bar{P}\right)\left(\frac{P_j}{N}(\sum_{k_2} V_{1,k_2}V_{j,k_2})^2 - \bar{P}\right) | P_1, P_2, \ldots\right]$$

equals

$$\frac{P_iP_j}{N^2}\mathbb{E}[(\sum_k V_{1,k}V_{i,k})^2(\sum_k V_{1,k}V_{j,k})^2] - \frac{\bar{P}P_j}{N}\mathbb{E}[(\sum_k V_{1,k}V_{j,k})^2] - \frac{\bar{P}P_i}{N}\mathbb{E}[(\sum_k V_{1,k}V_{i,k})^2] + \bar{P}^2 \tag{29}$$

Expanding out the first term on the right hand side, we obtain

$$\sum_{k_1}\sum_{k_2}\sum_{k_3}\sum_{k_4} \mathbb{E}[V_{i,k_1}V_{i,k_3}]\mathbb{E}[V_{j,k_2}V_{j,k_4}]\mathbb{E}[V_{1,k_1}V_{1,k_2}V_{1,k_3}V_{1,k_4}]$$

Now each of these expectations is zero except when $k_1 = k_3$ and $k_2 = k_4$, so it reduces to $\sum_{k_1}\sum_{k_2} \mathbb{E}[V_{1,k_1}^2 V_{1,k_2}^2]$. Now, $N(N-1)$ of these terms are unity, and N are $\mathbb{E}[V_{1,1}^4]$, which is $O(1)$, so it follows that the first term on the right hand side of (29) is $P_iP_j + O(1/N)$. In a similar manner, the second term can be shown to be $P_j\bar{P}$ and the third term is $P_i\bar{P}$.

Returning to the expansion of (28), we note that for all $i = 2, \ldots, K$,

$$\mathbb{E}\left[\left(\frac{P_i}{N}(\sum_{k_1} V_{1,k_1}V_{i,k_1})^2 - \bar{P}\right)^2 |P_1, P_2, \ldots\right]$$

equals

$$\frac{P_i^2}{N^2}\mathbb{E}[(\sum_k V_{1,k}V_{i,k})^4] - \frac{P_i\bar{P}}{N}\mathbb{E}[(\sum_k V_{1,k}V_{i,k})^2] + \bar{P}^2 \tag{30}$$

Expanding out the first term we obtain

$$\sum_{k_1}\sum_{k_2}\sum_{k_3}\sum_{k_4} \mathbb{E}[V_{1,k_1}V_{1,k_2}V_{1,k_3}V_{1,k_4}V_{i,k_1}V_{i,k_2}V_{i,k_3}V_{i,k_4}]$$

and each of these expectations is zero, unless $k_1 = k_2$ and $k_3 = k_4$ or $k_1 = k_3$ and $k_2 = k_4$ or $k_1 = k_4$ and $k_2 = k_3$. In each of these nonzero cases, the expectations are $O(1)$ and there are $O(N^2)$ of them, so the first term of (30) is $O(1)$. Similarly, for the other two terms. We conclude that

$$\text{Var}\,(\frac{1}{K}\sum_{i=2}^{K} P_i\xi_i^2|P_1, P_2, \ldots) = \frac{1}{K^2}\sum_{i=2}^{K}\sum_{j=2}^{K}(P_iP_j - P_j\bar{P}^{(K)} - P_i\bar{P}^{(K)} + (\bar{P}^{(K)})^2) + O(1/N)$$

as $N \uparrow \infty$. But by our assumption that the empirical distribution function of powers converges to a deterministic limit, it follows that

$$\frac{1}{K^2}\sum_{i=2}^{K}\sum_{j=2}^{K}(P_iP_j - P_j\bar{P}^{(K)} - P_i\bar{P}^{(K)} + (\bar{P}^{(K)})^2) \to 0$$

and hence that for any $\epsilon > 0, \limsup_K \text{ Var}\,(\frac{1}{K}\sum_{i=2}^{K} P_i\xi_i^2|P_1, P_2, \ldots) < \epsilon$, and this is true for any realization $P_1, P_2, \ldots$. Hence, for all $\epsilon > 0, \limsup_K \mathbb{E}[(\frac{1}{K}\sum_{i=2}^{K} P_i\xi_i^2 - \bar{P}^{(K)})^2] < \epsilon$. But $\bar{P}^{(K)} \to \int_0^\infty PdF(P)$, which implies mean-square convergence of $\frac{1}{K}\sum_{i=2}^{K} P_i\xi_i^2$ to $\int_0^\infty PdF(P)$, and hence convergence in probability. So we have

$$\sum_{i=2}^{K} P_i(\mathbf{s}_1^t\mathbf{s}_i)^2 = \frac{1}{N}\sum_{i=2}^{K} P_i\xi_i^2 \to \alpha\int_0^\infty PdF(P)$$

in probability. We conclude that

$$\beta_{1,MF}^{(N)} \to \frac{P_1}{\sigma^2 + \alpha\int_0^\infty PdF(P)} \text{ in probability}$$

