# PERFORMANCE OF LINEAR MULTIUSER RECEIVERS IN RANDOM ENVIRONMENTS

by

David Tse and Ofer Zeitouni

Memorandum No. UCB/ERL M99/4

15 January 1999

COVER

# PERFORMANCE OF LINEAR MULTIUSER RECEIVERS IN RANDOM ENVIRONMENTS

by

David Tse and Ofer Zeitouni

Memorandum No. UCB/ERL M99/4

15 January 1999

**ELECTRONICS RESEARCH LABORATORY**

College of Engineering
University of California, Berkeley
94720

# Performance of Linear Multiuser Receivers in Random Environments*

David Tse[†]
Dept. of EECS
University of California
Berkeley, CA 94720, USA
dtse@eecs.berkeley.edu

Ofer Zeitouni[‡]
Dept. of EE
Technion
Haifa 32000, Israel
zeitouni@ee.technion.ac.il

September 13, 1998

**Abstract**

We study the SIR performance of linear multiuser receivers in *random environments*, where signals from the users arrive in "random directions". Such random environment may arise in a DS-CDMA system with random signature sequences, or in a system with antenna diversity where the randomness is due to channel fading. Assuming that such random directions can be tracked by the receiver, the resulting SIR performance is a function of the directions and therefore also random. We study the asymptotic distribution of this random performance in the regime where both the number of users $M$ and the number of degrees of freedom $L$ in the system are large, but keeping their ratio fixed. Our results show that for both the decorrelator and the MMSE receiver, the variance of the SIR distribution decreases like $1/L$, and the SIR distribution is asymptotically Gaussian. We compute closed-form expressions for the asymptotic means and variances for both receivers. Simulation results are presented to verify the accuracy of the asymptotic results for finite-sized systems.

## 1 Introduction

In a direct-sequence code-division multiple access (DS-CDMA) system, each user modulates the information symbols onto its unique signature (or spreading) sequence. This

*Submitted to the IEEE Transactions on Information Theory.

[†]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.

[‡]The research of this author was done while visiting the department of EECS at Berkeley, under partial support from grant IRI 9310670.

spreading of information provide additional degrees of freedom for communication. To fully exploit the available degrees of freedom, linear *multiuser receivers* have been proposed to reduce or suppress the interference from other users. Prominent among these receivers are the decorrelator [8, 9] and the minimum mean square error (MMSE) receiver [26, 10, 14, 15].

A common performance measure for these linear receivers is the output signal-to-interference ratio (SIR). Clearly, the performance of these linear receivers depend on the signature sequences of the users. We focus on the common situation when the signature sequences of the users are *randomly* and *independently* chosen. This model is relevant in several scenarios: users may employ pseudo-random spreading sequences, or the transmitted signals from the users are distorted by independent multipath fading channels which randomize the received signature sequences. While the sequences are random, we assume in this paper that they are known perfectly at the receiver. In practice, this knowledge is obtained through adaptation, channel measurements or an initialization protocol. However, since the SIR performance of the users is a function of the signature sequences, it is also random. We are interested in characterizing their distributions.

The random sequence DS-CDMA model is an example of a system where multiuser receivers operate in a *random environment*. Another common example is a system with multiple antennas at the receiver. If the fading from a user to each of the receive antennas is independent, then diversity is achieved using multiple antennas to combat against the possibility of deep fade at any single antenna. Moreover, by tracking the fading of different users, linear multiuser receivers can be employed to suppress interference from other users while demodulating one particular user. The SIR performance is again random, being a function of the channel fading. This is very similar to the random signature sequence scenario since in both cases, signals from different users arrive from "random directions", where the directions are given by the signature sequences in the DS-CDMA system and by the fading patterns at the different antennas in the multi-antenna system. Indeed, one can think of the random sequence model considered in this paper as a canonical model for a multiuser system with diversity.

In this paper, we analyze the performance of both the decorrelator and the MMSE receiver in a random environment. The MMSE receiver is particularly interesting as it maximizes the SIR among all linear receivers. While it is known [10] that both receivers have the same near-far resistance (ability to reject worst-case interference), the MMSE receiver performs strictly better when the powers of the interferers are controlled or when they are relatively weak (such as from out of cell). The performance of the decorrelator, on the other hand, does not depend on the interferers' powers. The SIR performance of the MMSE receiver under power control will be studied.

In recent independent works [20, 24], it was shown that in a random environment, the SIR of a user under both the MMSE receiver and the decorrelator converges to a deterministic limit in a large system. The scaling is by letting the processing gain and the number of interferers go to infinity, keeping the number of interferers per unit processing gain fixed. In a finite system, however, the attained SIR will fluctuate around this limit. Such fluctuations determine important performance measures such as average probability of error and outage probability, i.e. the probability that the SIR of a user drops below a certain threshold. The main goal of this paper is to characterize such fluctuations in various scenarios. We provide Central Limit theorems which show that under appropriate scaling, the fluctuations are asymptotically Gaussian. Moreover, we give closed-form formulas for the variances of the fluctuations in terms of system parameters. Our results are obtained using techniques from random matrix theory. While the analysis of the deterministic limit involves only the asymptotic *eigenvalue* distribution of certain random matrices, the characterization of the SIR fluctuation requires understanding the asymptotic distribution of the *eigenvectors* as well as the fluctuation of the eigenvalue distribution around the asymptotic limit. Both of these are current research topics in random matrix theory and indeed our proofs exploit several recent results.

In related work, [6] has studied the problem of performance variability of linear multiuser detection under random signature sequences. They derived a heuristic approximation of the SIR performance of the decorrelator, and provided simulation results for the MMSE receiver. In contrast, our analytical results are justified by limit theorems and they apply both to the MMSE receiver and to the decorrelator. In the context of systems with antenna diversity, [25] have obtained related results on the performance of the decorrelator under flat Rayleigh fading. In this paper, our results apply to general fading distributions, not necessarily Rayleigh, which are of particular interest for distributed antenna systems, where the antennas can be placed at different locations of a room or a floor. In this scenario, the fading experienced consists of both small-scale (multipath) and large-scale effects, and cannot be accurately modeled as Rayleigh distributed. It turns out that relaxing the Rayleigh assumption complicates the analysis considerably.

Much of our results make only very weak assumptions on the distribution of the randomness and are therefore transparent to the specific random environment. For concreteness, we will focus on the DS-CDMA system with random signature sequences throughout most of the paper. In Section 2, we introduce the model. We analyse the performance of the decorrelator and the MMSE receiver in sections 3 and 4 respectively, with our main results being Theorems 3.3 and 4.5. Section 5 contains simulations validating the accuracy of our asymptotic results. In Section 6, we briefly comment on the application of our results to systems with antenna diversity. Section 7 contains our conclusions. The proofs of the results are found in the appendices.

During the final stage of the preparation of this paper, we were informed of independent work by Muller et al [13] on the performance of the decorrelator. The relationship between their results and ours will be discussed in Section 3. We were also informed of independent work by Kim and Honig [7] who have presented an approximation for the variance of the SIR under the MMSE. Unfortunately, no details were given about the derivation and we are unable to compare our approach here with theirs.

## 2 Linear Receivers for DS-CDMA Systems

In a DS-CDMA 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 sampled discrete-time model for a symbol-synchronous DS-CDMA system:

$$\mathbf{y} = \sum_{i=1}^{M} b_i \mathbf{s}_i + \mathbf{z}, \tag{1}$$

where $b_i \in \Re$ and $\mathbf{s}_i \in \Re^L$ are the transmitted symbol and signature sequence of user $m$ respectively, and $\mathbf{z}$ is $N(0, \sigma^2 I)$ background Gaussian noise. The length of the signature sequences is $L$, which is the number of degrees of freedom, and $M$ is the number of users. The received vector is $\mathbf{y} \in \Re^L$. We assume the $b_i$'s are independent and that $E[b_i] = 0$ and $E[b_i^2] = P_i$, where $P_i$ is the received power of user $i$ (energy per symbol).

We view multiuser receivers as *demodulators*, extracting good estimates of the (coded) symbols of each user as soft decisions to be used by the channel decoder [15]. From this point of view, the relevant performance measure is the signal-to-interference ratio (SIR) of the estimates. We shall now focus without loss of generality on the demodulation of user 1, assuming that the receiver has already acquired the knowledge of the spreading sequences. For user 1, the optimal demodulator $\mathbf{c}_1$ that generates a soft decision $\hat{b}_1 \equiv \mathbf{c}_1^t \mathbf{y}$ maximizing the signal-to-interference ratio (SIR) at the output :

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

is the MMSE receiver [10. 14, 15].

The formulae for the MMSE demodulator and its performance are well known [10]:

$$\hat{b}_{\text{mmse}}(\mathbf{y}) = \frac{P_1}{1 + P_1 \mathbf{s}_1^t (S_1 T S_1^t + \sigma^2 I)^{-1} \mathbf{s}_1} \mathbf{s}_1^t (S_1 T S_1^t + \sigma^2 I)^{-1} \mathbf{y} \tag{2}$$

and the signal to interference ratio $\beta$ for user 1 is

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

where $S_1 := [\mathbf{s}_2, \dots, \mathbf{s}_M]$ and $T := \text{diag}(P_2, \dots, P_M)$.

We observe that the MMSE receiver depends on the received powers of the interferers. The decorrelator is a simpler but sub-optimal linear receiver that operates without the need of knowing the received powers of the interferers. It simply nulls out the interference from other users by projecting the received signal onto the subspace orthogonal to the span of their signature sequences. The vector of symbol estimates $\hat{\mathbf{b}}_{dec}$ generated by the decorrelator for all users is given by:

$$\hat{\mathbf{b}}_{dec} := (S^t S)^{-1} S^t \mathbf{y},$$

where $S := [\mathbf{s}_1, \dots, \mathbf{s}_M]$. Here, the inverse is replaced by the pseudoinverse if $S^t S$ is not invertible. Observe that if there were no noise, the estimates will be exactly the original symbols, and hence it is the multiuser analog of the zero-forcing equalizer. Assuming that $S^t S$ is invertible, the SIR $\gamma$ of user 1 under the decorrelator is given by

$$\gamma = \frac{P_1}{\sigma^2[(S^t S)^{-1}]_{11}} \tag{4}$$

Note that the performance of the decorrelator does not depend on the powers of the interferers.

The formulas above for the SIR performance of various receivers can be numerically calculated given specific choices of the signature sequences. In this paper, however, we focus on the scenario when the sequences are *randomly* and independently chosen. In this case, the SIR performance of a receiver is a random variable, since it is a function of the spreading sequences, and we are interested in analysing its statistics. We will assume that though the sequences are randomly chosen, they are known to the receiver once they are picked. In practice, this means that the change in the spreading sequences is at a much slower time-scale than the symbol rate so that the receiver has the time to acquire the sequences. (There are known adaptive algorithms for which this can even be done blindly; see [5].) However, the *performance* of linear receivers depends on the initial choice of the sequences and hence is random.

The model for random sequences: let $\mathbf{s}_j = \frac{1}{\sqrt{L}}(v_{1j}, \dots, v_{Lj})^t$, j=1, ... M. The random variables $v_{ij}$'s are i.i.d., having zero mean, variance 1 and a symmetric distribution. The normalization by $\frac{1}{\sqrt{L}}$ ensures that $E[\|\mathbf{s}_j\|^2] = 1$, i.e. maintain a constant average power. In practice, it is quite common that the entries of the spreading sequences are 1 or $-1$, but our results hold for general distributions, which are useful when we look at other random environments such as systems with antenna diversity. We will also make the technical assumption that $E[v_{ij}^8] < \infty$. This last assumption can be relaxed (it is probably enough to assume the 4th moment is finite), but we chose this slightly stronger assumption in order to simplify the proofs.

# 3 Performance of the Decorrelator

We shall begin by studying the performance of the simpler decorrelator, before proceeding to the MMSE receiver. The following result shows that in a system with large processing gain and many users, the random SIR of a user converges to a deterministic limit. It is proved independently in [20, 21] and [24].

**Theorem 3.1** *Let $\gamma^{(L)}$ be the (random) SIR of the decorrelating receiver for user* 1 *when the spreading length is $L$ and the number of users $M = \lfloor \alpha L \rfloor$, where $\alpha > 0$ is a fixed constant. Then $\gamma^{(L)}$ converges to $\gamma^*$ in probability as $L \to \infty$, where $\gamma^*$ is given by*

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

In the scaling considered, the number of users per degree of freedom (or, equivalently, per unit bandwidth) $\alpha$ is fixed while the number of degrees of freedom grow. This scaling makes sense as more users can be supported by a larger bandwidth. Observe also that the above result holds regardless of the powers of the interferers, as the decorrelator nulls out all interferers and therefore its performance does not depend on the interferes' powers. Intuitively, this result says that for random signature sequences, the loss in SIR due to interference from other users is proportional to the number of interferers per degree of freedom.

