Unified Spatial Diversity Combining and Power Allocation for CDMA Systems in Multiple Time-Scale Fading Channels*

Junshan Zhang, † Edwin K. P. Chong‡ Ioannis Kontoyiannis§

January 29, 2001

Abstract

In a mobile wireless system, fading effects can be classified into large-scale (long-term) effects and small-scale (short-term) effects. We use transmission power control to compensate for large-scale fading and exploit receiver antenna (space) diversity to combat small-scale fading. We show that the interferences across the antennas are jointly Gaussian in a large system, and then characterize the signal-to-interference ratio (SIR) for both independent and correlated (across the antennas) small-scale fading cases. Our results show that when each user's small-scale fading effects are independent across the antennas, there is a clear separation between the gains of transmission power control and diversity combining, and the two gains are additive (in decibels). When each user's small-scale fading effects are correlated across the antennas, we observe that in general the gains of transmission power control and diversity combining are coupled. However, when the noise level diminishes to zero, using maximum ratio combining (MRC) "decouples" the gains and achieves the same diversity gain as in the independent case. We then characterize the Pareto-optimal (minimum) transmission power allocation for the cases of perfect and noisy knowledge of the desired user's large-scale fading effects. We find that using antenna diversity leads to significant gains for the transmission power.

Index Terms: CDMA, power control, space diversity, maximum ratio combining, selection combining, large-scale fading, small-scale fading, MMSE.

1 Introduction

In a mobile radio communication system, signal fading may severely degrade the system performance, and is a dominant source of impairment. Fading arises from randomly-delayed scattering, reflecting, and diffracting of electromagnetic waves in a random medium. According to their time scales, fading effects can be classified into two categories (as has been verified experimentally [14,

*J. Zhang and E. Chong were supported in part by the National Science Foundation through grant ECS-9501652.

†Junshan Zhang is with the Department of Electrical Engineering, Arizona State University, Tempe, AZ 85287-7206 (junshan.zhang@asu.edu).

‡Edwin K. P. Chong is with the School of Electrical and Computer Engineering, Purdue University, West Lafayette, IN 47907 (echong@ecn.purdue.edu).

§Ioannis Kontoyiannis is with the Department of Statistics, Purdue University, West Lafayette, IN 47907 (yiannis@stat.purdue.edu).

Chapter 8]): large-scale (long-term) effects and small-scale (short-term) effects. Large-scale fading is on the order of seconds, while small-scale fading is on the order of milliseconds. A more detailed description of fading effects is given in the following (see also [6] [8, Chapters 1, 2] [14, Chapter 8]):

- Large-scale effects include (1) distance-related attenuation, and (2) slow-shadowing fading, which is due to the terrain, buildings, and other obstacles that lie between the transmitter and receiver. Large-scale effects cause relatively slow variations in the (mean) signal strength as a mobile user moves through space. Large-scale fading is usually modeled as a log-normally distributed random variable.

- Small-scale effects are due to the scattering and/or reflections of the transmitted signals off surrounding objects. Small-scale effects may cause rapid and large swings in signal strength and are superimposed on top of the large-scale effects. Small-scale fading is typically modeled as a complex Gaussian random variable.

In current mobile wireless systems, the main traffic is typically voice and its transmission rate is around 10Kbps, which implies that large-sale fading may remain constant over a region spanning thousands of information symbols. Hence it is reasonable to assume that reliable estimates of large-sale fading are available. Future wireless systems are expected to be able to accommodate multimedia traffic and the data transmission rate will be much higher. Therefore, in these systems even small-scale fading may change little in the duration of many information symbols, which implies that it is also possible to get reliable side information about small-scale fading in such cases. Thus motivated, we assume in this paper that estimates of each user's large-scale fading are available at both its transmitter and receiver, and knowledge of small-scale fading depends on a specific communication scenario. For example, in a highly mobile communication system it is more reasonable to assume only partial side information about small-scale fading available, whereas in a fixed wireless communication scenario it is still possible to get reasonably reliable estimates of small-scale fading (see, e.g., [17]).

Little attention has been paid to the study of the system performance in a multiple time-scale fading environment. Our aim in this paper is to provide some first steps along this line. We consider a single-cell CDMA system, and our strategy is to use transmission power control to compensate for large-scale fading (including shadowing and distance-related attenuation) and to exploit antenna arrays to combat small-scale fading. The underlying rationale is as follows:

- Antenna arrays can provide space diversity to reduce the depth of fades and/or the fade duration caused by small-scale fading, by supplying the receiver with multiple replicas of the transmitted signal that have passed through different diversity channels [8, 18, 22].

- Because all the receiver antennas are placed at the same base station, large-scale fading affects all the diversity channels more or less identically [8, Chapter 5]. Therefore, we build on transmission power control to combat large-scale fading. In fact, practical power control algorithms can often respond quickly enough to compensate for large-scale effects but cannot compensate for small-scale effects [6].

Without loss of generality, we fix the number of antenna elments at the base station as L. We assume that a linear MMSE filter is applied to de-spread the received signal at each antenna,[1] and the antenna outputs are combined linearly. That is, the receiver is a concatenation of a bank of temporal MMSE filters and a linear spatial combiner. The focus of this paper is on large systems with both K (number of users) and N (processing gain) going to infinity while the ratio $\alpha = K/N$ is held fixed. We assume that a power control mechanism adjusts the transmission power as a function of the estimated large-scale fading. This model is applicable to many systems. For example, it is applicable to large systems with SIR-driven power control because in large systems the SIR (attained by the MMSE receiver) corresponding to unit received power converges to a constant (see, e.g., [20]) and hence the main task of power control is then to adjust each user's transmission power to combat the fading it experiences. It is also applicable to systems with *power balancing* [21].

We study the performance of a single-cell CDMA system with power control and space diversity. In our analysis, we assume that large-scale fading effects are known to the receiver although they are realizations of random variables. This is because large-scale fading effects remain roughly unchanged over thousands of information symbols. We assume throughout that the small-scale fading effects are independent across different users. We consider two contrasting cases—independent and correlated small-scale fading effects across the antennas for each user. Among the many diversity combining schemes, this paper focuses mainly on the maximum ratio combining (MRC) method, which requires perfect knowledge of the desired user's small-scale fading effects. We focus on the MRC method because it gives the best performance among all possible linear combiners and serves as a benchmark. We also present results on the performance of the selection combining (SC) method, which appears to be the simplest to implement [8]. The SC method chooses the branch with the highest signal-to-interference ratio (SIR) and does not require explicit information of fading effects, making it more amendable to implementation even in a highly mobile communication system.