□

Proof of Lemma 4.2:

Let $Y \equiv \|Q\mathbf{X}\|^2$. We compute the first and second moments of Y conditional on an arbitrary realization of $\mathcal{Q} = (q_{ij})$.

$$
\begin{aligned}
E[Y|Q] &= E\left[\sum_{i=1}^{m}\left(\sum_{j=1}^{n} q_{ij}V_j\right)^2\right] \\
&= E\left[\sum_{i=1}^{m}\sum_{j=1}^{n}\sum_{k=1}^{n} q_{ij}q_{ik}V_jV_k\right] \\
&= \sum_{i=1}^{m}\sum_{j=1}^{n} q_{ij}^2 \\
&= m, \\
E[Y^2|Q] &= E\left[\left\{\sum_{i=1}^{m}\left(\sum_{j=1}^{n} q_{ij}V_j\right)^2\right\}^2\right] \\
&= E\left[\sum_{i=1}^{m}\sum_{j=1}^{m}\sum_{k=1}^{n}\sum_{l=1}^{n}\sum_{r=1}^{n}\sum_{s=1}^{n} q_{ik}q_{il}q_{jr}q_{js}V_kV_lV_rV_s\right]
\end{aligned}
$$

Since the V_i's are independent and zero mean, the terms in the expectation above are zero whenever it has one random variable which has a different index than the other three. Hence,

$$
\begin{aligned}
& E[Y^2|Q] \\
= & \sum_{i=1}^{m}\sum_{j=1}^{m}\sum_{k=1}^{n} q_{ik}^2q_{jk}^2E[V_k^4] + \sum_{i=1}^{m}\sum_{j=1}^{m}\sum_{k\neq r} q_{ik}^2q_{jr}^2E[V_k^2]E[V_r^2] + 2\sum_{i=1}^{m}\sum_{j=1}^{m}\sum_{k\neq l} q_{ik}q_{il}q_{jk}q_{jl}E[V_k^2]E[V_l^2] \\
= & \sum_{i=1}^{m}\sum_{j=1}^{m}\sum_{k=1}^{n} q_{ik}^2q_{jk}^2\left(E[V_k^4]-3\right) + \sum_{i=1}^{m}\sum_{j=1}^{m}\sum_{k=1}^{n}\sum_{r=1}^{n} q_{ik}^2q_{jr}^2 + 2\sum_{i=1}^{m}\sum_{j=1}^{m}\sum_{k=1}^{n}\sum_{l=1}^{n} q_{ik}q_{il}q_{jk}q_{jl} \\
= & \sum_{i=1}^{m}\sum_{j=1}^{m}\sum_{k=1}^{n} q_{ik}^2q_{jk}^2\left(E[V_k^4]-3\right) + \sum_{i=1}^{m}\sum_{j=1}^{m}(\sum_{k=1}^{n} q_{ik}^2)(\sum_{r=1}^{n} q_{jr}^2) + 2\sum_{i=1}^{m}\sum_{j=1}^{m}(\sum_{k=1}^{n} q_{ik}q_{il})^2 \\
= & \sum_{k=1}^{n}(\sum_{i=1}^{m} q_{ik}^2)^2\left(E[V_k^4]-3\right) + m^2 + 2m
\end{aligned}
$$

the last step using the orthonormality of the rows of Q. Now if we add orthonormal rows to Q to construct a n by n orthogonal matrix Q', then the columns of Q' are orthonormal. This implies that for every column k,

$$\sum_{i=1}^{m} q_{ik}^2 \leq 1$$

Hence

$$E[Y^2|Q] \leq n|E[V_1^4]-3| + m^2 + 2m$$

and

$$E[(Y-m)^2|Q] \leq n|E[V_1^4]-3|+2m \tag{31}$$

and hence

$$E[(Y-m)^2] \leq n|E[V_1^4]-3|+2m$$

Using Chebychev's inequality, we have for every $\epsilon > 0$,

$$\begin{aligned}
\Pr\left[|\frac{Y}{n} - \frac{m}{n}| > \epsilon\right] &\leq \frac{E[(Y-m)^2]}{n^2\epsilon^2} \\
&\leq \frac{n|E[V_1^4]-3|+2m}{n^2\epsilon^2} \\
&\leq \frac{|E[V_1^4]-3|+2}{\epsilon^2} \cdot \frac{1}{n}.
\end{aligned}$$

Picking the constant $C \equiv \frac{|E[V_1^4]-3|+2}{\epsilon^2}$ yields the desired result.