Theorem 3.1 can be viewed as a law of large number. Though it gives the asymptotic limit, this result does not provide any information about the fluctuation around the limit for finite-sized system. This is the main consideration in this section. It is of interest to consider only the case when the number of users is less than the number of degrees of freedom, i.e. $\alpha < 1$, because otherwise the limiting SIR is zero. Moreover, since the performance of the decorrelator does not depend on the powers of the interferers, we can just focus on the case when the interferers have equal received power $P$, i.e. $T = PI$.

The first step is to obtain a formula for the SIR performance under the decorrelator, equivalent to but more useful for analysis than (4). It is known [10] that for the same signature sequences, the asymptotic efficiency of the decorrelator and MMSE receivers are identical, i.e.

$$\lim_{\sigma^2 \to 0} \gamma^{(L)} \sigma^2 = \lim_{\sigma^2 \to 0} \beta^{(L)} \sigma^2,$$

where $\beta^{(L)}$ is the SIR under the MMSE receiver in a system with processing gain $L$. Using eqn. (3) and (4), we therefore get

$$\frac{1}{[(S^t S)^{-1}]_{11}} = \lim_{\sigma^2 \to 0} \sigma^2 \mathbf{s}_1^t (S_1 T S_1^t + \sigma^2 I)^{-1} \mathbf{s}_1.$$

Let $S_1 T S_1^t = O F O^t$ be the spectral decomposition of $S_1 S_1^t$, where $F = \mathrm{diag}(\lambda_1, \dots, \lambda_L)$ is a diagonal matrix with decreasing eigenvalues and $O$ is an orthogonal matrix of the eigenvectors of $S_1 T S_1^t$. Putting this in the above expression and evaluating the limit, we get

$$\frac{1}{[(S^t S)^{-1}]_{11}} = \mathbf{s}_1^t O D O^t \mathbf{s}_1 \tag{5}$$

where $D = \mathrm{diag}(0, \dots, 0, 1, \dots, 1)$ and the number of 1's in the diagonal of $D$ is the number of zero eigenvalues of $S_1 S_1^t$.

To provide some background for our analysis of the random SIR performance for finite-sized systems, it is helpful to see first how Theorem 3.1 can be derived from the representation (5). The essence is based on the following lemma, proved in [12].

**Lemma 3.2** *Let* $\mathbf{s} = \frac{1}{\sqrt{L}}(v_1, \ldots, v_L)^t$ *where* $v_i$*'s are i.i.d. zero mean, unit variance random variable with finite 4th moment. Let* $A$ *be a deterministic* $L$ *by* $L$ *symmetric positive-definite matrix. Then*

$$E[\mathbf{s}^t A \mathbf{s}] = \frac{1}{L}\operatorname{tr} A$$

*and*

$$\operatorname{Var}\left[\mathbf{s}^t A \mathbf{s}\right] \leq \frac{1}{L} C_1 [\lambda_{\max}(A)]^2.$$

*for some constant* $C_1$ *which depends only on the fourth moment of* $v_1$.

This lemma holds for any deterministic matrix $A$. Applying this Lemma by conditioning on $A = ODO^t$ and observing that $A$ and $\mathbf{s}_1$ are independent, we obtain that

$$E[\mathbf{s}_1^t ODO^t \mathbf{s}_1] = \frac{1}{L} E[\operatorname{tr} D].$$

Also, $\lambda_{\max}(ODO^t) \leq 1$ and an application of Chebychev's inequality yields:

$$\mathbf{s}_1^t ODO^t \mathbf{s}_1 - \frac{1}{L}\operatorname{tr} D \xrightarrow{\mathcal{P}} 0 \tag{6}$$

Furthermore, Bai and Yin [1] showed that the smallest eigenvalue of the random matrix $S_1^t S_1$ converges almost surely to a positive number, when $\alpha < 1$. This implies that almost surely for large $L$, the signature sequences of the other users are linearly independent and the number of 1's in $D$ is $L - M + 1$. This together with (6) and (5) immediately yields Theorem 3.1.

Geometrically, the vector $\mathbf{w} = D^{\frac{1}{2}} O^t \mathbf{s}_1$ is the projection of the signature sequence of user 1 onto the subspace $V$ perpendicular to the signature sequences of other users. The decorrelator demodulates user 1 by projecting the received signal onto $\mathbf{w}$, and the SIR is determined by the length of $\mathbf{w}$. The above result says that in a large system, the amount of energy of $\mathbf{s}_1$ in $V$ is approximately proportional to the dimension of that space. This is what one would expect from the i.i.d. nature of the components of $\mathbf{s}_1$.

Observe that the above derivation of the asymptotic limit makes use of the convergence of $\operatorname{tr} D$ (i.e. the dimension of the subspace $V$) but not any properties of $O$, the eigenvectors of $S_1 S_1^t$. In fact, it depends only on the randomness of $\mathbf{s}_1$. However, when we are interested in characterizing the *fluctuations* of the SIR around the asymptotic limit, asymptotic properties of the eigenvectors are needed. The mathematical apparatus to deal with this is established in Appendix A. The solution depends on $\mu_D(\cdot)$, the asymptotic empirical distribution of the eigenvalues of $ODO^t$; this is given by $\mu_D(x) = \alpha\delta(x) + (1-\alpha)\delta(x-1)$. Applying Corollary A.2 to this problem, we can then conclude that

$$\sqrt{L}\left[\mathbf{s}_1^t ODO^t \mathbf{s}_1 - \frac{1}{L}\operatorname{tr} D\right] \xrightarrow{\mathcal{P}} N(0, a)$$

where

$$\begin{aligned} a &= 2\int x^2 \mu_D(x) + (E(v_{11}^4) - 3)\left(\int x\mu_D(x)\right)^2 \\ &= 2(1-\alpha) + (E(v_{11}^4) - 3)(1-\alpha)^2 \end{aligned}$$

This together with the fact that $\operatorname{tr} D$ converges almost surely to $L - M + 1$ yields the following theorem:

**Theorem 3.3** *For $\alpha < 1$, as $L \to \infty$,*

$$\sqrt{L}\left(\gamma^{(L)} - \frac{P_1}{\sigma^2}(1-\alpha)\right) \xrightarrow{\mathcal{D}} \mathcal{N}(0, (\frac{P_1}{\sigma^2})^2 a)$$

*where*

$$a = 2(1-\alpha) + (E[v_{11}^4] - 3)(1-\alpha)^2$$

This theorem says that the fluctuation of the SIR around the limit is approximately Gaussian with variance $\frac{1}{L}(\frac{P_1}{\sigma^2})^2 a$, decreasing like $1/L$ and with $a$ depending only on $\alpha$ and the fourth moment of $v_{11}$. Observe also the variance increases with $E[v_{11}^4]$, and hence is minimized when the entries take on $+1$ or $-1$ values only. It should be noted that while the asymptotic limit depends only on the second moment of $v_{ij}$'s, the amount of fluctuation around the limit depends on the fourth moment. and thus varies from one distribution to another.

Since the truth of Theorem 3.3 depends entirely on the machinery developed in Appendix A and the proofs there are rather technical, we would like to give some intuition as to why it holds. Define

$$\mathbf{u} = \frac{1}{\sqrt{L}}(u_1, \ldots u_L)^t := O^t \mathbf{s}_1.$$

Assuming that the signature sequences of the interfering users are linearly independent (which holds with probability 1 in a large system), we have

$$\gamma^{(L)} = \frac{P_1}{\sigma^2}\frac{1}{L}\sum_{i=M}^{L} u_i^2$$

First consider the special case when the entries $v_{1j}$ of the spreading sequences of user 1 are Gaussian. Then the $u_i$'s are i.i.d. Gaussian $N(0,1)$. In this case, $\gamma^{(L)}$ is Chi-square distributed. This is basically the main result of [25], except that they considered complex Gaussian $v_{ij}$ for their Rayleigh fading model. For large $L$, a direct application of the Central Limit Theorem yields Theorem 3.3, with $E[v_{11}^4] = 3$.

We observe that in the special case of Gaussian $v_{ij}$, the Central Limit approximation is actually not necessary as the explicit distribution of $\gamma^{(L)}$ can be obtained for finite $L$. Moreover, the properties of the eigenvectors $O$ play no role here, other than the fact that $O$ is independent of $\mathbf{s}_1$. The key reason is that the i.i.d. Gaussian distribution is *isotropic*, i.e. invariant to orthogonal transformations, so that whatever $O$ is , $O^t\mathbf{s}_1$ is also isotropic.

Let us now consider the general case where the $v_{ij}$'s are not necessarily Gaussian so that $\mathbf{s}_1$ may not be isotropic. In this case, $O^t\mathbf{s}_1$ has a complicated distribution dependent on both the distribution of $O$ and $\mathbf{s}_1$, and need not be isotropic. To analyse this problem, we need to exploit a special property of the eigenvectors of $S_1 S_1^t$. In particular, we show that even though $\mathbf{s}_1$ may not be isotropic, as $L \to \infty$, the vector $\mathbf{u} := O^t \mathbf{s}_1$ will be asymptotically isotropic and moreover independent of $\|\mathbf{s}_1\|$. (This fact is made precise in Theorem A.1 of Appendix A.) In essence, we show there that there is enough randomness in $O$ to make $O^t \mathbf{s}_1$ close to being isotropic.

This implies that we can write $\mathbf{u}$ as

$$\frac{\mathbf{u}}{\|\mathbf{s}_1\|} \approx \frac{1}{\|\mathbf{r}\|}(r_1, \ldots, r_L)^t$$

where the $r_i$'s are i.i.d. Gaussian $N(0,1)$ and independent of $\|\mathbf{s}_1\|$. Thus,

$$O^t \mathbf{s}_1 \approx \frac{\|\mathbf{s}_1\|}{\|\mathbf{r}\|}(r_1, \ldots, r_L)^t$$

and

$$\gamma^{(L)} \approx \frac{P_1}{\sigma^2} \frac{\|\mathbf{s}_1\|^2}{\|\mathbf{r}\|^2} \sum_{i=M}^{L} r_i^2. \tag{7}$$

Using Central Limit Theorem, it can be seen that

$$L\|\mathbf{s}_1\|^2 \approx 1 + \frac{1}{\sqrt{L}}\phi_1 \tag{8}$$

$$\|\mathbf{r}\|^2 \approx 1 + \frac{1}{\sqrt{L}}\phi_2 \tag{9}$$

$$\frac{1}{L}\sum_{i=M}^{L} r_i^2 \approx (1-\alpha) + \frac{1}{\sqrt{L}}\phi_3 \tag{10}$$

where the $\phi_i$'s are zero-mean jointly Gaussian and $\phi_1$ independent of $\phi_2$ and $\phi_3$. The second moments of these random variables can be calculated as:

$$E[\phi_1^2] = E[v_{11}^4] - 1; \qquad E[\phi_2^2] = 2; \qquad E[\phi_3^2] = 2(1-\alpha)^2$$

and

$$E[\phi_2\phi_3] = 2(1-\alpha).$$

Using (8), (9), (10), we can perform a Taylor-series expansion of (7), keeping the first and second order terms only, and obtain

$$\gamma^{(L)} \approx \frac{P_1}{\sigma^2}\left[1 - \alpha + \frac{1}{\sqrt{L}}((1-\alpha)\phi_1 - (1-\alpha)\phi_2 + \phi_3)\right].$$

Direct computation reveals that the variance of the Gaussian fluctuation $(1-\alpha)(\phi_1 - \phi_2) + \phi_3$ is precisely $a$ given in Theorem 3.3.

The essence of the above argument is based on the fact that the eigenvector matrix $O$ of $S_1 S_1^t$ itself is in some sense asymptotically *isotropic*. A version of this phenomenon has been proved by Silverstein [17]: he showed that given any deterministic vector $\mathbf{s}_1$ whose entries are either $+1/\sqrt{L}$ or $-1/\sqrt{L}$, the random vector $O^t\mathbf{s}_1$ is asymptotically isotropic, to the accuracy of the Central-Limit approximation. We show that this is true also when $\mathbf{s}_1$ is a random vector with i.i.d. elements of general distribution, but independent of $O$. This fact is made precise in Theorem A.1 in Appendix A, using the theory of weak convergence.

An interesting observation from the above heuristic derivation is that the asymptotic distribution of the SIR under the decorrelator depends on the distribution of $v_{ij}'s$ only through that of $\|\mathbf{s}_1\|^2$, i.e. the fluctuation of the received energy of the signal from user 1. In the special case when the signature sequence entries takes on $+1/\sqrt{L}$ or $-1/\sqrt{L}$, $\|\mathbf{s}_1\|^2 = 1$ and eqn. (7) simplifies to

$$\gamma^{(L)} \approx \frac{P_1}{\sigma^2}\frac{1}{\|\mathbf{r}\|^2}\sum_{i=M}^{L} r_i^2.$$

A similar approximation was proposed independently in [13]. However, the assumption of $O^t\mathbf{s}_1$ being asymptotically isotropic was made without justification. As was pointed out, this matter is rather subtle as the property depends both on the distributions of $O$ and $\mathbf{s}_1$.