In the case of independent small-scale fading effects, our results on the SIR show that there is a clear separation between the gains of transmission power control and of diversity combining (in a sense to be made clear in Theorem 3.1). Because of this separation, the two gains are additive (in decibels), and the diversity combining behaves the same here as in a single-user system with antenna arrays. Based on the above results, we characterize the Pareto-optimal (minimum) transmission power allocation for the following two cases: perfect and noisy knowledge of the desired user's large-scale fading effects. We find that using antenna diversity leads to significant gains for the transmission power. For example, there is nearly a 6dB gain for the transmission power allocation with four versus two receiver antennas when the MRC method is used and the load α is moderate; the gain in the SC method is still pronounced although it is smaller than in the MRC method. Moreover, the increase of network capacity by using antenna diversity is significant.

In the case of correlated small-scale fading effects, we find that in general the gains of transmission power control and of diversity combining are coupled (made precise in Theorem 4.1). However, when the noise level diminishes to zero, using MRC results in the "decoupling" of the gains and

[1] A similar but simpler analysis can also be carried through when matched filters or decorrelators are employed.

achieves the same diversity gain as in the case of independent small-scale fading. The intuition behind this fact is that CDMA systems are interference-limited and correlated interferences across the antennas can be whitened.

In related work, it was proposed in [12] to employ antenna arrays at the base station to increase the network capacity of CDMA systems. Reference [12] assumed a matched filter for each user and *perfect instantaneous power control* for intracell users, that is, all the users within one cell have the same received power at any instant. Recent work [5] studied the network capacity region with and without power constraints for CDMA systems with antenna arrays. In [5], it was assumed that perfect knowledge of the fading effects is available to construct receivers with various degrees of complexity. In [15], an iterative algorithm was developed that jointly updates transmission powers and beamforming weights so that it converges to the jointly optimal beamforming and power vector. Reference [23] extended the above idea to joint power control and optimal filtering in both temporal and spatial domains. The algorithms in both [15] and [23] assumed fixed channel gains (fading effects). A key feature distinguishing our work from the previous ones is that we explicitly take into account the different time scales of the large-scale and small-scale fading effects and impose more realistic assumptions on the fading effects and hence on power control and diversity combining.

An outline of the rest of this paper is given as follows. The next section contains our model description for the channel and signals, and the assumptions on the side information of fading effects. We also describe the receiver structure. In Sections 3 and 4, we study system performance for independent and correlated small-scale fading effects, respectively. We then address the problems of transmission power allocation in Section 5. To illustrate our results, we provide some numerical examples in Section 6. Finally, we draw our conclusions in Section 7.

2 System Models

2.1 Channel and Signal Models

In a mobile wireless communication system, users encounter both large-scale fading and small-scale fading. As described in [8, Chapter 2], large-scale fading effects reflect the variation of the local average signal, and are well modeled as positive random variables. For simplicity, we quantize the large-scale fading into M levels $h_1 > h_2 > \cdots > h_M > 0$, which naturally captures the salient features of practical systems that employ discrete channel estimates. We assume that all the users experience independently and identically distributed (i.i.d.) large-scale fading with distribution $P\{g = h_m\} = \pi_m$, $m = 1, \cdots, M$, where g represents the large-scale fading.

We consider the uplink of a single-cell symbol-synchronous CDMA system. As in [5, 12, 23], we assume *frequency-flat* fading. Figure 1 depicts a simplified diagram of the wireless link of a CDMA system with L antennas. Because large-scale fading is due to large obstacles (such as terrain and buildings) between transmitter and receiver, we assume that each user's large-scale effects are identical across the antennas. The received signal before filtering at the ℓth antenna can

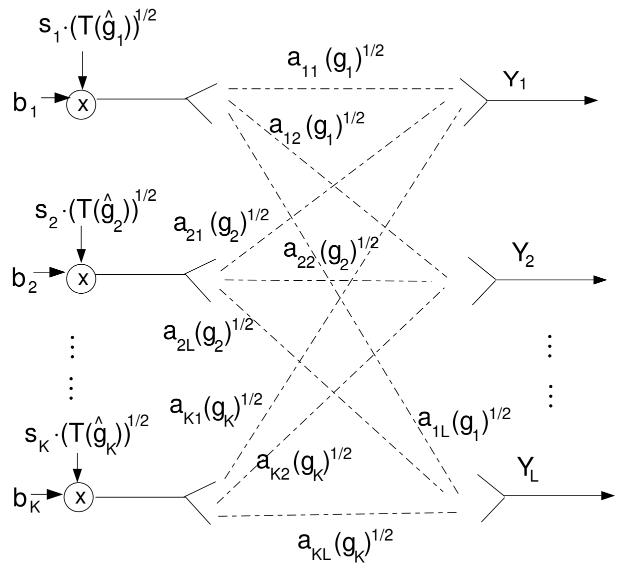

Figure 1: A simplified diagram of the wireless link of a CDMA system with L antennas. The g_k's represent large-scale fading effects, and the $a_{k\ell}$'s represent small-scale fading effects.

be written as

$$Y_\ell = \sum_{k=1}^{K} a_{k\ell}\sqrt{T(\hat{g}_k)g_k}\, b_k s_k + n_\ell, \tag{1}$$

where $a_{k\ell}$ is the kth user's (normalized) small-scale fading in the ℓth diversity channel and has a distribution $\mathbb{C}N(0,1)$,[2] g_k is the kth user's large-scale fading, $\hat{g}_k$ is the estimator of g_k, $T(\hat{g}_k)$ is the kth user's transmission power, b_k is the kth user's transmitted information symbol, s_k is the kth user's spreading signature, and n_ℓ is proper complex white Gaussian noise with positive variance η. In (1), s_k is identical for all L diversity channels because fading is frequency-flat (see also [5, 12, 23]). As noted before, we assume here that a power control mechanism adjusts the transmission power as a function of the estimated large-scale fading (denoted as $T(\hat{g}_k)$). Then, one important question to ask is what is a good way to allot the transmission power. We elaborate on this issue in Section 5.