□

Proof of Lemma 4.3:

¿From eqn. (3),

$$\beta_1^{(N)} = \mathbf{s}_1^t(S_1 D_1 S_1^t + \sigma^2 I)^{-1}\mathbf{s}_1 P_1$$

Let $\lambda_1^{(N)}, \ldots, \lambda_N^{(N)}$ be the eigenvalues of $S_1 D_1 S_1^t$. Write $S_1 D_1 S_1^t + \sigma^2 I$ as $Q^t \Lambda Q$, where $\Lambda = \mathrm{diag}(\lambda_1^{(N)} + \sigma^2, \ldots, \lambda_N^{(N)} + \sigma^2)$. Let $\mathbf{u}^{(\mathbf{N})} = Q\mathbf{s}_1$. Then

$$\beta_1^{(N)} = \sum_{i=1}^{N} \frac{[u_i^{(N)}]^2 P_1}{\lambda_i^{(N)} + \sigma^2}$$

Fix a $\delta_1 > 0$. Pick a finite partition $\mathcal{I} = \{I_1, I_2, \ldots, I_M\}$ of $(0, \infty)$ such that

$$\sum_{k=1}^{M} G^*(I_k)\frac{P_1}{l(I_k)+\sigma^2} - \int_0^\infty \frac{P_1}{\lambda+\sigma^2} dG^*(\lambda) < \delta_1 \tag{32}$$

and

$$\int_0^\infty \frac{P_1}{\lambda+\sigma^2} dG^*(\lambda) - \sum_{k=1}^{M} G^*(I_k)\frac{P_1}{r(I_k)+\sigma^2} < \delta_1 \tag{33}$$

where $l(I_k), r(I_k)$ are the left and right endpoints of the interval I_k respectively.

Let G_N be the empirical distribution of the eigenvalues of $S_1 D_1 S_1^t$. Fix $\delta_2 > 0$, and consider the events:

$$
\begin{aligned}
E_1 &= \left\{\left|\sum_{\{i:\lambda_i^{(N)}\in I_k\}} (u_i^{(N)})^2 - G_N(I_k)\right| < \frac{\delta_2}{M} \quad \text{for all } k=1,\ldots,M\right\} \\
E_2 &= \left\{|G_N(I_k) - G^*(I_k)| < \frac{\delta_2}{M} \text{for all } k=1,\ldots,M\right\}
\end{aligned}
$$

If both events E_1 and E_2 hold, then we have

$$
\begin{aligned}
\beta_1^{(N)} &= \sum_{i=1}^{N} \frac{u_i^{(N)} P_1}{\lambda_i^{(N)} + \sigma^2} \\
&\le \sum_{k=1}^{M} \Big(\sum_{\{i:\lambda_i^{(N)}\in I_k\}} (u_i^{(N)})^2 \Big) \frac{P_1}{l(I_k)+\sigma^2} \\
&\le P_1 \sum_{k=1}^{M} \frac{G^*(I_k) + 2\frac{\delta_2}{M}}{l(I_k)+\sigma^2} \\
&\le \int_0^\infty \frac{P_1}{\lambda+\sigma^2} dG^*(\lambda) + \delta_1 + \frac{2\delta_2}{\sigma^2} \qquad \text{from eqn. (32)}
\end{aligned}
$$

and similarly,

$$
\begin{aligned}
\beta_1^{(N)} &\ge \sum_{k=1}^{M} \Big(\sum_{\{i:\lambda_i^{(N)}\in I_k\}} (u_i^{(N)})^2 \Big) \frac{P_1}{r(I_k)+\sigma^2} \\
&\ge P_1 \sum_{k=1}^{M} \frac{G^*(I_k) - 2\frac{\delta_2}{M}}{l(I_k)+\sigma^2} \\
&\ge \int_0^\infty \frac{P_1}{\lambda+\sigma^2} dG^*(\lambda) - \delta_1 - \frac{2\delta_2}{\sigma^2} \qquad \text{from eqn. (32)}
\end{aligned}
$$