# 4 Performance of MMSE Receiver

We now turn to analysing the performance of the MMSE receiver. In [21], it is shown that in a large system, the SIR under the MMSE receiver converges to a deterministic limit. While the results there apply to the general setting of users with unequal received powers, we focus here on the case when the users are controlled to equal received power. This would be the case when users are all in a single cell and have the same SIR requirements. In this case, the limiting SIR has a simple closed-form expression, which is also obtained independently in [24].

**Theorem 4.1** *[21] Let $\beta^{(L)}$ be the (random) SIR of the MMSE receiver for user* 1 *when the spreading length is $L$. Suppose the received powers of the users are all equal to $P$. Then $\beta^{(L)}$ converges to $\beta^*$ in probability as $L \to \infty$, where $\beta^*$ is given by:*

$$\beta^* = \frac{(1-\alpha)P}{2\sigma^2} - \frac{1}{2} + \sqrt{\frac{(1-\alpha)^2P^2}{4\sigma^4} + \frac{(1+\alpha)P}{2\sigma^2} + \frac{1}{4}} \qquad (11)$$

To provide some background in understanding our approach to analysing the random performance in a finite-sized system, it helps to first give the basic intuition behind the proof of Theorem 4.1. Recall from eqn. (3) that the SIR of user 1 under the MMSE receiver is given by:

$$\beta = P\mathbf{s}_1^t(S_1TS_1^t + \sigma^2 I)^{-1}\mathbf{s}_1$$

where $S_1 = [\mathbf{s}_2, \dots, \mathbf{s}_M]$ and $T = PI$. In terms of the spectral decomposition $STS^t = OFO^t$ introduced in the previous section, where $F = \mathrm{diag}(\lambda_1, \dots, \lambda_L)$, we have

$$\beta = P\mathbf{s}_1^t O(F + \sigma^2 I)^{-1}O^t\mathbf{s}_1.$$

Comparing this to the performance of the decorrelator:

$$\gamma = \frac{P}{\sigma^2}\mathbf{s}_1^t ODO^t\mathbf{s}_1,$$

where $D = \mathrm{diag}(0, \dots, 0, 1, \dots, 1)$, we see that the expression for the MMSE receiver is more complicated as it depends on the random eigenvalues of $S_1TS_1^t$ as well. This reflects the fact that the MMSE receiver attains a better performance by taking into account the strength of the interferers rather than just nulling them out.

Nevertheless. Theorem 4.1 can be proved by, first, using Lemma 3.2 to show that for large L

$$\beta^{(L)} \approx \frac{P}{L}\mathrm{tr}\,(F + \sigma^2 I)^{-1}.$$

Second, using results from random matrix theory [12, 16], it can further be deduced that the empirical distribution of the eigenvalues of $S_1TS_1^t$ converges to some limiting distribution $G^*$. Combining these facts. we obtain that $\beta^{(L)}$ converges in probability to

$$P\int \frac{1}{\lambda + \sigma^2}dG^*(\lambda).$$

In [21], it is further shown how this limit can be explicitly computed to be (11). This calculation is also done in Appendix C.

Following this train of thought, the random fluctuation of $\beta^{(L)}$ around the limit $\beta^*$ can be dealt with by decomposing into three terms:

$$\beta^{(L)} - \frac{P}{L}\mathrm{tr}\,(F + \sigma^2 I)^{-1} \tag{12}$$

and

$$\frac{P}{L}\mathrm{tr}\,(F + \sigma^2 I)^{-1} - E[\beta^{(L)}], \tag{13}$$

and

$$E[\beta^{(L)}] - \beta^* \tag{14}$$

Note that the first term depends on $\mathbf{s}_1$ and the eigenvector matrix $O$, while the second and third terms depend only on the fluctuation of the empirical eigenvalue distribution of $S_1TS_1$ around the limiting distribution $G^*$. Just as for the decorrelator, the first fluctuation can be characterized using the theory developed in Appendix A. Applying Corollary A.2 there, we obtain:

**Lemma 4.2**
$$\sqrt{L}\left(\beta^{(L)} - \frac{P}{L}\mathrm{tr}\,(F+\sigma^2 I)^{-1}\right) \xrightarrow{\mathcal{D}} N(0,b)$$

*where*

$$b = 2\int\left[\frac{P}{(\lambda+\sigma^2)}\right]^2 dG^*(\lambda) + (E(v_{11}^4)-3)\left[\int\frac{P}{\lambda+\sigma^2}dG^*(\lambda)\right]^2 \qquad (15)$$

Note that
$$\int\frac{P}{\lambda+\sigma^2}dG^*(\lambda) = \beta^*$$

and
$$\int\frac{P}{(\lambda+\sigma^2)^2}dG^*(\lambda) = -\frac{d\beta^*}{d(\sigma^2)}.$$

Thus, to compute the second integral, we need only to differentiate eqn. (11) with respect to $\sigma^2$. We therefore get

$$b = \frac{2\beta^*(1+\beta^*)^2}{\frac{\sigma^2}{P}(1+\beta^*)^2+\alpha} + \left(E[v_{11}^4]-3\right)(\beta^*)^2.$$

The above lemma says that the fluctuation of the first term (12) is of the order of $1/\sqrt{L}$. Regarding the fluctuation of $\frac{P}{L}\mathrm{tr}\,(F+\sigma^2 I)^{-1}$, we have the following result, the proof of which can be found in Appendix B.

**Lemma 4.3**
$$\limsup_{L\to\infty}\mathrm{Var}\left[\sum_{i=1}^{L}\frac{1}{\lambda_i+\sigma^2}\right] < \infty$$

This says that the fluctuation of $\frac{P}{L}\mathrm{tr}\,(F+\sigma^2 I)^{-1}$ around its mean is of the order at most $1/L$, negligible compared to the first source of fluctuation (12).

Finally, concerning the deviation of the mean SIR from the limit $\beta^*$, we have the following result, proved in Appendix C.

**Lemma 4.4**
$$\limsup_{L\to\infty} L\left(E[\beta^{(L)}]-\beta^*\right) < \infty$$

This shows that the mean SIR is of order at most $1/L$ from the limit $\beta^*$. Combining lemmas 4.2, 4.3 and 4.4, we have the following main result characterizing the asymptotic distribution of the performance under the MMSE receiver.

**Theorem 4.5** *As $L\to\infty$,*

$$\sqrt{L}(\beta^{(L)}-\beta^*) \xrightarrow{\mathcal{D}} N(0,b)$$

*where*

$$b = \frac{2\beta^*(1+\beta^*)^2}{\frac{\sigma^2}{P}(1+\beta^*)^2 + \alpha} + \left(E[v_{11}^4] - 3\right)(\beta^*)^2$$

*and*

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

*Moreover,*

$$\limsup_{L\to\infty} L\left(E[\beta^{(L)}] - \beta^*\right) < \infty$$

This theorem says that while the asymptotic limit $\beta^*$ can be expected to be a very accurate approximation of the mean SIR for reasonably sized system (difference of order $1/L$) , the fluctuations can be significantly larger (of order $1/\sqrt{L}$). This will be validated by the simulation results in the next section.

We would like to give some intuition underlying the proof of this result. This is similar to our heuristic discussion on the deccorelator. Because of the asymptotically isotropic distribution of $O^t\mathbf{s}_1$, we can write

$$O^t\mathbf{s}_1 \approx \frac{\|\mathbf{s}_1\|}{\|\mathbf{r}\|}(r_1, \ldots, r_L)^t$$

where the $r_i$'s are i.i.d. Gaussian $N(0,1)$ and independent of $\|\mathbf{s}_1\|$. Thus,

$$\beta^{(L)} \approx \frac{\|\mathbf{s}_1\|^2}{\|\mathbf{r}\|^2}\sum_{i=1}^{L}\frac{P}{\lambda_i+\sigma^2}r_i^2. \tag{16}$$

Using a process-level Central Limit argument together with the fact that the random eigenvalue fluctuation is small, it can be shown that

$$L\|\mathbf{s}_1\|^2 \approx 1 + \frac{1}{\sqrt{L}}\phi_1 \tag{17}$$

$$\|\mathbf{r}\|^2 \approx 1 + \frac{1}{\sqrt{L}}\phi_2 \tag{18}$$

$$\frac{1}{L}\sum_{i=1}^{L}\frac{P}{\lambda_i+\sigma^2}r_i^2 \approx \beta^* + \frac{1}{\sqrt{L}}\phi_3 \tag{19}$$

where the $\phi_i$'s are zero-mean jointly Gaussian and $\phi_1$ independent of $\phi_2$ and $\phi_3$. The second moments of these random variables can be calculated as:

$$E[\phi_1^2] = E[v_{11}^4] - 1; \qquad E[\phi_2^2] = 2; \qquad E[\phi_3^2] = 2\int_0^\infty \left[\frac{P}{(\lambda+\sigma^2)}\right]^2 dG^*(\lambda)$$

and

$$E[\phi_2\phi_3] = 2\beta^*.$$

Using (17), (18), (19), we can expand (16) in the Taylor-series expansion, keeping the first and second order terms only, and obtain the Gaussian approximation given in the main theorem.

# 5 Simulations and Numerical Results

To see how accurate the limit theorems are for finite-sized system, we compare the theoretical results with actual values obtained by simulations. All simulation results are obtained by averaging over 10,000 independently generated samples, and will be considered as the actual values of the statistics. Users are received at equal power $P$, and the SNR $P/\sigma^2$ is set at 20dB. The chips of the signature sequences have values $+1/\sqrt{L}$ or $-1/\sqrt{L}$. Fig. 1 and 2 display results for the MMSE receiver. In Fig. 1, we plot the limiting SIR $\beta^*$ (given by formula (11)), the mean SIR $\bar{\beta}_1^{(L)}$, and the actual and theoretical SIR level at one standard deviation below the mean. These curves are plotted as a function of $\alpha$ (number of users per degree of freedom) for different system sizes: $L = 16, 32, 64, 128$. The theoretical SIR level 1 standard deviation from the mean is given by

$$\beta^* - \sqrt{\frac{b}{L}}$$

where $b$ is given by eqn. (15).

We make several observations. First, the mean $\bar{\beta}^{(L)}$ is very close $\beta^*$, and their difference is much smaller than 1 standard deviation. For $L = 32$ and greater, the $\beta^*$ and $\bar{\beta}^{(L)}$ curves are almost indistinguishable. This confirms our theoretical results, which predict that $\beta^* - \bar{\beta}^{(L)}$ is going to zero at least as fast as $1/L$, while the standard deviation goes to zero like $1/\sqrt{L}$. Second, the theoretical prediction $\sqrt{b/L}$ of the standard deviation is quite close to the actual standard deviation. Again, the two corresponding curves are almost indistinguishable for $L \geq 32$. Third, the standard deviation compared to the mean SIR is small where there are few users per unit processing gain, but quite significant when there are many users. This is true even for $L = 64$.

Next, we investigate how accurate the Central-Limit results are in predicting the tail of the SIR distribution. In Fig. 2, we compare the actual 1%-outage SIR with that predicted by Theorem 4.5. (The 1% outage level is the value $x$ such that $\Pr(SIR < x) = 0.01$. ) We see that while the theoretical result is accurate when $\alpha$ is small (less than 0.5), it tends to be over-pessimistic for $\alpha$ larger, when the achieved SIR is small. The accuracy of the theoretical results becomes good for the entire range only when $L = 128$.