We assume that the b_k's are independent proper complex random variables with $\mathbb{E}[b_k] = 0$ and $\mathbb{E}[|b_k|^2] = 1$. The model for random signatures is as follows [20]: $s_k = \frac{1}{\sqrt{N}}[s_{k1}, \ldots, s_{kN}]^T$, where the s_{kn}'s are i.i.d. proper complex random variables with zero mean and covariance 1, that is, $\mathbb{E}[|s_{in}|^2] = 1$. We further assume that $\mathbb{E}[|s_{in}|^4] < \infty$.

[2]We use $\mathbb{C}N(0,1)$ to denote a proper complex Gaussian distribution with covariance 1 (see [13] for the definitions of proper complex random variables). Proper complex random variables are also known as *circularly symmetric* complex random variables.

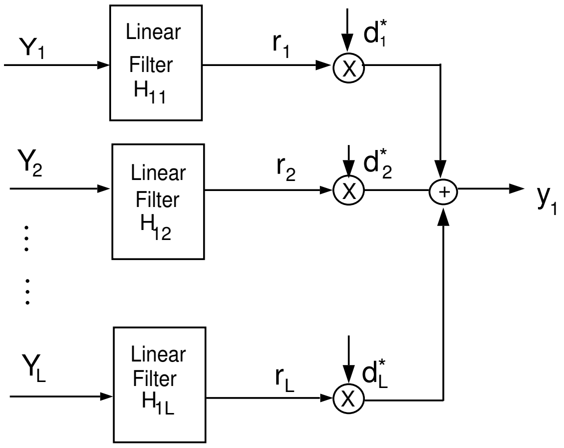

Figure 2: A receiver structure with linear filters followed by a linear combiner.

2.2 Channel Side Information

We assume that linear filtering is applied to the received signal at each antenna, and then the antenna outputs are linearly combined. Figure 2 depicts a typical receiver structure with linear temporal filters followed by a linear spatial combiner. (The d_ℓ's are combiner weighting factors.)

We consider user 1 without loss of generality. We impose the following assumptions on knowledge of the fading effects:

(C1) Transmitter 1 has no information about the interferers' large-scale and small-scale fading effects;

(C2) Receiver 1 has knowledge of the interferers' large-scale fading effects, but no knowledge of the interferers' small-scale fading effects;

(C3) Transmitter 1 has knowledge of the large-scale fading effects of user 1 but not its small-scale fading effects;

(C4) Receiver 1 has knowledge of the large-scale fading effects of user 1. The information of small-scale fading effects available at Receiver 1 depends on a specific system.

The reasoning behind (C1) is that, typically, side information at the transmitter is obtained via a feedback channel, and it is unrealistic to provide much information via feedback. We impose (C2) because it is difficult, if not impossible, for multiuser receivers (e.g., MMSE filters) to adapt to the short-term changes of interference structure; however, it is possible to incorporate the large-scale fading effects since they remain roughly unchanged over a range of thousands of information symbols. As mentioned before, we assume that each user's transmission power is allotted based on the channel side information available at the transmitter. Thus, under the assumption (C3), the kth user's transmission power is determined by $\hat{g}_k$, $k = 1, \cdots, K$. More specifically, when a user experiences large-scale fading g with $\hat{g}$ equal to h_m, its transmission power is allotted as $T_m = T(h_m)$, where T is some fixed mapping. We assume that the $\hat{g}_k$'s are i.i.d., which implies that the $T(\hat{g}_k)$'s are i.i.d. as well.

In a practical communication system, due to time-varying channel conditions, feedback errors and delays are often inevitable, resulting noisy side information about large-scale fading at the transmitter. For future convenience, define for $m, n \in \{1, \cdots, M\}$,

$$P_e^{(m,n)} \triangleq P\{\hat{g} = h_m | g = h_n\}, \tag{2}$$

that is, $P_e^{(m,n)}$ represents the error probability when the actual large-scale fading is h_n but the estimate indicates that the large-scale fading is h_m with $m \neq n$. We further note that in (C4), different assumptions on the side information of small-scale fading effects lead to different diversity combining methods.

2.3 Linear MMSE Filter, Linear Combiner, and Concatenation

We focus on the receiver that consists of a bank of temporal MMSE filters and a linear spatial combiner. More specifically, a linear MMSE filter is applied to the received signal at each antenna. We note that conditions (C2) and (C4) indicates that the signatures and large-scale fading effects of the interferers, but not the small-scale fading effects of the interferers, are available at the MMSE filters. The MMSE filters outputs are then combined via a spatial combiner. In particular, we are interested in the following two cases: In the first case, the receiver has knowledge of the small-scale fading effects of the desired user, and combine the MMSE filter outputs to maximize the SIR; in contrast, in the second case, the receiver has no knowledge of this, and does selection combining.

We note that if the receiver had knowledge of small-scale fading of all the users in all the L diversity channels, the corresponding linear MMSE receiver would be different from the one considered here. However, observe that the large dimension (NL) of the signals may incur possibly very high computational complexity, making the linear (end-to-end) MMSE receiver difficult to implement. Thus, we confine ourselves to the receiver proposed above, which is more practically appealing.

A linear MMSE filter is applied to de-spread the received signal at each antenna. The MMSE filter at antenna l is the one minimizing

$$\mathbb{E}\left[\left| c_\ell^H Y_\ell - b_1 \right|^2 \middle| a_{1\ell}, s_k, T(\hat{g}_k) g_k,\ k = 1, \cdots, K \right].$$

(It turns out that in a large system, there is no need to know $a_{1\ell}$ for the construction of the MMSE filter [25].) It can be shown that the linear MMSE filters for all the diversity channels share the following common form [25]: $c = \text{scalar} \cdot M_I^{-1} s_1$, where $M_I \triangleq \sum_{k=2}^K T(\hat{g}_k) g_k s_k s_k^H + \eta I$. As pointed out in [25] (also indicated by Theorems 3.1 and 4.1 below), the SIR is the key parameter that governs the performance in a large system. Since any (positive) scaled version of linear MMSE filters results in the same SIR, we process the received signals at all the antennas with $M_I^{-1} s_1$ (which is a scaled version of the MMSE filter). Then the output of the linear filter in the ℓth diversity channel is

$$r_\ell = \sum_{k=1}^{K} a_{k\ell}\sqrt{T(\hat{g}_k)g_k}b_k s_1^H M_I^{-1} s_k + s_1^H M_I^{-1} n_\ell.$$