Hence, given any $\epsilon > 0$, one can pick $\delta_1, \delta_2 > 0$ and M such that:

$$
\left|\beta_i^{(N)} - \int_0^\infty \frac{P_1}{\lambda+\sigma^2} dG^*(\lambda)\right| < \epsilon
$$

whenever events E_1 and E_2 occur. Thus, by the union of events bound,

$$
\Pr\left[\left|\beta_i^{(N)} - \int_0^\infty \frac{P_1}{\lambda+\sigma^2} dG^*(\lambda)\right| > \epsilon\right] \le \Pr[E_1^c] + \Pr[E_2^c] \tag{34}
$$

and

$$
\Pr[E_2^c] \le \sum_{k=1}^{M} \Pr\left[|G_N(I_k) - G^*(I_k)| > \frac{\delta_2}{M}\right]
$$

By Theorem 4.1, each of the probabilities in the sum go to zero as $N \to \infty$. Hence $\Pr[E_2^c] \to 0$. Now,

$$
\begin{aligned}
\Pr[E_1^c] &= \Pr\left[\left|\sum_{\{i:\lambda_i^{(N)}\in I_k\}} (u_i^{(N)})^2 - G_N(I_k)\right| > \frac{\delta_2}{M} \quad \text{for some } k\right] \\
&\le \sum_{k=1}^{M} \Pr\left[\left|\sum_{\{i:\lambda_i^{(N)}\in I_k\}} (u_i^{(N)})^2 - G_N(I_k)\right| > \frac{\delta_2}{M}\right] \\
&\le \sum_{k=1}^{M} \frac{C}{N} \qquad \text{by Lemma 4.2} \\
&= \frac{MC}{N}
\end{aligned}
$$

which approaches 0 as $N \to \infty$. Hence, from eqn. (34), we can conclude that

$$\beta_1^{(N)} \to \int_0^\infty \frac{P_1}{\lambda + \sigma^2} dG^*(\lambda) \qquad \text{in probability.}$$

□

Biographies:

David Tse received the B.A.Sc. degree in systems design engineering from University of Waterloo, Canada, in 1989, and the M.S. and Ph.D. degrees in electrical engineering from Massachusetts Institute of Technology in 1991 and 1994 respectively. From 1994 to 1995, He was a postdoctoral member of technical staff at A.T. & T. Bell Laboratories. Since 1995, he has been an assistant professor in the Department of Electrical Engineering and Computer Sciences in the University of California at Berkeley. He received a 1967 NSERC 4-year graduate fellowship from the government of Canada in 1989, a NSF CAREER award in 1998, and the Best Paper Award at the Infocom 1998 conference for his work with Stephen Hanly. His current research interests include resource allocation problems for broadband and wireless networks, and information theory.

Stephen Hanly received the B.Sc.(Hon) degree in mathematics and computer Science, and the M.Sc. in mathematics, from the University of Western Australia, in 1988 and 1990, respectively. In 1989, he was awarded a Commonwealth Scholarship to study in Britain for three years, and in 1994, he received the Ph.D. degree in mathematics from Cambridge University. From 1993 to 1995, he was a postdoctoral member of technical staff at A.T. & T. Bell Laboratories. From 1996 to 1997, he was a research fellow, and is now a lecturer, in the Department of Electrical Engineering, at the University of Melbourne, Australia. He is a co-recipient of the 1998 Infocom best paper award, for joint work with David Tse. His research interests are in mobile radio, information theory, and resource allocation problems in networks.