For the decorrelator, as we mentioned in Section 3 and was also independently pointed out in [13], an alternative approximation is suggested by the heuristic (7). This does not assume a Gaussian approximation to the various sums, but is only based on the fact that $O_1^s$ is asymptotically isotropic. The random variable

$$\frac{\sum_{i=M}^{L} r_i^2}{\sum_{i=1}^{L} r_i^2}$$

follows a Beta distribution, since the $r_i$'s are i.i.d. $N(0,1)$. (See for example [3].) Hence an approximation to $\gamma^{(L)}$ is a product of the independent random variable $\|\mathbf{s}_1\|^2$ and a Beta distributed random variable. In the special case of $+1, -1$ sequences, $\|\mathbf{s}_1\|^2 = 1$ and this approximation simply becomes a Beta distribution. The approximation is applied

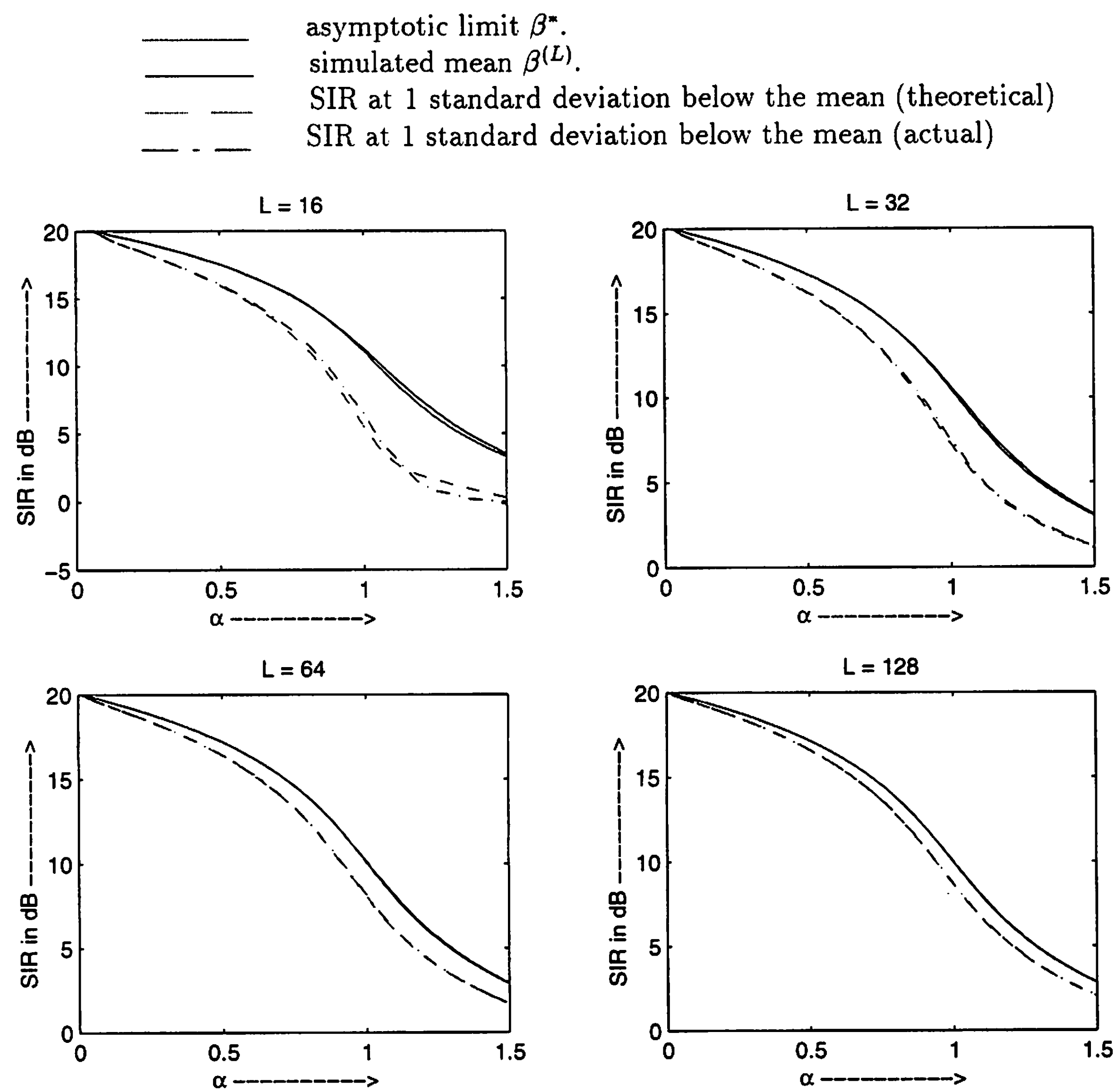


Figure 1: Asymptotic limit, mean, and SIR at one standard deviation below the mean, theory and actual, for the MMSE receiver.

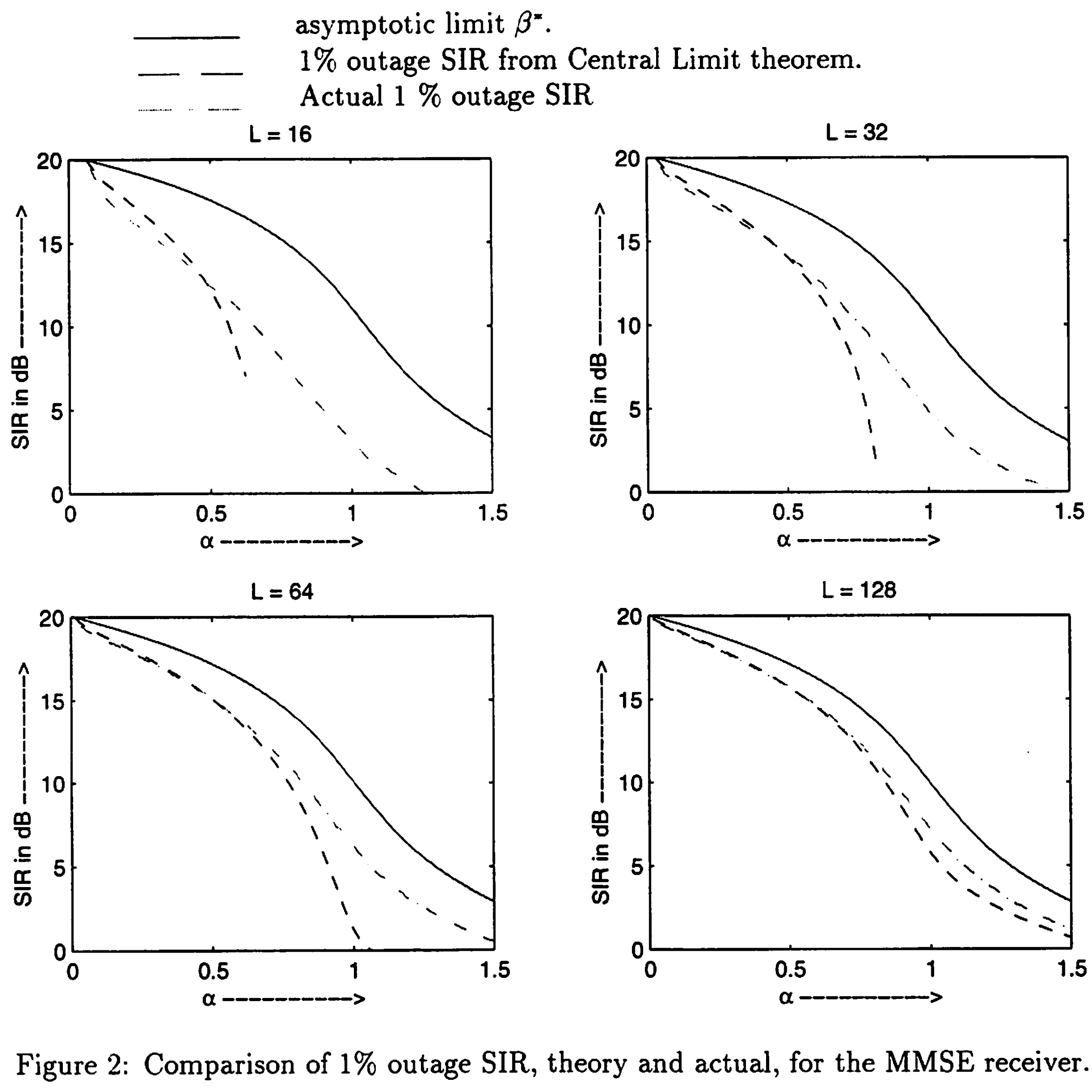


Figure 2: Comparison of 1% outage SIR, theory and actual, for the MMSE receiver.

to calculate the 1%-outage level for the deccorelator in Fig. 3. The result is compared to the actual 1%-outage level, as well as the Central-Limit approximation provided by Theorem 3.3. We see that even for $L = 16$, the Beta distribution approximation is very accurate, and in fact indistinguishable from the actual values for $L \geq 32$. On the other hand, the Central Limit approximation, while accurate for small $\alpha$, tends to be over-pessimistic for $\alpha$ close to 1. This suggests that for moderate $L$, $O^t \mathbf{s}_1$ is already very close to perfectly isotropic. On the other hand, the Gaussian approximation to $\sum_{i=1}^{L} r_i^2$ and $\sum_{i=M}^{L} r_i^2$ introduces errors which are only negligible when $L$ is quite large. Thus, when $L = 128$, all three curves (actual, Beta distribution approximation, Central Limit Theorem approximation) merge.

# 6 Antenna Diversity

In the previous sections, we have focused on the DS-CDMA system with random signature sequences. Another example of a random environment in which linear multiuser receivers operate is a system with multiple antennas for providing spatial diversity. These antennas can be arranged in an array located at a single basestation, or they can be distributed in geographically different locations in which case they provide *macro-diversity*. Antenna elements co-located in an array mainly serves to combat *multipath fading*, while a distributed antenna system can combat larger scale fading effects. In any case, performance can be improved by adaptive combination of signals received at the various antenna elements depending on the channel strengths. A general baseband model for such a system with flat fading is given by:

$$\mathbf{y} = \sum_{i=1}^{M} b_i \mathbf{h}_i + \mathbf{z},$$

where $b_i$ is the transmitted symbol of the $i$th user, and $\mathbf{y}$ is a $L$-dimensional vector of received symbols at the antennas. The vector $\mathbf{z}$ is i.i.d. complex circular symmetric Gaussian noise with variance per component $\sigma^2$. The vector $\mathbf{h}_i$ represents the (flat) fading of the $i$th user at each of the antennas. Let $\mathbf{h}_j = (v_{1j}/\sqrt{L}, \dots v_{Lj}/\sqrt{L})^t$. We will assume a fading model in which $v_{ij}$ are i.i.d. circular symmetric random variables with variance $E[|v_{ij}|^2]$ normalized to be 1, to keep the total received energy at the antennas constant, irrespective of the number of antennas. The circular symmetry arises naturally when shifting from a high carrier frequency to the baseband. We will also assume the signal constellation is circular symmetric as well, so that $b_i$ is circular symmetric. The average received power of all users are assumed to be the same, $E[|b_i|^2] = P$. We let $\alpha = \lfloor M/L \rfloor$ be the number of users per antenna element.

Assuming that the receiver can track the fading perfectly, the MMSE receiver is the optimal linear receiver in maximizing the SIR of each user 1. The decorrelator nulls out the interference from other users. The performance of both of these receivers is a function of the channel fading at the current time, and is therefore random.

The similarity of the multi-antenna model with the DS-CDMA system is obvious, with

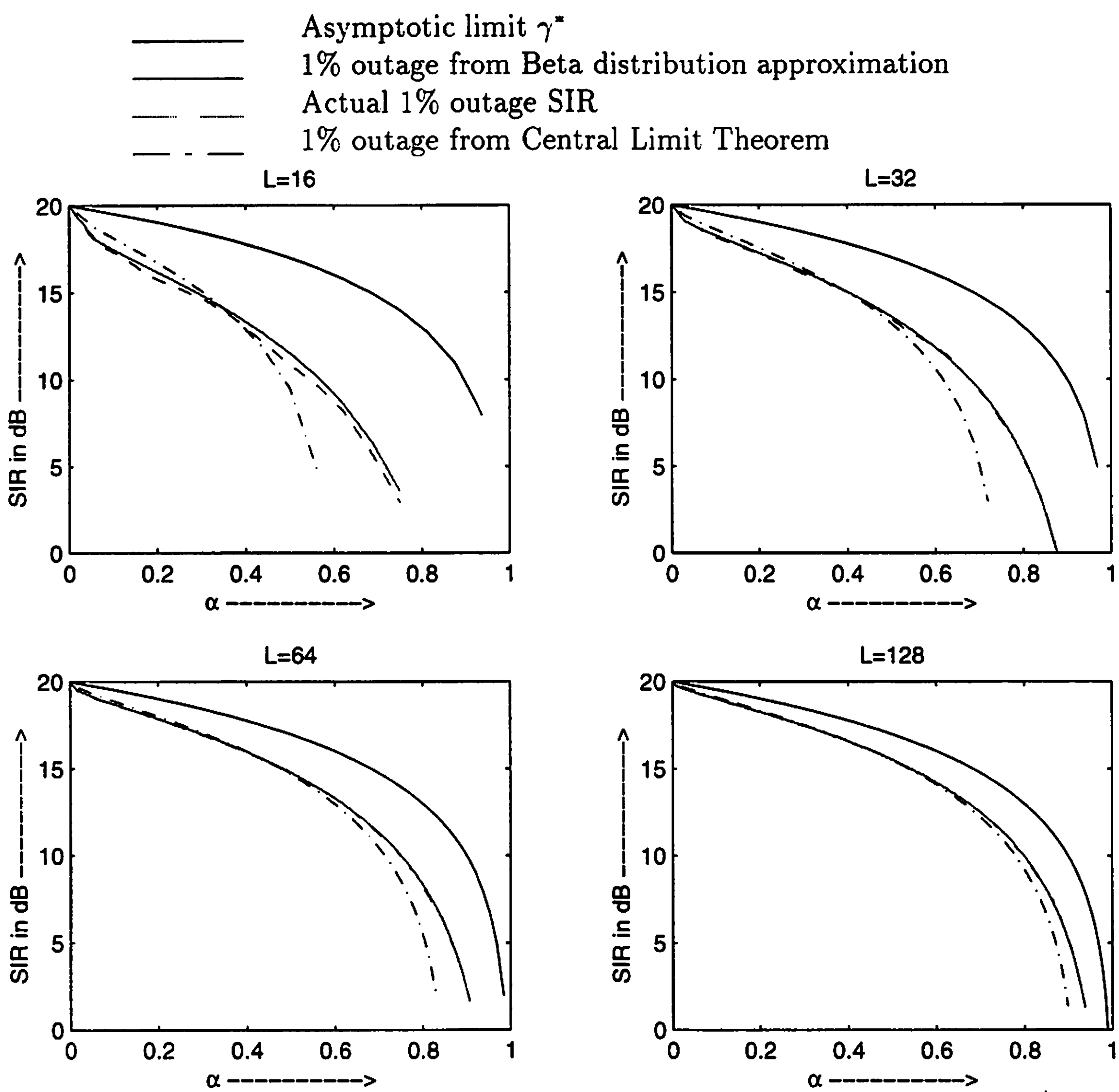


Figure 3: Comparison of 1% outage SIR for the decorrelator, Central Limit approximation, Beta distribution approximation and actual.