For convenience, define

$$\boldsymbol{a} \triangleq [a_{11}, a_{12}, \cdots, a_{1L}]^t, \qquad \boldsymbol{d} \triangleq [d_1, d_2, \cdots, d_L]^t,$$

and for $l = 1, \cdots, L$,

$$\begin{aligned}
\mathcal{I}_{\mathrm{MAI},\ell}^{(N)} &\triangleq \sum_{k=2}^{K} a_{k\ell}\sqrt{T(\hat{g}_k)g_k}b_k s_1^H M_I^{-1} s_k, \\
\mathcal{I}_{n,\ell}^{(N)} &\triangleq s_1^H M_I^{-1} n_\ell, \\
\mathcal{I}_{\ell}^{(N)} &\triangleq \mathcal{I}_{\mathrm{MAI},\ell}^{(N)} + \mathcal{I}_{n,\ell}^{(N)}.
\end{aligned}$$

The output of any linear combiner has the following form:

$$\begin{aligned}
y_1 &= \sum_{\ell=1}^{L} d_\ell^* r_\ell \\
&= \sum_{\ell=1}^{L} d_\ell^* a_{1\ell}\sqrt{T(\hat{g}_1)g_1}b_1 s_1^H M_I^{-1} s_1 + \mathcal{I}^{(N)},
\end{aligned} \tag{3}$$

where $\mathcal{I}^{(N)} = \mathcal{I}_{\mathrm{MAI}}^{(N)} + \mathcal{I}_n^{(N)}$ with $\mathcal{I}_{\mathrm{MAI}}^{(N)} = \sum_{\ell=1}^{L} d_\ell^* \mathcal{I}_{\mathrm{MAI},\ell}^{(N)}$ and $\mathcal{I}_n^{(N)} = \sum_{\ell=1}^{L} d_\ell^* \mathcal{I}_{n,\ell}^{(N)}$.

It is easy to show that the desired signal power is

$$U_1 = T(\hat{g}_1)g_1 |\langle \boldsymbol{d}, \boldsymbol{a}\rangle|^2 \left|s_1^H M_I^{-1} s_1\right|^2, \tag{4}$$

where $\langle \cdot, \cdot \rangle$ denotes the standard inner product in $\mathbb{C}^L$. According to the assumptions (C2) and (CC4), receiver 1 has knowledge of all the users' large-scale fading effects. We further assume that receiver 1 has knowledge of the s_k's and $T(\hat{g}_k)g_k$'s (cf. [20, 25]). Let $\mathbb{E}_I$ denote the conditional expectation given the above information and the interferers' small-scale fading effects. Then the instantaneous interference power in a symbol interval is

$$\begin{aligned}
\mathbb{E}_I\left[|\mathcal{I}^{(N)}|^2\right] = \mathbb{E}_I \Bigg\{ \sum_{\ell_1,\ell_2} &\left(\sum_{k_1=2}^{K} d_{\ell_1}^* a_{k_1\ell_1}\sqrt{T(\hat{g}_{k_1})g_{k_1}}b_{k_1} s_1^H M_{I_1}^{-1} s_{k_1} + d_{\ell_1}^* s_1^H M_{I_1}^{-1} n_{\ell_1} \right) \\
&\cdot \left(\sum_{k_2=2}^{K} d_{\ell_2} a_{k_2\ell_2}^* \sqrt{T(\hat{g}_{k_2})g_{k_2}} b_{k_2}^* s_{k_2}^H M_{I_2}^{-1} s_1 + d_{\ell_2} n_{\ell_2}^H M_{I_2}^{-1} s_1 \right) \Bigg\},
\end{aligned} \tag{5}$$

where $M_{I_1} \triangleq \sum_{k_1=2}^{K} T(\hat{g}_{k_1})g_{k_1} s_{k_1} s_{k_1}^H + \eta I$ and $M_{I_2} \triangleq \sum_{k_2=2}^{K} T(\hat{g}_{k_2})g_{k_2} s_{k_2} s_{k_2}^H + \eta I$. Combining (4) and (5), the SIR can be expressed as follows:

$$\mathrm{SIR}_1^{(N)} = T(\hat{g}_1)g_1 \frac{|\langle \boldsymbol{d}, \boldsymbol{a}\rangle|^2 \left|s_1^H M_I^{-1} s_1\right|^2}{\mathbb{E}_I\left[|\mathcal{I}^{(N)}|^2\right]}. \tag{6}$$

3 The Case of Independent Small-Scale Fading

In this section, we assume that each user's small-scale fading effects are independent across different antennas. Let F_μ denote the discrete probability distribution with

$$P\{\mu = T_m h_n\} = \pi_n P_e^{(n,m)}, \qquad m, n \in \{1, \cdots, M\}.$$

We need the following lemmas to establish our main results (the proofs of these lemmas have been relegated to the Appendix).

Lemma 3.1 *The empirical distribution of* $\{T(\hat{g}_1)g_1, \cdots, T(\hat{g}_K)g_K\}$ *converges weakly (as* $K \to \infty$*) to* F_μ *with probability one.*

A simple application of the Glivenko-Cantelli Theorem proves the above lemma.

Based on Lemma 3.1, using [19, Theorem 1.1] it can be shown that $s_1^H M_I^{-1} s_1$ converges in probability to β_0, where β_0 is the unique positive solution to the following fixed point equation [20, 25]:

$$\beta_0 = \frac{1}{\eta + \alpha \int_0^\infty \frac{\mu}{1+\mu\beta_0} dF_\mu(\mu)}. \tag{7}$$

Moreover, $s_1^H M_I^{-2} s_1$ converges in probability to $-\frac{\partial \beta_0}{\partial \eta}$.

Lemma 3.2 *(i)*

$$\mathbb{E}\left[\left|s_1^H M_I^{-1} s_k\right|^4\right] \leq \frac{C_1}{N^2 \eta^4},$$

where C_1 *is a constant independent of* N.

(ii) Suppose $a_{k\ell_1}$ *and* $a_{k\ell_2}$ *are independent for any* $\ell_1 \neq \ell_2$, $\ell_1, \ell_2 \in \{1, \cdots, L\}$*. Then for* $\ell_1 \neq \ell_2$,

$$s_1^H M_I^{-1} \left[\sum_{k=2}^{K} a_{k\ell_1} a_{k\ell_2}^* T(\hat{g}_k) g_k s_k s_k^H\right] M_I^{-1} s_1 \xrightarrow{P} 0.$$

(iii) For $\ell = 1, \cdots, L$*, we have that*

$$s_1^H M_I^{-1} \left[\sum_{k=2}^{K} \left(|a_{k\ell}|^2 - 1\right) T(\hat{g}_k) g_k s_k s_k^H\right] M_I^{-1} s_1 \xrightarrow{P} 0.$$

We are now ready to present our first main result.