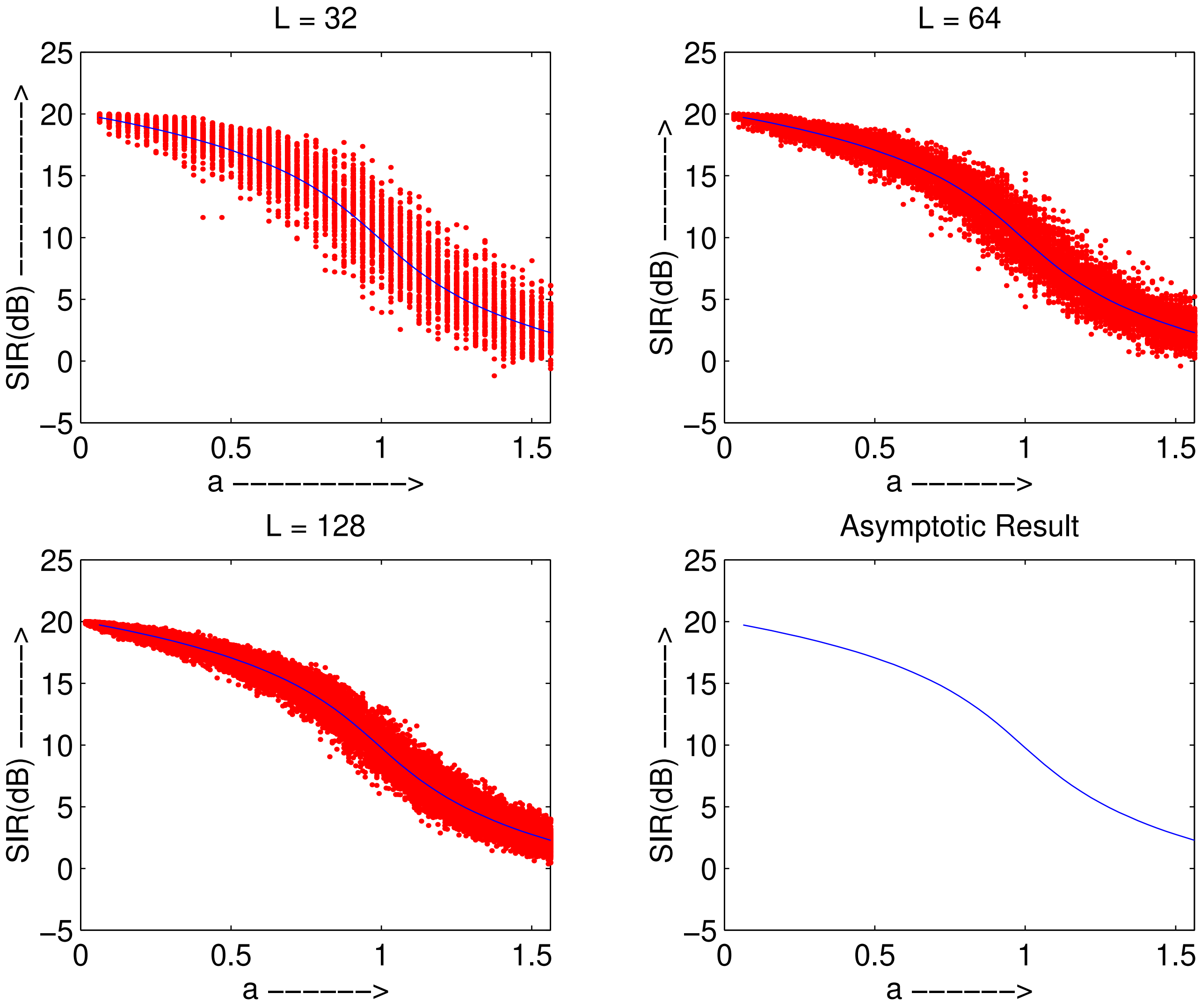

Figure 1: Randomly generated MMSE SIR's for user 1 compared to asymptotic limit eqn. (8) in the equal-power regime, for $N = 32, 64, 128$. Here, $\frac{P}{\sigma^2} = 20\text{dB}$.