the signature sequences replaced by the channel fading vectors. The only difference is that the entries of $H$ are now complex rather than real as in the signature sequences. Rigorously speaking, Theorem A.1 which we used for analysing the DS-CDMA problem, although undoubtedly true for general symmetric distribution of $v_{ij}$'s, is only proved for *real* $v_{ij}$. (The proof of Lemmas 4.3 and 4.4 carries over verbatim to the complex case, c.f. for a similar argument [2]). However, we expect that Theorem A.1 can be generalized to the case when $v_{ij}$ is complex circular symmetric. Based on this assumption, the performance of the decorrelator in the multi-antenna system can be approximated by (in analogy to (7)):

$$\gamma^{(L)} \approx \frac{P}{\sigma^2}\frac{\|\mathbf{h}_1\|^2}{\|\mathbf{r}\|^2}\sum_{i=M}^{L}|r_i|^2$$

where $r_i$'s are i.i.d. zero-mean complex circular symmetric Gaussian random variables with $E[|r_i|^2] = 1$. This assumes $\alpha < 1$. Note that in the case of the Rayleigh fading model, $v_{ij}$'s are circular symmetric Gaussian and the approximation becomes exact, and this specializes to the result of [25]. For large $L$, applying the Central Limit Theorem, $\gamma^{(L)}$ can be further approximated by a Gaussian random variable with mean $(1-\alpha)P/\sigma^2$ and variance

$$\frac{1}{L}\left(\frac{P}{\sigma^2}\right)^2\left\{2(1-\alpha)+(E[|v_{11}|^4]-3)(1-\alpha)^2\right\}.$$

For the MMSE receiver, the SIR performance can be approximated, for large $L$, by a Gaussian random variable with mean $\beta^*$ and variance

$$\frac{1}{L}\left\{\frac{2\beta^*(1+\beta^*)^2}{\frac{\sigma^2}{P}(1+\beta^*)^2+\alpha}+\left(E[|v_{11}|^4]-3\right)(\beta^*)^2\right\},$$

where $\beta^*$ is given by eqn. (11).

# 7 Conclusions

In this paper, we studied the SIR performance of the decorrelator and the MMSE receiver in a random environment. Such random environment may arise in a DS-CDMA system with random signature sequences, or in a system with antenna diversity where the randomness is due to channel fading. We showed that for the two receivers considered, the variance of the SIR distribution decreases like $1/L$, and the SIR distribution is asymptotically Gaussian. We computed closed-form expressions for the variances for both receivers, and observed that the relative amount of fluctuation is large when there are many users per degree of freedom and the achieved SIR is low.

Simulation results show that the asymptotic mean and variance computed from the theory are very accurate approximations for even moderate system size and for a wide range of $\alpha$ (number of users per degree of freedom). On the other hand, when the achieved SIR is small and system size only moderate, the Gaussian approximation is not very good for approximating the tail of the SIR distribution (1% outage, for example.) Based on insights gained from the theory, an alternative approximation based on the Beta distribution is derived for the performance of the decorrelator. This approximation, observed independently in [13], is very accurate for moderate system size and for the whole range of $\alpha$.

There are several interesting directions for future work. One remedy to offset the random fluctuation of the SIR is through power control. The interesting question is then to characterize the distribution of power required to keep the SIR at a desired level. The problem is complicated by the fact that all users will vary their powers simultaneously to achieve their individual desired SIR. However, we conjecture that in the scaling considered in this paper, the performance of a user is insensitive to the power variations of other users and depends mainly on its own power. This would then imply that the power distribution can be computed as that of the reciprocal of the SIR calculated in this paper.

Another interesting question is to characterize the empirical distribution of SIR levels of the users across the system. Contrast this with the SIR distribution of a *particular* user, which is what we computed in this paper. We conjecture that in a large system, some kind of "weak asymptotic independence" between users will hold and with high probability the two distributions are very close.

# Acknowledgments

The help from Kiran on the simulations is very much appreciated. The first author would like to thank Prof. Jack Salz for discussions in the initial phase of this work. We would also like to thank Prof. Jack Silverstein for pointing out to us that the proof of Lemma 4.3 is contained in [2], and Prof. S. Shamai for the reference to [13].

# 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] Bai, Z.D. and J. W. Silverstein, " No eigenvalues outside the support of the limiting spectral distribution of large dimensional sample covariance matrices", *Annals of Probability* 26(1) (1998), pp. 316-345.

[3] Bickel, P.J. and K.A. Doksum, *Mathematical Statistics: Basic Ideas and Selected Topics*, Golden Day, 1977.

[4] Burkholder, D.L., "Distribution function inequalities for martingales", *Annals Probab.*, 1 (1973), pp. 19-42.

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

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

[7] Kim, J.B. and M. Honig, "Outage Probability of Multi-code DS-CDMA with Linear Interference Suppression", to appear in MILCOM.

[8] 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.

[9] 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.

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

[11] 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.*

[12] Marcenko, V.A. and L.A Pastur, (1967) "Distribution of eigenvalues for some sets of random matrices" *Math. USSR-Sb* 1. (1967), pp. 457-483.

[13] Müller, R. R., P. Schramm and J. B. Huber, "Spectral efficiency of CDMA systems with linear interference suppression", IEEE Workshop on Communication Engineering, Ulm, Germany, January 1997, pp. 93–97. See also "Spectral efficiency of multiuser systems based on CDMA with linear MMSE interference suppression", ISIT 97, Ulm, Germany, pp. 266.

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

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

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

[17] Silverstein, J.W., "Weak convergence of random functions defined by the eigenvectors of sample covariance matrices", *Annals of Probability*, v.18, no. 3, 1990, pp. 1174-1194.

[18] Silverstein, J.W., "On the eigenvectors of large dimensional sample covariance matrices", *J. Multivariate Analysis*, v.30, 1989, pp. 1-16.

[19] Sinai, Ya and Soshnikov, A., "Central limit theorem for traces of large random symmetric matrices with independent matrix elements", preprint, 1997.

[20] Tse. D. and S.V. Hanly, "Multiuser demodulation: effective interference, effective bandwidth and capacity", presented at Allerton Conference, Oct. 1997.

[21] Tse, D. and S.V. Hanly, "Linear multiuser receivers: effective interference, effective bandwidth and user capacity", to appear in *IEEE Trans. on Information Theory.*

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

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

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

[25] Winters J.H., J. Salz. and R. Gitlin, "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.

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

# Appendices

## A Asymptotically Isotropic Eigenvectors

In this section, we develop the machinery required to prove Theorem 3.3 and Lemma 4.2. Theorem A.1 quantifies precisely what it means to say that the eigenvectors of a random matrix are *asymptotically isotopic.* Its proof uses heavily ideas from Silverstein [17] and so we adopt his notations.

**Notations** We let $\mathbf{x}_n = (x_1, \cdots, x_n)$ denote random vectors with $\|\mathbf{x}_n\| = \sqrt{\sum_{i=1}^n x_i^2} = 1$, and let $\mathbf{z}_n = (z_1, \cdots, z_n)$ denote arbitrary random vectors in $\Re^n$. As in [17], we let $\{v_{ij}\}_{\substack{i=1,\cdots \\ j=1,\cdots}}$ be i.i.d. random variables with $Ev_{ij}^2 = 1$, $Ev_{ij} = 0$, $Ev_{ij}^4 < \infty$ and symmetric distribution. With $V_n = \{v_{ij}\}_{\substack{i=1 \\ j=1,\cdots s(n)}}^n$, let $M_n = \frac{1}{s(n)} V_n V_n^t$, and define $\frac{n}{s(n)} \underset{n \to \infty}{\longrightarrow} y \in (0, \infty)$.

Let $M_n = O_n \Lambda_n O_n^t$ denote the spectral resolution of $M_n$, with $\Lambda_n$ a diagonal matrix whose entries, the eigenvalues of $M_n$, are arranged in nondecreasing order, and $O_n$ denoting an orthonormal matrix consisting of the eigenvectors of $M_n$.

Let $z_1, z_2, \cdots$ denote a sequence of i.i.d., independent of $\{v_i\}$ random variables, with $Ez_i = 0$, $Ez_i^2 = 1$, $Ez_i^4 < \infty$, with symmetric distribution. Let

$$\eta_n = \frac{1}{\sqrt{n}} \sum_{i=1}^n (z_i^2 - 1) \xrightarrow{\mathcal{D}} \eta$$

where $\eta$ denotes a Gaussian random variable with zero mean and variance $\sigma_\eta^w = Ez_1^4 - 1$, and the convergence, due to the CLT, is in the sense of distributions. Letting $W^o(\cdot)$ denote a Brownian bridge, define $\mathbf{y}_n = (y_1, \cdots y_n)^t = O_n^t \mathbf{z}_n$ where $\mathbf{z}_n = (z_1, \cdots, z_n)$. Introduce the process

$$Z_n(t) = \sqrt{\frac{1}{2n}} \sum_{i=1}^{[nt]} \left( y_i^2 - \frac{1}{n} \right)$$

The main result of this section is the

**Theorem A.1**

$$\{Z_n(t)\}_{t \in [0,1]} \xrightarrow{\mathcal{D}} \{W^o + \frac{\eta}{\sqrt{2}} t\}_{t \in [0,1]}, \tag{20}$$

*where $W^o$ and $\eta$ are independent and the convergence is in the sense of distributions in $D[0,1]$, the space of right continuous functions with left limits (RCLL), equipped with the Skorohod topology.*

To see why this says that the vector $\mathbf{y}_n$ is asymptotically isotropic, consider the special case when $\mathbf{z}_n$ is an i.i.d. Gaussian random vector, i.e. isotropic to start with. Then $\mathbf{y}_n$ is also an i.i.d. zero-mean Gaussian vector which is clearly isotropic. It is not difficult in this case to verify by a standard functional Central Limit Theorem that Theorem A.1 holds. What Theorem A.1 says is that $\mathbf{y}_n$ will be asymptotically isotropic even in the general case when the $z_i$'s are not. It's truth depends on the asymptotic isotropic property of the eigenvector matrix $O_n$.