Theorem 3.1 *Suppose that each user's small-scale fading effects across different antennas are independent. Then as* $N \to \infty$,

(a) The interferences across the antennas, $\mathbb{I}^{(N)} \triangleq \left[\mathcal{I}_1^{(N)}, \mathcal{I}_2^{(N)}, \cdots, \mathcal{I}_L^{(N)}\right]^t$*, have a limiting proper complex Gaussian distribution* $\mathbb{C}N(\mathbf{0}, \beta_0 \boldsymbol{I})$;

(b)

$$\mathrm{SIR}_1^{(N)} \xrightarrow{P} T(\hat{g}_1) g_1 \beta_0 \frac{|\langle \boldsymbol{d}, \boldsymbol{a} \rangle|^2}{|\boldsymbol{d}|^2}, \tag{8}$$

where β_0 *is the unique positive solution to (7).*

We note that the convergence in part (a) is in the weak sense, that is, the joint distribution of the interferences at the output of the antennas converges weakly to a proper complex Gaussian measure $\mathbb{C}N(\mathbf{0}, \beta_0 \boldsymbol{I})$. More importantly, the Gaussianity of the interferences reveals that from the viewpoints of detection and channel capacity, the SIR is of fundamental interest. We provide the proof of Theorem 3.1 in the following.

Proof: The proof of part (a) makes use of the Cramér-Wold Theorem [2, Theorem 29.4] and the dependent central limit theorem in [10]. Let $\boldsymbol{Z} = [z_1, z_2, \cdots, z_L]^t$ denote a proper Gaussian random vector with mean $\mathbf{0}$ and covariance matrix $\boldsymbol{I}$. By the Cramér-Wold Theorem, it suffices to show for any $\boldsymbol{d} = [d_1, d_2, \cdots, d_L]^t$ in $\mathbb{C}^L$, $\sum_{\ell=1}^{L} d_\ell^* \mathcal{I}_\ell^{(N)}$ converges in distribution to $\sum_{\ell=1}^{L} d_\ell^* z_\ell$. To this end, define

$$t_k^{(N)} \triangleq \sum_{\ell=1}^{L} d_\ell^* a_{k\ell} \sqrt{T(\hat{g}_k) g_k} b_k s_1^t M_I^{-1} s_k, \qquad k = 2, \cdots, K.$$

Then $\mathcal{I}_{\text{MAI}}^{(N)} = \sum_{k=2}^{K} t_k^{(N)}$, and $\sum_{\ell=1}^{L} d_\ell^* \mathcal{I}_\ell^{(N)} = \mathcal{I}_{\text{MAI}}^{(N)} + \mathcal{I}_n^{(N)}$. It is straightforward to see that $\mathbb{I}^{(N)}$ is proper. In what follows, we first show that $\mathcal{I}_{\text{MAI}}^{(N)}$ has a limiting proper complex Gaussian distribution. Define $\mathcal{F}_{N,k}$ as the σ-algebra generated by $\{t_2^{(N)}, \cdots, t_k^{(N)}\}$, that is,

$$\mathcal{F}_{N,k} \triangleq \sigma\left(t_2^{(N)}, \cdots, t_k^{(N)}\right), \qquad k = 2, \cdots, K.$$

It is clear that the triangular array $\{t_k^{(N)}\}$ is a martingale difference array with respect to $\{\mathcal{F}_{N,k}\}$. Thus, based on [10], it suffices to verify that the following three conditions are satisfied:

1. $\max_{2 \le k \le K} |t_k^{(N)}|$ is bounded in L_2 norm,

2. $\max_{2 \le k \le K} |t_k^{(N)}|$ converges to 0 in probability as $N \to \infty$,

3. $\sum_{k=2}^{K} |t_k^{(N)}|^2 - \sum_l |d_\ell|^2 s_1^H (M_I^{-1} - \eta M_I^{-2}) s_1$ converges to 0 in probability as $N \to \infty$.

Although more complicated, the proof of the above three conditions essentially follows the same line as that of [25, Theorem 3.1]. We omit the details here.

It is easy to show that $\mathcal{I}_n^{(N)}$ has a limiting proper complex Gaussian distribution, and that $\mathcal{I}_{\text{MAI}}^{(N)}$ and $\mathcal{I}_n^{(N)}$ are uncorrelated. Thus, $\mathcal{I}^{(N)} = \sum_{\ell=1}^{L} d_\ell^* \mathcal{I}_\ell^{(N)} = \mathcal{I}_{\text{MAI}}^{(N)} + \mathcal{I}_n^{(N)}$ has a limiting proper complex Gaussian distribution. Then, it remains to calculate the variance of $\mathcal{I}^{(N)}$. Since the $a_{k\ell}$'s are independent, we have that

$$\mathbb{E}_I\left[|\mathcal{I}^{(N)}|^2\right] = \mathbb{E}_I\left\{\sum_{\ell_1,\ell_2} \sum_{k=2}^{K} d_{\ell_1}^* d_{\ell_2} a_{k\ell_1} a_{k\ell_2}^* T(\hat{g}_k) g_k |b_k|^2 \left|s_1^H M_I^{-1} s_k\right|^2 + \sum_l \left|d_\ell s_1^H M_I^{-1} n_\ell\right|^2\right\}.$$

Using Lemma 3.2, it can be shown that

$$\mathbb{E}_I\left[|\mathcal{I}^{(N)}|^2\right] - \sum_l |d_\ell|^2 \left(\sum_k T(\hat{g}_k) g_k \left|s_1^H M_I^{-1} s_k\right|^2 + \eta s_1^H M_I^{-2} s_1\right) \xrightarrow{P} 0,$$

which yields

$$\mathbb{E}_I\left[|\mathcal{I}^{(N)}|^2\right] - \sum_l |d_\ell|^2 s_1^H M_I^{-1} s_1 \xrightarrow{P} 0. \tag{9}$$

It then follows that

$$\mathbb{E}_I\left[|\mathcal{I}^{(N)}|^2\right] \xrightarrow{P} \sum_l |d_\ell|^2 \beta_0,$$

thus concluding the proof of part (a).