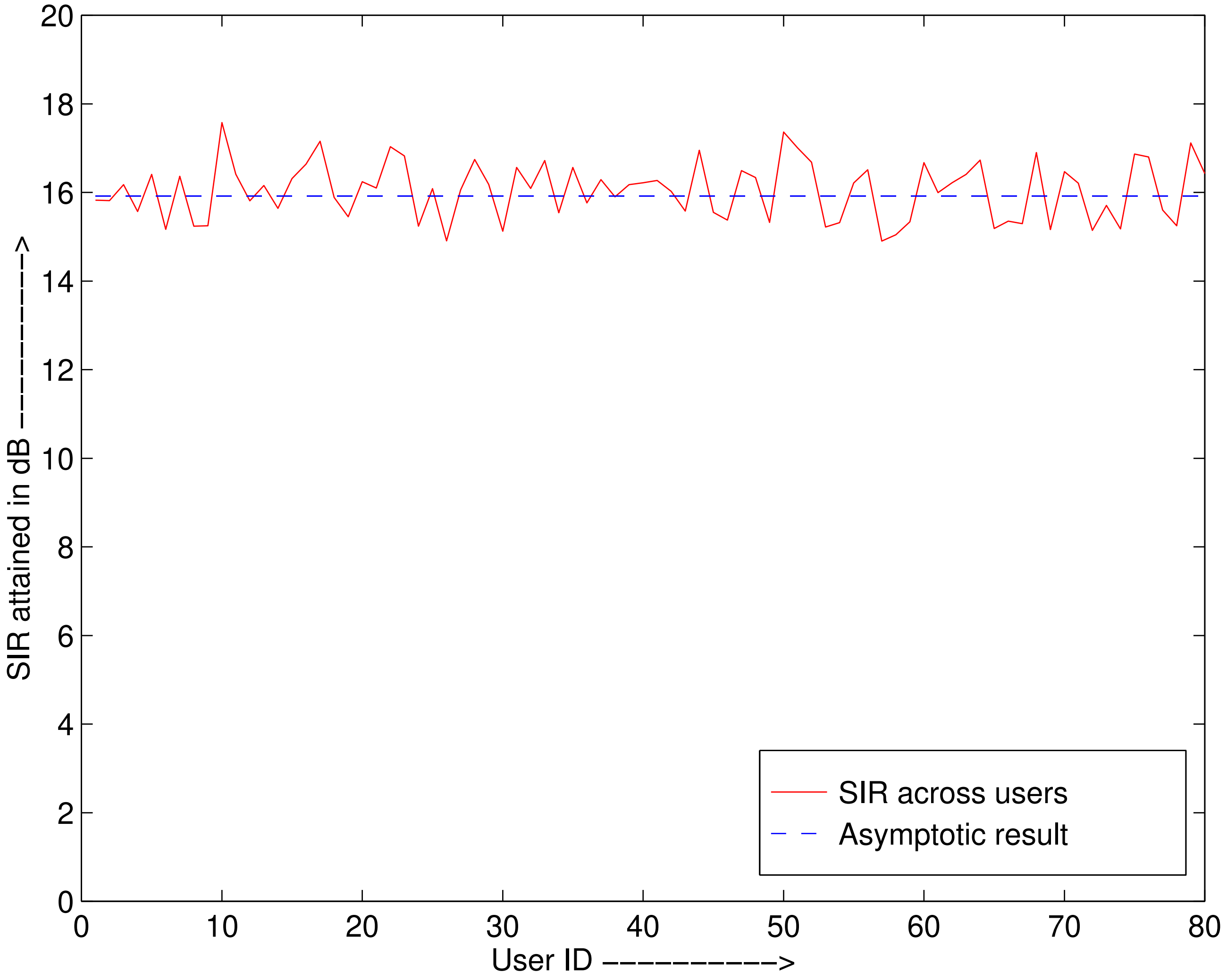

Figure 2: Randomly generated MMSE SIR's across users for one realization of the spreading sequences. Here, spreading length $N = 128$, number of users $K = 80$ and $\frac{P}{\sigma^2} = 20$dB.