A consequence of Theorem A.1, of use to us in this paper, is the following. Suppose $D_n$ is a diagonal matrix (possibly random) with entries monotone on the diagonal and eigenvalue distribution $L_n^D := \frac{1}{n}\sum_{i=1}^n \delta_{(D_n)_{ii}} \Rightarrow \mu_D$, such that $\mu_D$ possesses a continuous c.d.f. $F_D(x)$ with support on some compact set $[0, a]$, with possibly a jump at 0. Denote by $N_n = O_n D_n O_n^t$, and let

$$\theta_n = \mathbf{y}_n D_n \mathbf{y}_n^t$$

Here and in the sequel, we use $\mathbb{E}$ to denote expectation w.r.t. the randomness incurred in $D_n$, $\{z_i\}$ and $M_n$. Note that

$$\frac{1}{n}\mathbb{E}\theta_n = \frac{1}{n}\mathbb{E}\mathrm{tr}\, D_n \to_{n\to\infty} \int_0^a x\mu_D(dx)\,.$$

**Corollary A.2** *As $n \to \infty$,*

$$\sqrt{\frac{1}{2n}}\Big(\theta_n - \mathrm{tr}\, D_n\Big) \xrightarrow{\mathcal{D}} N(0, \sigma_\theta^2)\,. \tag{21}$$

*where*

$$\sigma_\theta^2 = \int_0^a x^2\mu_D(dx) + \left(\frac{\sigma_\eta^2}{2} - 1\right)\left(\int_0^a x\mu_D(dx)\right)^2\,. \tag{22}$$

**Proof of Corollary A.2** Let $F_n(x) = \int_0^{x^-} L_n^D(dy)$ denote the number of eigenvalues of $D_n$ not larger than $x$. Then, (assuming $D_n$ does not possess multiple eigenvalues except possibly at 0),

$$\begin{aligned}
\theta_n - \mathrm{tr}\, D_n &= \sum_{i=1}^n \left(y_i^2 - \frac{1}{n}\right)(D_n)_{ii} \\
&= \sqrt{2n}\sum_{i=1}^n (D_n)_{ii}\Big(Z_n((D_n)_{ii}^+) - Z_n((D_n)_{ii}^-)\Big) \\
&= \sqrt{2n}\int_0^a x dZ_n\Big(F_n(x)\Big) \\
&= -\sqrt{2n}\int_0^{-a} Z_n\Big(F_n(x)\Big)dx\,.
\end{aligned}$$

Since $F_n(x)$ converges to the continuous $F_D(x)$, the weak convergence of $Z_n(\cdot)$ is carried over to that of $Z_n(F_n(\cdot))$, and the latter converges in distribution to the Gaussian process

$$\{W^o\Big(F_D(x)\Big) + \frac{\eta}{\sqrt{2}} F_D(x)\}_{x\in[0,\infty)} .$$

Hence,

$$\frac{1}{\sqrt{2n}}(\theta_n - \operatorname{tr} D_n) \xrightarrow{\mathcal{D}} \int_0^a x \left[ d\Big(W^o(F_D(x)\Big) + \frac{\eta}{\sqrt{2}} dF_D(x) \right]$$
$$= \frac{\eta}{\sqrt{2}} \int_0^a (x dF_D(x) + \int_0^a d\Big(W^o(F_D(x)\Big) .$$

The convergence (21) and the value of the variance in (22) follow from evaluating the variance of the limiting Gaussian process. □

**Proof of Theorem A.1** The proof is a modification of the argument presented in [17]. Let $L_n^{\Lambda_n} = \frac{1}{n}\sum_{i=1}^n \delta_{(\Lambda_n)_{ii}}$, and $F_n^{\Lambda}(x) = \int_0^{x^-} L_n^{\Lambda_n}(dy)$ denote the number of eigenvalues of $\Lambda_n$ smaller than $x$. It is well known that $L_n^{\Lambda_n} \xrightarrow{w} \mu_\Lambda$, in the sense of weak convergence of distributions, with $F^\Lambda(y) = \int_0^{x^-} \mu_\Lambda(dy)$ denoting the appropriate distribution function. Then $F_n^\Lambda(x) \to F^\Lambda(x)$, uniformly, c.f. the argument in [17, pg. 1176].

By Theorem 4.1 in [17], for any fixed sequence of vectors with

$$\|\mathbf{x}_n\| = 1 , \quad \sum_{i=1}^n x_{n,i}^4 \underset{n\to\infty}{\longrightarrow} 0 ,$$

we have that

$$\left\{ \sqrt{\frac{n}{2}} \Big( \mathbf{x}_n^t M_n^r \mathbf{x}_n - \frac{1}{n} \operatorname{tr}(M_n^r) \Big) \right\}_{r=1}^\infty \xrightarrow{\mathcal{D}} \left\{ \int x^r dW^o\Big(F^\Lambda(x)\Big) \right\}_{r=1}^\infty$$

in the sense of convergence of laws in $\Re^\infty$. In particular, it follows that for any sequence $\varepsilon_n \to 0$,

$$\bar{d} := \tag{23}$$
$$\sup_{\{\mathbf{x}_n : \|\mathbf{x}_n\|=1, \sum_{i=1}^n x_{n,i}^4 < \varepsilon_n\}} d_{\mathcal{L}} \left( \left\{ \sqrt{\frac{n}{2}} \left( \mathbf{x}_n^t M_n^r \mathbf{x}_n - \frac{1}{n} \operatorname{tr}(M_n^r) \right) \right\}_{r=1}^\infty , \left\{ \int x^r dW^o\Big(F^\Lambda(x)\Big) \right\}_{r=1}^\infty \right) \underset{n\to\infty}{\longrightarrow} 0 ,$$

where $d_{\mathcal{L}}(a,b)$ denote the distance between the laws of the random variables $a, b$ in, say, the Lévy-Prohorov metric.

Next, note that

$$
\begin{aligned}
& \left\{\sqrt{\frac{1}{2n}}\left(\mathbf{z}_n^t M_n^r \mathbf{z}_n - \operatorname{tr}(M_n^r)\right)\right\}_{r=1}^{\infty} \\
= & \left\{\sqrt{\frac{1}{2n}}\left(\frac{\mathbf{z}_n^t}{\|\mathbf{z}_n\|} M_n^r \frac{\mathbf{z}_n}{\|\mathbf{z}_n\|} - \frac{\operatorname{tr}(M_n^r)}{n}\right)\right\}_{r=1}^{\infty} \|\mathbf{z}_n\|^2 \\
+ & \left\{\sqrt{\frac{1}{2n}} \frac{(\operatorname{tr} M_n^r)}{n} (\|\mathbf{z}_n\|^2)\right\}_{r=1}^{\infty} \\
\triangleq & \; Z_n^1 + Z_n^2,
\end{aligned}
$$

where $Z_n^i$ are random variables taking values in $\Re^\infty$.

Let $\varepsilon_n = n^{-\frac{1}{2}}$, and let $A_n = \{\mathbf{x}_n : \|\mathbf{x}_n\| = 1, \sum_{i=1}^n x_{n,i}^4 < \varepsilon_n\}$. Note that

$$
\sum_{i=1}^n \left(\frac{z_i}{\|\mathbf{z}_n\|}\right)^4 = \frac{\frac{1}{n}\sum_{i=1}^n z_i^4}{n \cdot \left(\frac{1}{n}\sum_{i=1}^n z_i^2\right)^2} \xrightarrow{P} 0,
$$

and further

$$
P\left(\frac{\mathbf{z}_n}{\|\mathbf{z}_n\|} \notin A_n\right) \leq C\frac{\varepsilon_n}{n} \underset{n\to\infty}{\longrightarrow} 0.
$$

Let $L_1^{\mathbf{z}_n}$ denote the law of $Z_n^1$ *conditioned* on $\mathbf{z}_n$, let $L^W$ denote the law of $\{\int x^r dW^o(F^\Lambda(x))\}_{r=1}^\infty$, and let $d(v,\mu)$ denote the Lévy-Prohorov distance between the laws $v, \mu$. Then, using (23),

$$
d(L_1^{\mathbf{z}_n}, L^W) \leq P\left(\frac{\mathbf{z}_n}{\|\mathbf{z}_n\|} \in A_n\right) + \bar{d} \underset{n\to\infty}{\longrightarrow} 0. \tag{24}
$$

This, together with the convergence $\frac{\operatorname{tr} M_n^r}{n} \to \int x^r dF_n^\Lambda(x)$ and $\frac{(\|\mathbf{z}_n\|^2 - n)}{\sqrt{n}} \xrightarrow{\mathcal{D}} N(0,\eta)$, imply that

$$
Z_n^1 + Z_n^2 \xrightarrow{\mathcal{D}} \left\{\int x^r dW^o\left(F^\Lambda(x)\right)\right\}_{r=1}^{\infty} + \left\{\frac{\eta}{\sqrt{2}}\int x^r dF_n^\Lambda(x)\right\}_{r=1}^{\infty} \tag{25}
$$

where $\eta$ is a $N(0,\sigma_\eta)$ random variable independent of $W^o$.

The next step consists of inverting the time change in (25). In view of the argument in Page 1191 of [17], (20) follows from (25) as soon as some tightness holds, that is as soon as one shows that for some $C > 0$,

$$
\mathbb{E}\Big(Z_n\big(F_n(0)\big)\Big)^4 \leq CE\Big(F_n(0)\Big)^2 \tag{26}
$$

and for any $0 \leq x_1 \leq x_2$,

$$
\mathbb{E}\Big(Z_n(F_n(x_2)) - Z_n(F_n(x_1))\Big)^4 \leq CE\Big(F_n(x_2) - F_n(x_1)\Big)^2, \tag{27}
$$

(compare with Theorem 4.2 of [17]). In fact, the proof of these facts follows closely the proof in [17], whose notations we adopt here. Since the proof of (26) is similar, we consider below only the proof of (27). Let $P^n = \{P_{ij}\}$ denote the projection matrix on the subspace of $\Re^n$ spanned by the eigenvectors of $M_n$ having eigenvalues in $[x_1, x_2]$. One checks immediately that

$$Z_n\Big(F_n(x_2)\Big) - Z_n\Big(F_n(x_1)\Big) = \sqrt{\frac{1}{2n}}(\mathbf{z}_n^t P^n \mathbf{z}_n - \operatorname{tr} P^n).$$

With $\gamma_{ij} = z_i z_j$, one sees that the LHS of (27) satisfies

$$\begin{aligned}
&\mathbb{E}\Big(Z_n(F_n(x_2)) - Z_n(F_n(x_1))\Big)^4 \\
&= \frac{1}{4n^2}\mathbb{E}\left(\sum_{i,j}\gamma_{ij}P_{ij} - \sum_i P_{ii}\right)^4 \\
&\le \frac{c}{n^2}\left(\mathbb{E}\left(\sum_{i\neq j}\gamma_{ij}P_{ij}\right)^4 + \mathbb{E}\left(\sum_{i=1}^n (z_i^2-1)P_{ii}\right)^4\right) \\
&\triangleq c(I_1 + I_2),
\end{aligned}$$

for some constant $c$ independent of $n$.

Following the same argument that led to (4.10) in [17], one finds that

$$\begin{aligned}
I_1 \quad = \quad & \frac{n-1}{n}\Big(12(n-2)Ez_1^4(Ez_1^2)^2E(P_{12}^2P_{13}^2) + 3(n-2)(n-3)(Ez_1^2)^4E(P_{12}^2P_{34}^2) \\
& +12(n-2)(n-3)(Ez_1^2)^4E(P_{12}P_{23}P_{34}P_{14}) + 2E(P_{12}^4)(Ez_1^4)^2\Big)
\end{aligned}$$

Using the fact that $P$ is a projection matrix, we have that

$$P_{13} = \sum_{j\geq 4} P_{3j}P_{1j} + P_{31}P_{11} + P_{32}P_{12} + P_{33}P_{13}\,.$$

Using the fact that expectations are invariant with respect to permutations, we conclude, as in [17], that

$$\begin{aligned}
(n-2)(n-3)E(P_{12}P_{23}P_{34}P_{14}) &\le EP_{12}^2 + 2E(P_{11}P_{22}P_{12}^2) + 2EP_{11}^2P_{12}^2 \\
&\le EP_{11}P_{22} + 2E(P_{11}^2P_{22}^2) + 2E(P_{11}P_{22}P_{11}^2) \\
&\le EP_{11}P_{22} + 2E(P_{11}P_{22}) + 2E(P_{11}P_{22}) \\
&= 5E(P_{11}P_{22})
\end{aligned}$$

where we made repeated use of $P_{12}^2 \le P_{11}P_{22}$ and $P_{11} \ge \max(P_{11}P_{22}, P_{11}^2)$. Similarly,

$$(n-2)E(P_{12}^2P_{13}^2) \le E(P_{11}P_{12}^2) \le E(P_{11}^2P_{22}) \le E(P_{11}P_{22})$$

and

$$(n-3)E(P_{12}^2P_{34}^2) \leq E(P_{12}^2P_{33})$$

leading to

$$(n-2)(n-3)E(P_{12}^2P_{34}^2) \leq (n-2)E(P_{12}^2P_{33}).$$

However, using the identity

$$P_{33}\left(\sum_{j\neq 1,3} P_{ij}P_{ij} + P_{11}P_{11} + P_{13}P_{13}\right) = P_{33}P_{11},$$

we obtain

$$(n-2)E(P_{33}P_{12}^2) + EP_{11}^2P_{33} + EP_{13}^2P_{33} = EP_{11}P_{33},$$

leading to

$$(n-2)EP_{33}P_{12}^2 \leq EP_{11}P_{33} = EP_{11}P_{22}.$$

Hence,

$$(n-2)(n-3)E(P_{12}^2P_{34}^2) \leq EP_{11}P_{22}.$$

Combining the above, we conclude that

$$I_1 \leq K_1E(P_{11}P_{22})$$

for some constant $K_1$ which depends on $Ez_1^4$ only. Similarly, for some constant $K_2$ independent of $n$,

$$\begin{aligned} I_2 &\leq \frac{1}{n}\mathbb{E}P_{11}^4 + \frac{3(n-1)}{n}(\mathbb{E}P_{11}^2P_{22}^2)\Big(E(z_i^2-1)^2\Big)^2 \\ &\leq \frac{1}{n}\mathbb{E}P_{11}^4 + K_2\mathbb{E}(P_{11}P_{22}). \end{aligned}$$

Hence, for some $K_3$ independent of $n$,

$$\begin{aligned} c(I_1+I_2) &\leq K_3\left(\frac{1}{n}\mathbb{E}P_{11}^4\frac{(n-1)}{n}\mathbb{E}(P_{11}P_{22})\right) \\ &= K_3\mathbb{E}\left(\frac{1}{n}\sum_i P_{ii}\right)^2 = K_3\Big(F_n(x_2)-F_n(x_1)\Big)^2, \end{aligned}$$

as required. □

# B Proof of Lemma 4.3

Throughout this proof, we follow the notations of Section 4, while for simplicity taking $P=1$. That is, we consider the matrix $S_1 := [\mathrm{s}_2,\ldots,\mathrm{s}_M]$, and denote by $\{\lambda_i\}_{i=1}^L$ the eigenvalues of $S_1S_1^t$. The case for general $P$ follows directly from a rescaling of $\sigma^2$.