The proof of part (b) follows by Combining (6), (7), and (9). ∎

We observe from part (b) of Theorem 3.1 that the limiting SIR expression can be factorized into two components: $T(\hat{g}_1)g_1\beta_0$ and $|\langle \boldsymbol{d}, \boldsymbol{a}\rangle|^2/|\boldsymbol{d}|^2$. Clearly, $T(\hat{g}_1)g_1\beta_0$ is a function of only the large-scale fading effects and hence depends only on transmission power control; in contrast, $|\langle \boldsymbol{d}, \boldsymbol{a}\rangle|^2/|\boldsymbol{d}|^2$ is is a weighted sum of the small-scale fading effects of user 1 and hence depends only on diversity combining schemes. Thus, there is a clear *separation* between the gains of transmission power control and diversity combining, and more importantly, the two gains are *additive* (in decibels). Because of this separation, diversity combining behaves the same here as in a single-user system with antenna arrays, and its gain depends on the side information of the desired user's small-scale fading effects.

Heuristically, the above result tells us that in a large system, the SIR is approximately $T(\hat{g}_1)g_1\beta_0|\langle \boldsymbol{d}, \boldsymbol{a}\rangle|^2/|\boldsymbol{d}|^2$. Henceforth we use the above limiting expression to analyze two diversity combining schemes—MRC and SC.

3.1 Maximum Ratio Combining

Part (a) in Theorem 3.1 establishes that the joint distribution of the overall interferences at the output of the antennas converges weakly to a proper complex Gaussian measure $\mathbb{C}N(\boldsymbol{0}, \beta_0\boldsymbol{I})$, which implies that the traditional MRC method is the best linear combiner (see, e.g., [14, Chapter 8]). MRC requires information about the $a_{1\ell}$'s [8]. Given the $a_{1\ell}$'s, the principle of MRC is to weight the outputs of the linear filters appropriately and sum them up to maximize the SIR. A simple application of the Cauchy-Schwartz Inequality shows that the choice of $\boldsymbol{d}$ that maximizes the SIR is $d_\ell = e_0 a_{1\ell}$, $l = 1, \cdots, L$, where $e_0 \in \mathbb{C}$ is some nonzero constant [8, Chap. 5]. The SIR is then of the following form,

$$\Gamma_{\mathrm{mrc}} = T(\hat{g}_1)g_1\beta_0 X, \tag{10}$$

where X denotes $\sum_{\ell=1}^L |a_{1\ell}|^2$. It is straightforward to see that $X \sim \mathrm{Gamma}(L, 1)$.

We note that the MRC choice $\boldsymbol{d}$ (under the right choice of e_0) also minimizes the mean-square error

$$\mathbb{E}\left[\left|\sum_{\ell=1}^L d_\ell^* r_\ell - b_1\right|^2 \middle| \boldsymbol{a}\right].$$

Indeed, for a given $\boldsymbol{a}$, there exists a functional relationship between the MMSE (over all the antennas) and the SIR (denoted as MSIR) attained by the MRC method [9]:

$$\mathrm{MSIR} = \frac{1}{\mathrm{MMSE}} - 1. \tag{11}$$

Diversity combining is used to improve the mean SIR and reduce the dynamic range of the signal strength [8, 14], which indicates that the mean and variance of the SIR are important performance measures. Since large-scale fading effects may remain constant over thousands of information symbols, it is of more interest to study the "local" mean and variance of the SIR. (Roughly speaking, "local" here refers to the duration in which the large-scale fading effects remain constant.) The "local" mean and variance of the SIR corresponds to finding the conditional mean and variance of Γ_{mrc} given $T(\hat{g}_1)g_1$. Using (10) we have that

$$\mathbb{E}[\Gamma_{\text{mrc}}|T(\hat{g}_1)g_1] = LT(\hat{g}_1)g_1\beta_0, \qquad \text{var}(\Gamma_{\text{mrc}}|T(\hat{g}_1)g_1) = L(T(\hat{g}_1)g_1\beta_0)^2. \tag{12}$$

3.2 Selection Combining

For the system output, the ideal selection diversity combiner chooses, during each "instant", the signal from the filter that has the largest SIR [8, 18]. In practice the antenna signals could be sampled, for example, and the best one is sent to the decoder.

Define $a_{(L)} \triangleq \max_\ell \left(|a_{1\ell}|^2\right)$. It can easily be shown that the SIR achieved by SC (denoted as Γ_{sc}) is simply $\Gamma_{\text{sc}} = T(\hat{g}_1)g_1\beta_0 a_{(L)}$.

As in the MRC method, we are interested in the "local" mean and variance of Γ_{sc}. The calculation boils down to computing the mean and variance of $a_{(L)}$. The following result is useful in the calculation [1, Chapter 3]:

$$a_{(L)} = \sum_{j=0}^{L-1} \frac{D_j}{L-j}, \tag{13}$$

where the D_j's are exponentially distributed independent random variables with mean 1.[3] Using (13), we can readily calculate the mean and variance of $a_{(L)}$ (cf. [8, Chapter 5]):

$$\mathbb{E}[a_{(L)}] = \sum_{j=1}^{L} \frac{1}{j}, \qquad \text{var}(a_{(L)}) = \sum_{j=1}^{L} \frac{1}{j^2}.$$

Then it is straightforward to see that

$$\mathbb{E}[\Gamma_{\text{sc}}|T(\hat{g}_1)g_1] = T(\hat{g}_1)g_1\beta_0 \sum_{j=1}^{L} \frac{1}{j}, \qquad \text{var}(\Gamma_{\text{sc}}|T(\hat{g}_1)g_1) = (T(\hat{g}_1)g_1\beta_0)^2 \sum_{j=1}^{L} \frac{1}{j^2}. \tag{14}$$

Comparing (12) and (14), we conclude that the mean of the SIR in the MRC method (Γ_{mrc}) is larger than that (Γ_{sc}) in the SC method. Moreover, as L increases, Γ_{mrc} grows linearly in L while Γ_{sc} grows approximately at the rate of $\log L$. From a theoretical viewpoint, given a realization of $\boldsymbol{a}$, MRC and SC are essentially equivalent to taking $(\|\boldsymbol{a}\|_2)^2$ and $(\|\boldsymbol{a}\|_1)^2$, respectively. It is clear that $\|\boldsymbol{a}\|_2$ is always greater than or equal to $\|\boldsymbol{a}\|_1$, which explains why Γ_{mrc} is always larger than Γ_{sc}. However, we note that the variance of the SIR in the SC method is smaller than that in the MRC method.