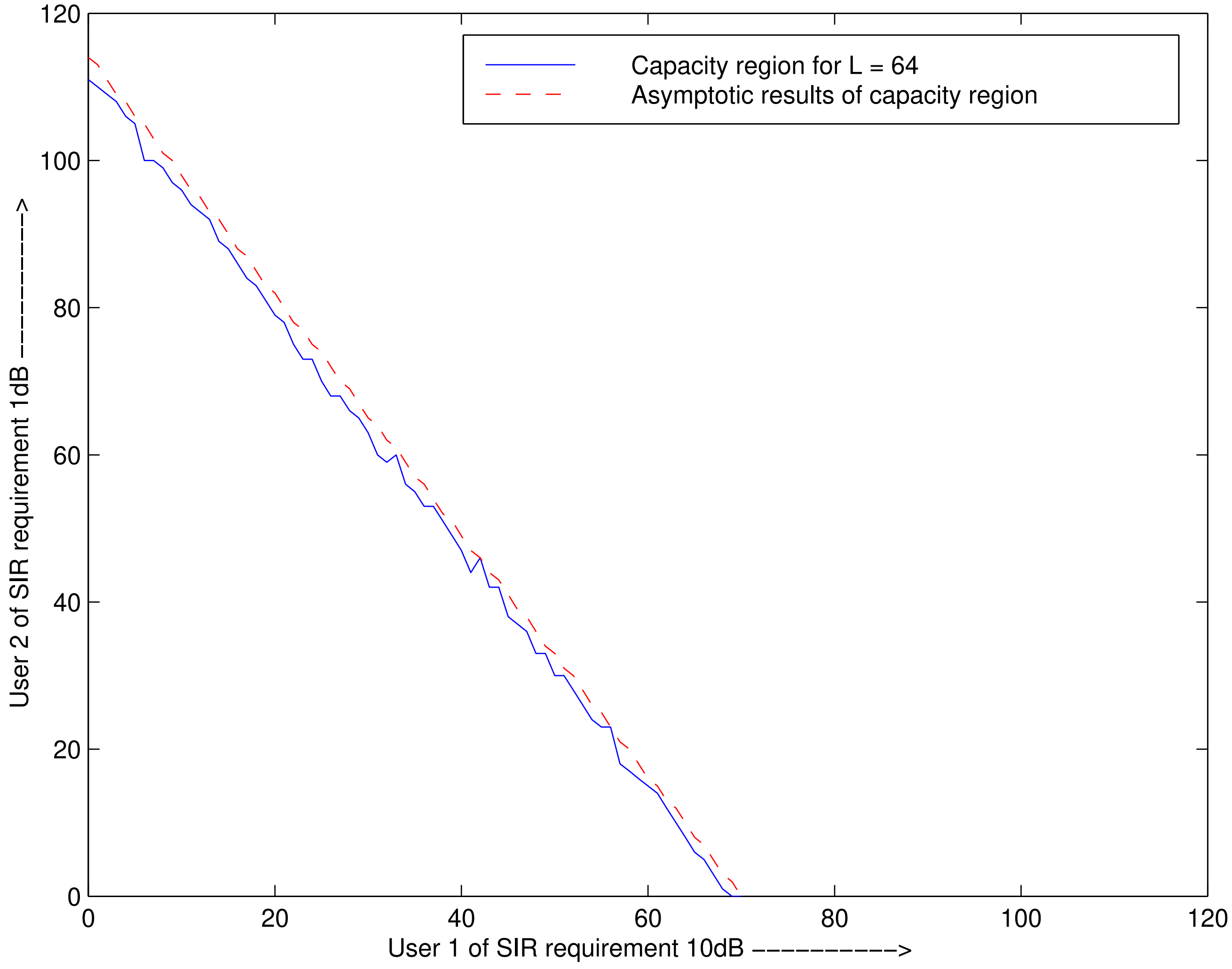

Figure 3: User capacity region for two classes of users, with $\frac{P_1}{\sigma^2} = 29\text{dB}, \frac{P_2}{\sigma^2} = 20\text{dB}$