We use various constants $C, C_i$, whose values may change from line to line and are always independent of $L$ (but may depend on $\sigma$). We also use constants $K_p$, whose values may change from line to line, and which are independent of $L$ and $\sigma$.

Before starting, we recall the Burkholder inequality, c.f. [4]: If $\{\theta_i\}$ is a martingale difference sequence with respect to an increasing filtration $\mathcal{G}_i$, i.e. $\theta_i$ is $\mathcal{G}_i$ measurable and $E(\theta_i|\mathcal{G}_{i-1}) = 0$, then, for any $p > 1$,

$$E\left|\sum_{i=1}^{k}\theta_i\right|^p \leq K_p E\left|\sum_{i=1}^{k}\theta_1^2\right|^{p/2} . \tag{28}$$

Using the fact that if $\{\theta_i\}$ is square integrable then $\theta_i^2 - E(\theta_i^2|\mathcal{G}_{i-1})$ is again a martingale difference sequence, and iterating $\lceil \log_2 p \rceil$ times this inequality, one also gets that for $p \geq 2$,

$$E\left|\sum_{i=1}^{k}\theta_i\right|^p \leq K_p E\sum_{i=1}^{k}|\theta_1^p| + K_p E\left(\sum_{i=1}^{k}E(\theta_i^2|\mathcal{G}_{i-1})\right)^{p/2} . \tag{29}$$

We emphasize that in (28) and (29), $K_p$ does not depend on $k$. Let $A := (S_1 S_1^t + \sigma^2 I)^{-1}$. Noting that

$$S_1 S_1^t = \sum_{j=2}^{M} \mathrm{s}_j \mathrm{s}_j^t ,$$

we let $A_j := (S_1 S_1^t + \sigma^2 I - \mathrm{s}_j \mathrm{s}_j^t)^{-1}$. Since

$$\operatorname{tr} A = \sum_{i=1}^{L} \frac{1}{\lambda_i + \sigma^2} ,$$

we need only estimate $E(\operatorname{tr} A - E\operatorname{tr} A)^2$.

Let $\mathcal{F}_j = \sigma(\mathrm{s}_i, 2 \leq i \leq j)$, and write $E_j(\cdot) = E(\cdot|\mathcal{F}_j)$. Using the identity

$$\operatorname{tr} A_j - \operatorname{tr} A = \frac{\mathrm{s}_j^t A_j^2 \mathrm{s}_j}{1 + \mathrm{s}_j^t A_j \mathrm{s}_j} ,$$

we have that

$$\begin{aligned}
\operatorname{tr} A - E\operatorname{tr} A &= \sum_{j=2}^{M} E_j \operatorname{tr} A - E_{j-1}\operatorname{tr} A \\
&= \sum_{j=2}^{M}(E_j - E_{j-1})\operatorname{tr} A_j + \sum_{j=2}^{M}(E_j - E_{j-1})\frac{\mathrm{s}_j^t A_j^2 \mathrm{s}_j}{1 + \mathrm{s}_j^t A_j \mathrm{s}_j} \\
&= \sum_{j=2}^{M}(E_j - E_{j-1})\frac{\mathrm{s}_j^t A_j^2 \mathrm{s}_j}{1 + \mathrm{s}_j^t A_j \mathrm{s}_j} .
\end{aligned}$$

We now define

$$a_j = \frac{1}{L}\mathrm{tr}\,(A_j^2), \quad \alpha_j = \mathbf{s}_j^t A_j^2 \mathbf{s}_j - a_j \quad \omega_j = \frac{1}{1+\mathbf{s}_j^t A_j \mathbf{s}_j}, \quad b_L = \frac{1}{1+L^{-1}E\mathrm{tr}\,A},$$

$$\zeta_j = \mathbf{s}_j^t A_j \mathbf{s}_j - L^{-1}E\mathrm{tr}\,A_j \quad \hat{\zeta}_j = \mathbf{s}_j^t A_j \mathbf{s}_j - L^{-1}\mathrm{tr}\,A_j\,.$$

Using some algebra, one arrives at

$$\begin{aligned}
\sum_{j=2}^{M}(E_j - E_{j-1})\frac{\mathbf{s}_j^t A_j^2 \mathbf{s}_j}{1+\mathbf{s}_j^t A_j \mathbf{s}_j} &= b_L \sum_{j=2}^{M} E_j \alpha_j - b_L^2 \sum_{j=2}^{M} E_j a_j \hat{\zeta}_j \\
&\quad - b_L^2 \sum_{j=2}^{M} (E_j - E_{j-1})(\alpha_j \zeta_j - \mathbf{s}_j^t A_j^2 \mathbf{s}_j \omega_j \zeta_j^2) \\
&:= W_1 - W_2 - W_3\,.
\end{aligned}$$

Hence, since $0 < \alpha < \infty$ and $b_L$ is uniformly bounded due to $\sigma > 0$, it will be enough to estimate $E(W_1/b_L)^2$, $E(W_i/b_L^2)^2$, $i = 2, 3$.

Turning to the first term, note that $E_j(\alpha_j)$ is a martingale difference sequence, i.e. $E_{j-1}(E_j\alpha_j) = 0$. Hence, by Burkholder's inequality (29), for each $p \geq 2$, there exists some universal constant $K_p$ such that

$$E(W_1/b_L)^2 = E(\sum_{j=2}^{M}(E_j\alpha_j))^2 \leq K_p \left( E\left(\sum_{j=2}^{M} E_{j-1}(E_j\alpha_j)^2\right)^{p/2} + E\left(\sum_{j=2}^{M}(E_j\alpha_j)^p\right)\right). \tag{30}$$

Let $\mathbf{1}_L$ denote the vector $(1, 1, \ldots, 1)$ in $\Re^L$. Then, since the eigenvalues of the matrices $A_j$ are bounded above by $\sigma^2$, we have for $p = 2, 4$ that

$$\sum_{j=2}^{M}(E_j\alpha_j)^p \leq \sum_{j=2}^{M} E\alpha_j^p = ME\alpha_2^p \leq MC_5 E\|\mathbf{s}_2^t - L^{-1/2}\mathbf{1}_L\|^{2p} \leq C_5 M/L^p\,. \tag{31}$$

On the other hand,

$$\begin{aligned}
E_{j-1}(E_j\alpha_j)^2 &\leq E_{j-1}\left((\mathbf{s}_j - L^{-1/2}\mathbf{1}_L)^t A_j^2(\mathbf{s}_j - L^{-1/2}\mathbf{1}_L) + 2L^{-1/2}\mathbf{1}^t A_j^2(\mathbf{s}_j - L^{-1/2}\mathbf{1}_L)\right)^2 \\
&\leq C_6 L^{-1}\,.
\end{aligned}$$

Combining the above estimates, with $p = 2$, one gets that

$$E(W_1/b_L)^2 \leq C_7\,.$$

The estimate for $W_2$ is similar, modulo the following auxiliary results, valid for $p = 2, 4$:

$$E(|\hat{\zeta}_j|^p) \leq C_7 L^{-p/2}, \quad E(|\zeta_j - \hat{\zeta}_j|^p \leq C_7 L^{-p/2}, \tag{32}$$

implying that

$$E(|\zeta_j|^p) \leq C_8 L^{-p/2} \,. \tag{33}$$

To see (32), recall Lemma 2.7 of [2], which represents a variant of Lemma 3.2: There exists for each $p \geq 2$ a universal constant $K_p$ such that, for any deterministic matrix $D$, and any vector of i.i.d. random variables $\mathbf{x} = (x_1, \ldots, x_n)$ with $Ex_1 = 0$ and $Ex_1^2 = 1$,

$$E|\mathbf{x}^t D\mathbf{x} - \operatorname{tr} D|^p \leq K_p (E(|x_1|^4 \operatorname{tr} DD^t)^{p/2} + E|x_1|^{2p} \operatorname{tr} (DD^t)^{p/2}) \,.$$

Taking $p = 2, 4$, $\mathbf{x} = \mathbf{s}_j\sqrt{L}$, $D = A_j$ and the expectation with respect to $\mathbf{s}_j$, and using the independence of $A_j$ and $\mathbf{s}_j$ together with $\operatorname{tr} A_j^k \leq L\sigma^{-2k}$, we obtain that

$$E_{j-1}(|\hat{\zeta}_j|^p) \leq C_9 L^{-p/2} \,, \tag{34}$$

and hence $E(|\hat{\zeta}_j|^p) \leq C_7 L^{-p/2}$. On the other hand, letting for any $j > 2$ $A_{2j} = (A_2^{-1} - \mathbf{s}_j^t \mathbf{s}_j)^{-1}$, and noting that still $\operatorname{tr} A_{2j}^k \leq L\sigma^{2k}$,

$$\begin{aligned} E|\zeta_j - \hat{\zeta}_j|^p &= E|\zeta_2 - \hat{\zeta}_2|^p = E\left|\frac{1}{L}\sum_{j=3}^{M}(E_j \operatorname{tr} A_2 - E_{j-1}\operatorname{tr} A_2)\right|^p \\ &= E\left|\frac{1}{L}\sum_{j=3}^{M}(E_j \operatorname{tr}(A_2 - A_{2j}) - E_{j-1}\operatorname{tr}(A_2 - A_{2j}))\right|^p \\ &= E\left|\frac{1}{L}\sum_{j=3}^{M}(E_j - E_{j-1})\frac{\mathbf{s}_j^t A_{2j}^2 \mathbf{s}_j}{1 + \mathbf{s}_j^t A_{2j}\mathbf{s}_j}\right|^p \\ &\leq K_p \frac{1}{L^p} E\left|\sum_{j=3}^{M} |(E_j - E_{j-1})\frac{\mathbf{s}_j^t A_{2j}^2 \mathbf{s}_j}{1 + \mathbf{s}_j^t A_{2j}\mathbf{s}_j}|^2\right|^{p/2} \leq C_{10}\frac{1}{L^{p/2}} \,, \end{aligned}$$

where the Burkholder inequality (28) was used in the next to last step.

Returning to the estimate on $W_2/b_L^2$, recall that all $a_j$-s are deterministically bounded, uniformly in $L$. Therefore, using Burkholder's inequality (29) in the first step, and (34) in the third,

$$\begin{aligned} E|\sum_{j=2}^{M} E_j(a_j\hat{\zeta}_j)|^2 &\leq C_{11}\left(E(\sum_{j=2}^{M} E_{j-1}(E_j(a_j\hat{\zeta}_j))^2) + \sum_{j=2}^{M} E|E_j(a_j\hat{\zeta}_j)|^2\right) \\ &\leq C_{12}\left(E(\sum_{j=2}^{M} E_{j-1}(\hat{\zeta}_j)^2) + \sum_{j=2}^{M} |E(\hat{\zeta}_j)^2|\right) \leq C_{13}(ML^{-1}) \,, \end{aligned}$$

proving the desired estimate on $W_2$.