[3] The D_j's are called normalized spacings of the order statistics [1, Chapter 3].

4 The Case of Correlated Small-Scale Fading

In the preceding sections, the analysis has been premised upon the assumption that each user's small-scale fading effects at various antennas are independent. There will be cases where this is difficult to achieve, for example, because of insufficient antenna spacing. In what follows, we consider a more realistic scenario where each user's small-scale fading effects are correlated across the antennas at the base station. In such scenario, it is natural to ask whether it is possible to achieve the same diversity gain as in the independent case, and if so when. The answer is yes, and indeed when the noise level diminishes to zero, using MRC leads to the same diversity gain as in the independent case. However, to achieve the same diversity gain, the correlation among the small-scale fading effects does incur extra processing complexity. To be more specific, we begin with the following lemma. We assume that $\mathbb{E}[a_{k\ell_1}^* a_{k\ell_2}]$ does not depend on k and use R_{ℓ_1,ℓ_2} to denote $\mathbb{E}[a_{k\ell_1}^* a_{k\ell_2}]$.

Lemma 4.1 *Suppose that* $\mathbb{E}\left[|a_{k\ell_1}|^2|a_{k\ell_2}|^2\right] < \infty$ *where* $\ell_1, \ell_2 \in \{1, \cdots, L\}$. *Then for any* $\ell_1 \neq \ell_2$,

$$\sum_{k=2}^{K} s_1^H M_I^{-1} a_{k\ell_1}^* a_{k\ell_2} T(\hat{g}_k) g_k s_k s_k^H M_I^{-1} s_1 - R_{\ell_1,\ell_2} \sum_{k=2}^{K} s_1^H M_I^{-1} T(\hat{g}_k) g_k s_k s_k^H M_I^{-1} s_1 \xrightarrow{P} 0.$$

The proof of Lemma 4.1 follows the same line as that of part (b) in Lemma 3.2.

Define the complex covariance for $\boldsymbol{a}$ by

$$\boldsymbol{R} \triangleq \frac{1}{2} \begin{bmatrix} 1 & R_{12} & \cdots & R_{1L} \\ R_{21} & 1 & \cdots & R_{2L} \\ \vdots & & \ddots & \vdots \\ R_{L1} & \cdots & \cdots & 1 \end{bmatrix}.$$

Using Lemma 4.1, it can easily be shown that

$$\mathbb{E}_I\left[|\mathcal{I}^{(N)}|^2\right] - 2\boldsymbol{d}^H \boldsymbol{R} \boldsymbol{d} \left(s_1^H M_I^{-1} s_1 - \eta s_1^H M_I^{-2} s_1\right) - \eta \boldsymbol{d}^H \boldsymbol{d} s_1^H M_I^{-2} s_1 \xrightarrow{P} 0.$$

Combining the above with (6), we obtain that

$$\mathrm{SIR}_1^{(N)} \xrightarrow{P} \frac{T(\hat{g}_1) g_1 (\beta_0)^2 |\langle \boldsymbol{d}, \boldsymbol{a} \rangle|^2}{2\boldsymbol{d}^H \boldsymbol{R} \boldsymbol{d} \left(\beta_0 + \eta \frac{\partial \beta_0}{\partial \eta}\right) - \boldsymbol{d}^H \boldsymbol{d} \eta \frac{\partial \beta_0}{\partial \eta}}. \tag{15}$$

For convenience, define

$$\boldsymbol{W} \triangleq \frac{1}{\beta_0}\left[2\left(\beta_0 + \eta \frac{\partial \beta_0}{\partial \eta}\right) \boldsymbol{R} - \eta \frac{\partial \beta_0}{\partial \eta} \boldsymbol{I}\right].$$

(Note that $\boldsymbol{W}$ is positive definite.) We now have the following result.

Theorem 4.1 *Suppose that each user's small-scale fading effects are correlated across the antennas, with covariance matrix* $\boldsymbol{R}$. *Then as* $N \to \infty$, *(a) The interferences across the antennas,* $\mathbb{I}^{(N)} = \left[\mathcal{I}_1^{(N)}, \mathcal{I}_2^{(N)}, \cdots, \mathcal{I}_L^{(N)}\right]^t$*, have a limiting proper complex Gaussian distribution* $\mathbb{C}N(\mathbf{0}, \beta_0 \boldsymbol{W})$*;*

(b)

$$\mathrm{SIR}_1^{(N)} \xrightarrow{P} T(\hat{g}_1) g_1 \beta_0 \frac{|\langle \boldsymbol{d}, \boldsymbol{a}\rangle|^2}{\boldsymbol{d}^H \boldsymbol{W} \boldsymbol{d}}. \tag{16}$$

The proof of Theorem 4.1 follows essentially the same line as that of Theorem 3.1.

Based on the limiting SIR expression given in (16), it is interesting to note that in the case where small-scale fading effects are correlated across the antennas, the separation between the gains of power control and diversity combining no longer exists. This is because the gain of diversity combining depends on $\boldsymbol{W}$, which is a function of β_0, and β_0 is determined by transmission power control. Therefore, in general, the gains of transmission power control and diversity combining are *coupled.*

4.1 Maximum Ratio Combining

Part (a) in Theorem 4.1 establishes that the joint distribution of the overall interference at the output of the antennas converges weakly to a proper complex Gaussian measure $\mathbb{C}N(\mathbf{0}, \beta_0 \boldsymbol{W})$. When each user's small-scale fading effects are correlated across the antennas, in general $\boldsymbol{W}$ is not diagonal, which implies that the traditional MRC method is no longer the best linear combiner. Instead, the MRC principle leads to a *new* combiner in the correlated small-scale fading case. It is worth noting that (11) still holds, and the MRC approach and the MMSE criterion still lead to the same optimal linear combiner.