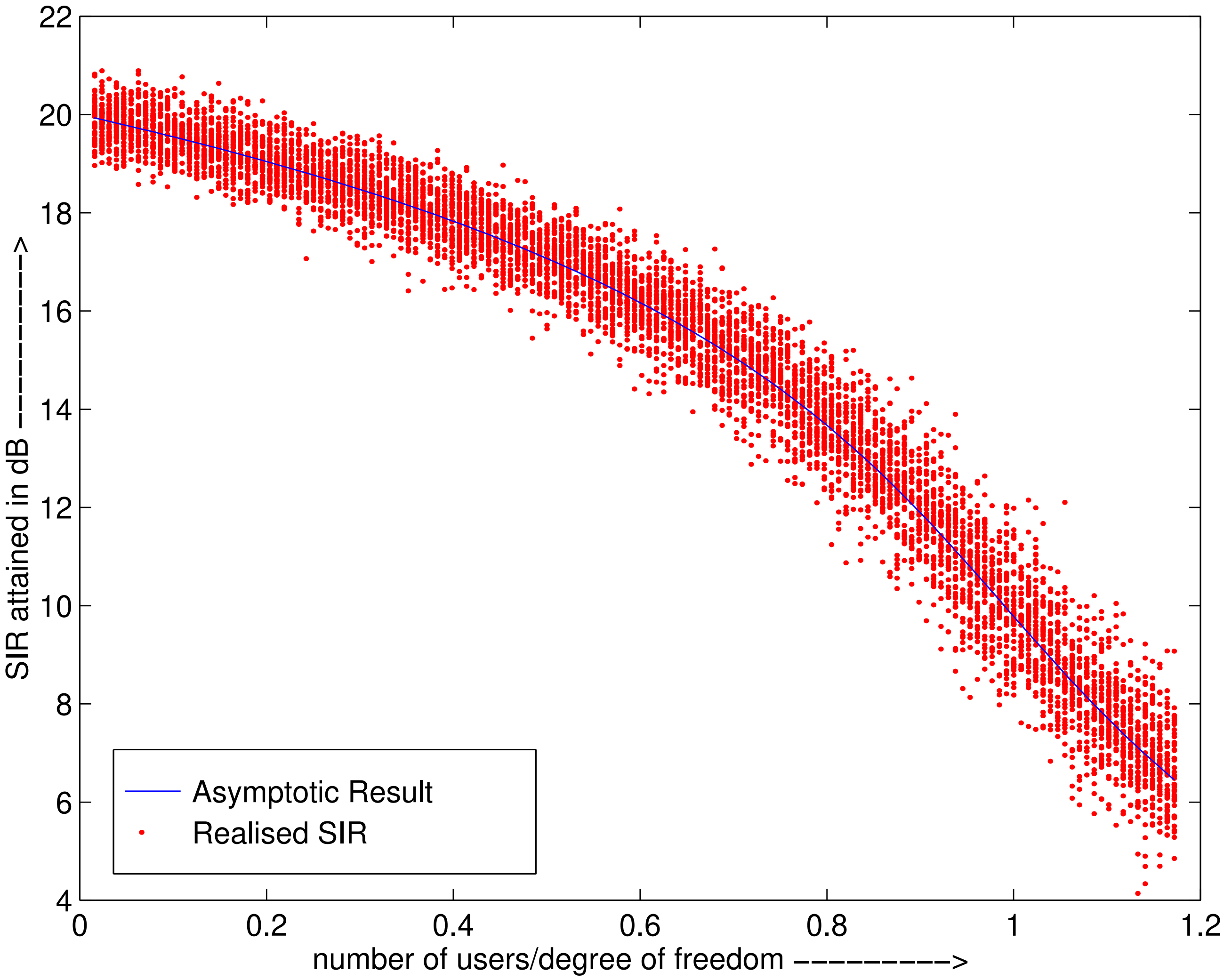

Figure 4: Random SIR's for user 1 in Rayleigh fading environment, compared to asymptotic limit eqn. (8) Here, $\frac{P}{\sigma^2} = 20$dB.

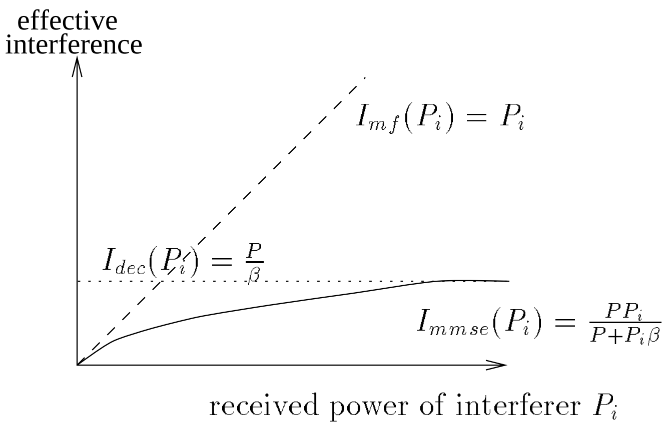

Figure 5: Effective interference for the 3 receivers as a function of interferer's received power P_i. Here,P is the received power of the user to be demodulated, and β is the SIR achieved.

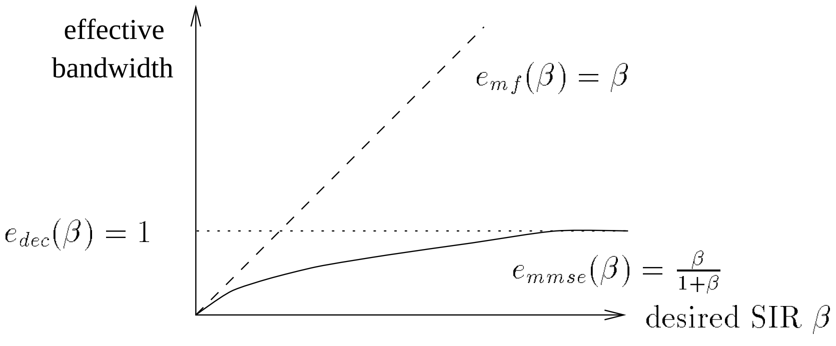

Figure 6: Effective bandwidths for 3 receivers as a function of SIR.