The argument involving $W_3$ is similar, only simpler: First, note that $\omega_j \mathbf{s}_j^t A_j^2 \mathbf{s}_j$ is bounded uniformly, and so is $\omega_j$. Therefore, using again Burkholder's inequality (28) in the first step,

$$
\begin{aligned}
E(W_3/b_L^2)^2 &= E\left(\sum_{j=2}^{M}(E_j - E_{j-1})(\alpha_j\zeta_j - \mathbf{s}_j^t A_j^2 \mathbf{s}_j \omega_j \zeta_j^2)\right)^2 \\
&\le K_p E\left(\sum_{j=2}^{M}((E_j - E_{j-1})(\alpha_j\zeta_j - \mathbf{s}_j^t A_j^2 \mathbf{s}_j \omega_j \zeta_j^2))^2\right) \\
&\le C_{14}\sum_{j=2}^{M}\left(E(\alpha_j\zeta_j)^2 + E(\zeta_j^4)\right) = (M-1)(E(\alpha_2^2\zeta_2^2) + E(\zeta_2^4))
\end{aligned}
$$

In view of (33), we need only to recall from (31) that $E(\alpha_2^4) \le C_{15}L^{-p}$. □

**Remark** Another possible route to the proof of Lemma 4.3 is as follows. Note that the function $(1 + x/\sigma^2)^{-1}$ is analytic in the (open) disk $|x| < \sigma^2$. For $\sigma^2/P > (1+\sqrt{\alpha})^2$, this disk includes the support of the limiting measure $G^*(\cdot)$, which is $[P(1-\sqrt{(}\alpha)^2, P(1+\sqrt{\alpha})^2]$. Expanding in Taylor series the function $(1+x/\sigma^2)^{-1}$ up to order $k = C\log L$, and controlling the reminder by using the analyticity, it follows that Lemma 4.3 holds as soon as, for some $C_1$ large enough, $k \le C_1(\log L)$,

$$
\mathrm{Var}\left(\sum_{i=1}^{L}\frac{\lambda_i^k}{(1+\sqrt{\alpha})^k}\right) < C_2 . \tag{35}
$$

(All constants $C_i$ are taken to be independent of $L$). Such an estimate is the main result in [19], who under stronger moment assumptions deal with Wigner matrices and not with sample covariance matrices, and show that (35) actually holds for $k = o(\sqrt{L})$. But for sample covariance matrices, one may use the same construction as [19], except that instead of considering paths $i_0 \to i_1 \to i_2 \to \ldots \to i_k$ as in [19], one considers paths $i_0 \to j_0 \to i_1 \to j_1 \to \ldots \to i_k \to j_k$, where $i_m \in \{1,\ldots,L\}$ and $j_m \in \{1,\ldots,M\}$. The parameter $s$ in [19] is replaced by a pair $s_1, s_2$ with $s_1 + s_2 = s$, keeping the random walk parameterization as in [19]. We do not see however how to extend this argument to the range $\sigma^2 \le P(1+\sqrt{\alpha})^2$. On the other hand, for $\sigma^2 > P(1+\sqrt{\alpha})^2$, the argument in [19] actually shows that $L(\beta^{(L)} - \beta^*)$ is asymptotically normal.

We note that for $k = o(\log L/\log\log L)$, the cruder estimates contained in [18, Lemma 1] are enough to yield (35): indeed, for $p$ finite and independent of $L$ this is the content of the proof there, while by bounding there the number of set partitions of $\{i_1,\ldots,i_r,i_1',\ldots,i_r'\}$, $\{k_1,\ldots,k_r,k_1',\ldots,k_r'\}$ one extends the conclusion to $k$'s as above. Unfortunately, this technique seems to break down when $k = O(\log L)$.

We finally note that besides the condition $E(v_{ij}^4) < \infty$, the assumption $E(v_{ij}^8) < \infty$ was used only in bounding $W_3$; a tightening of this argument, valid under the condition $E(v_{ij}^4) < \infty$, seems possible, following [2, Section 4], but we do not pursue this direction here.

# C Proof of Lemma 4.4

Define

$$S_i := [\mathbf{s}_1, \mathbf{s}_2, \ldots, \mathbf{s}_{i-1}, \mathbf{s}_{i+1}, \ldots, \mathbf{s}_M]$$

and

$$S := [\mathbf{s}_1, \mathbf{s}_2, \ldots, \mathbf{s}_M].$$

As in Appendix B, we will for simplicity take $P = 1$. The general result follows from a rescaling of $\sigma^2$. Now,

$$\beta_i^{(L)} = \mathbf{s}_i^t (S_i S_i^t + \sigma^2 I)^{-1} \mathbf{s}_i . \tag{36}$$

is the SIR attained by user $i$. In [21, Eq. (27)], a key equation relating the achieved SIR's of the users and the trace of $(SS^t + \sigma^2 I)^{-1}$ was derived:

$$\frac{1}{L} \sum_{i=1}^{M} \frac{\beta_i^{(L)}}{1 + \beta_i^{(L)}} = 1 - \frac{\sigma^2}{L} \mathrm{tr}\,(SS^t + \sigma^2 I)^{-1}. \tag{37}$$

It follows from Lemma 3.2 that

$$\bar{\beta}_i^{(L)} := E(\beta_i^{(L)}) = E\left[L^{-1} \mathrm{tr}\,(S_i S_i^t + \sigma^2 I)^{-1}\right].$$

Let

$$\bar{\beta}^{(L)} := E\left[L^{-1} \mathrm{tr}\,(SS^t + \sigma^2 I)^{-1}\right].$$

It follows from lemmas 4.2 and 4.3 that for large $L$, each of the $\beta_i^{(L)}$ is close to $\bar{\beta}_i^{(L)}$, which in turn is close to $\bar{\beta}^{(L)}$. Moreover, $L^{-1}\mathrm{tr}\,(SS^t + \sigma^2 I)^{-1}$ is also close to $\bar{\beta}^{(L)}$. Substituting these approximations into (37) gives us an approximate fixed point equation in $\bar{\beta}^{(L)}$:

$$\alpha \frac{\bar{\beta}^{(L)}}{1 + \bar{\beta}^{(L)}} \approx 1 - \sigma^2 \bar{\beta}^{(L)}.$$

The *exact* fixed point equation has a unique positive solution, which is precisely the limiting value $\beta^*$. (In fact, the formula (11) for $\beta^*$ is obtained by solving this quadratic equation.) Thus, to estimate how far $\bar{\beta}^{(L)}$ is from $\beta^*$, we need estimates on how far each of the $\beta_i^{(L)}$ deviates from $\bar{\beta}^{(L)}$. This is the main idea of the following development.

One can write

$$\bar{\beta}_i^{(L)} = \bar{\beta}^{(L)} + \delta_i^{(L)}$$

$$\delta_i^{(L)} = \frac{1}{L} E\left[\mathrm{tr}\,(S_i S_i^t + \sigma^2 I)^{-1} - \mathrm{tr}\,(SS^t + \sigma^2 I)^{-1}\right].$$

By the matrix-inversion lemma,

$$(S_i S_i^t + \sigma^2 I)^{-1} - (SS^t + \sigma^2 I)^{-1} = (S_i S_i^t + \sigma^2 I)^{-1} \mathbf{s}_i \mathbf{s}_i^t (SS^t + \sigma^2 I)^{-1}$$

so

$$
\begin{aligned}
|\delta_i^{(L)}| &= |\frac{1}{L}E\left[\mathrm{tr}\ ((S_iS_i^t+\sigma^2 I)^{-1}\mathbf{s}_i\mathbf{s}_i^t(SS^t+\sigma^2 I)^{-1})\right]| \\
&= \frac{1}{L}E\left[\mathrm{tr}\ (\mathbf{s}_i^t(SS^t+\sigma^2 I)^{-1}(S_iS_i^t+\sigma^2 I)^{-1}\mathbf{s}_i)\right] \\
&\leq \frac{1}{L}E\left[\frac{1}{\sigma^4}\|\mathbf{s}_i\|^2\right] \\
&= \frac{1}{\sigma^4 L} \qquad (38)
\end{aligned}
$$

Also,

$$
\begin{aligned}
\mathrm{Var}[\beta_i^{(L)}] &= E\left[\left(\beta_i^{(L)} - L^{-1}\mathrm{tr}\,(S_iS_i^t+\sigma^2 I)^{-1} + L^{-1}\mathrm{tr}\,(S_iS_i^t+\sigma^2 I)^{-1} - \bar{\beta}_i^{(L)}\right)\right] \\
&= E\left[\left(\beta_i^{(L)} - \mathrm{tr}\,(S_iS_i^t+\sigma^2 I)^{-1}\right)^2\right] \\
&\quad +2E\left[\left(\beta_i^{(L)} - L^{-1}\mathrm{tr}\,(S_iS_i^t+\sigma^2 I)^{-1}\right)\left(L^{-1}\mathrm{tr}\,(S_iS_i^t+\sigma^2 I)^{-1} - \bar{\beta}_i^{(L)}\right)\right] \\
&\quad +E\left[\left(L^{-1}\mathrm{tr}\,(S_iS_i^t+\sigma^2 I)^{-1} - \bar{\beta}_i^{(L)}\right)^2\right]
\end{aligned}
$$

By Lemma 3.2, the first term above is bounded by $C_1\sigma^4/L$ for some constant $C_1$ that depends only on the fourth moment of $v_{11}$, and the second term is 0. By Lemma 4.3, the third term is bounded by $C_2/L^2$ for some constant $C_2$ independent of $L$ and $i$. Hence $\mathrm{Var}[\beta_i^{(L)}] \leq C_3/L$ for some constant $C_3$ independent of $L$ and $i$.

Combining this with (38), we can now write:

$$\beta_i^{(L)} = \bar{\beta}^{(L)} + \Delta_i^{(L)}, \qquad (39)$$

where

$$|E[\Delta_i^{(L)}]| \leq \frac{1}{\sigma^4 L} \qquad (40)$$

and

$$E[(\Delta_i^{(L)})^2] \leq \frac{C_4}{L}$$

for some constant $C_4$ independent of $L$ and $i$.

Substituting (39) into the key equation (37),

$$\frac{M}{L} - \frac{1}{L}\sum_{i=1}^{M}\frac{1}{1+\bar{\beta}^{(L)}+\Delta_i^{(L)}} = 1 - \frac{\sigma^2}{L}\mathrm{tr}\,(SS^t+\sigma^2 I)^{-1}. \qquad (41)$$

Let $\nu_i^{(L)} := \Delta_i^{(L)}/(1+\bar{\beta}^{(L)})$. Then,

$$\frac{1}{1+\bar{\beta}^{(L)}+\Delta_i^{(L)}} = \frac{1}{1+\bar{\beta}^{(L)}}\frac{1}{1+\nu_i^{(L)}} = \frac{1}{1+\bar{\beta}^{(L)}}(1-\nu_i+\nu_i^2\frac{1}{(1+\xi_i)^3}),$$

for some $\xi_i := \xi_i(\nu_i^{(L)})$ satisfying $\xi_i \in [0, \nu_i^{(L)}] \cup [\nu_i^{(L)}, 0]$. Note that

$$\begin{aligned}\nu_i^{(L)} &= \frac{\beta_i^{(L)} - \bar{\beta}^{(L)}}{1 + \bar{\beta}^{(L)}} \\ &\geq -1 + \frac{1}{1 + \bar{\beta}^{(L)}} \qquad \text{since } \beta_i^{(L)} \geq 0 \\ &\geq -1 + \frac{\sigma^2}{1 + \sigma^2} \qquad \text{since } \bar{\beta}^{(L)} \leq 1/\sigma^2.\end{aligned}$$

Hence $(1 + \xi_i)^{-1} \leq C_5$ for some deterministic constant $C_5$ independent of $L$. Substituting in (41) and taking expectations, using the fact from (40) that

$$|E[\Delta_i^{(L)}]| \leq \frac{1}{\sigma^4 L}$$

and that

$$E[L^{-1}\mathrm{tr}\,(SS^t + \sigma^2 I)^{-1}] = \bar{\beta}^{(L)},$$

we get, for some $C_6$ independent of $L$,

$$\left| \frac{M}{L} - \frac{M}{L(1 + \bar{\beta}^{(L)})} - 1 + \sigma^2 \bar{\beta}^{(L)} \right| \leq \frac{C_6}{L},$$

and hence, for some $C_7$ independent of $L$,

$$\left| \alpha - \frac{\alpha}{1 + \bar{\beta}^{(L)}} - 1 + \sigma^2 \bar{\beta}^{(L)} \right| \leq \frac{C_7}{L}.$$

But $\beta^* := \int (\lambda + \sigma^2)^{-1} dG^*(\lambda)$ is the unique solution of the equation

$$\alpha - \frac{\alpha}{1 + \beta} - 1 + \sigma^2 \beta = 0,$$

and moreover, it can be easily seen that the solution of this equation is a differentiable function of the right-hand side at 0. Hence,

$$\limsup_{L \to \infty} L(\bar{\beta}^{(L)} - \beta^*) < \infty$$

The lemma now follows from (38). □