By definition, MRC chooses $\boldsymbol{d}$ to maximize the SIR. Observe that the maximization of the SIR boils down to the following optimization problem:

$$\max_{\boldsymbol{d} \in \mathbb{C}^L} \frac{\boldsymbol{d}^H \boldsymbol{a}\boldsymbol{a}^H \boldsymbol{d}}{\boldsymbol{d}^H \boldsymbol{W} \boldsymbol{d}}. \tag{17}$$

Define $\widetilde{\boldsymbol{d}} = \boldsymbol{W}^{1/2}\boldsymbol{d}$. Then the optimization problem in (17) can be written as

$$\max_{\widetilde{\boldsymbol{d}} \in \mathbb{C}^L} \frac{\widetilde{\boldsymbol{d}}^H \boldsymbol{W}^{-1/2} \boldsymbol{a}\boldsymbol{a}^H \boldsymbol{W}^{-1/2} \widetilde{\boldsymbol{d}}}{\widetilde{\boldsymbol{d}}^H \widetilde{\boldsymbol{d}}}. \tag{18}$$

Appealing to [7, p. 176], it follows that the maximum of the objective function is the largest eigenvalue λ_1 of $\boldsymbol{W}^{-1/2}\boldsymbol{a}\boldsymbol{a}^H\boldsymbol{W}^{-1/2}$, and is achieved when $\boldsymbol{d}$ is the eigenvector corresponding to λ_1. Because the matrix $\boldsymbol{W}^{-1/2}\boldsymbol{a}\boldsymbol{a}^H\boldsymbol{W}^{-1/2}$ is of rank 1 (except when $\boldsymbol{a} = \mathbf{0}$, which happens with probability zero), we have that

$$\lambda_1 = \mathrm{trace}\left(\boldsymbol{W}^{-1/2}\boldsymbol{a}\boldsymbol{a}^H\boldsymbol{W}^{-1/2}\right) = \boldsymbol{a}^H \boldsymbol{W}^{-1} \boldsymbol{a},$$

which bears a Hermitian quadratic form in complex random variables. To characterize the distribution of λ_1, we follow [18] and further impose that $\boldsymbol{a}$ is a proper complex random vector. Then by [13, Theorem 1], the probability density function of $\boldsymbol{a}$ is given by[4]

$$f(\boldsymbol{a}) = \frac{1}{\pi^L \det(2\boldsymbol{R})} \exp\left\{-\boldsymbol{a}^H (2\boldsymbol{R})^{-1}\boldsymbol{a}\right\}.$$

The moment generating function of λ_1 is $G_{\lambda_1}(t) = \mathbb{E}[\exp(t\lambda_1)]$, where the expectation is taken with respect to $f(\boldsymbol{a})$. (Because λ_1 is complex Gaussian, $G_{\lambda_1}(t)$ exists.) Appealing to [18, Appendex B], $G_{\lambda_1}(t)$ can be shown to be

$$G_{\lambda_1}(t) = \frac{1}{\det\left(\boldsymbol{I} - 2t\boldsymbol{R}^H \left(\boldsymbol{W}\right)^{-1}\right)}.$$

After some algebra, we get that

$$G_{\lambda_1}(t) = \frac{1}{\prod_{\ell=1}^{L}\left(1 - t\frac{\beta_0}{\beta_0 + \eta\frac{\partial\beta_0}{\partial\eta}(1-\frac{1}{\nu_\ell})}\right)}, \tag{19}$$

where the ν_ℓ's are eigenvalues of the matrix $2\boldsymbol{R}$.

Based on (19), we can use the inverse Laplace transform to obtain the probability density function of λ_1. In particular, we are interested in the case where the noise level diminishes to zero ($\eta = 0$). In this high SNR region, the moment generating function of λ_1 can further be simplified to

$$G_{\lambda_1}(t) = (1-t)^{-L}. \tag{20}$$

It then follows that that λ_1 has a distribution Gamma$(L, 1)$. Combining (20) and Theorem 4.1 leads to the following result.

Theorem 4.2 *(Correlated fading and MRC: SIR) When the noise level diminishes to zero ($\eta = 0$), as $N \to \infty$,*

$$\mathrm{SIR}_{1,mrc}^{(N)} \xrightarrow{P} T(\hat{g}_1) g_1 \beta_0 X, \tag{21}$$

where $X \sim$ Gamma$(L, 1)$. Moreover, the optimal combining vectors share the common form $e_1\boldsymbol{R}^{-1}\boldsymbol{a}$, where $e_1 \in \mathbb{C}$ is some nonzero constant.

The physical implications of this result are as follows. In the high SNR region ($\eta = 0$), the MRC approach used in the presence of correlated small-scale fading can still achieve the same diversity gain as in the independent case. That is to say, even when the small-scale fading is correlated across the antennas, the MRC approach is as effective as in the independent case. Moreover, using MRC leads to the "decoupling" of the gains of transmission power control and diversity combining. The underlying reasoning behind Theorem 4.2 is that CDMA systems are interference-limited and correlated interference may be whitened. Indeed, when $\eta = 0$, the optimization problem in (18) reduces to

$$\max_{\tilde{\boldsymbol{d}} \in \mathbb{C}^L} \frac{\tilde{\boldsymbol{d}}^H \boldsymbol{R}^{-1/2}\boldsymbol{a}\boldsymbol{a}^H\boldsymbol{R}^{-1/2}\tilde{\boldsymbol{d}}}{\tilde{\boldsymbol{d}}^H\tilde{\boldsymbol{d}}}.$$

[4] Without loss of generality, we assume henceforth that no deterministic linear relationship exists among any of the $a_{k\ell}$'s, and hence $\boldsymbol{R}$ is Hermitian positive definite [18].

Then the solution to the above problem is given by $\tilde{\boldsymbol{d}} = e_1 \boldsymbol{R}^{-1/2}\boldsymbol{a}$, where $e_1 \in \mathbb{C}$ is some nonzero constant. Thus we have that $\boldsymbol{d} = e_1 \boldsymbol{R}^{-1}\boldsymbol{a}$. Intuitively, we can first use $\boldsymbol{R}^{-1}$ to *whiten* the correlated interference, and then apply the traditional MRC method as in the independent case.

As pointed out in [8, Chapter 5] [18, Chapter 10], when selection combining is adopted, it does not appear that one can handle analytically more than two antennas $L \geq 2$. We also note that in the presence of background noise, closed-form solutions seem hard to attain for either MRC or SC. However, based on our preceding analysis, one can expect that when the users' SNR is reasonably high, a large portion of the diversity effectiveness will be retained even when significant correlations exist, as in a single-user system [8, 18].

5 Transmission Power Allocation

In this section, we study the allocation of transmission power based on the side information about large-scale fading. Our desired allocation is to make transmission power consumption as low as possible while keeping all the users' QoS requirements satisfied. Recall that from the viewpoints of detection and channel capacity, the SIR is the key performance measure for QoS requirements. Because of the randomness of the received powers and signatures, the SIR is random as well. Also observe that large-sale fading may remain constant over thousands of information symbols, which possibly span dozens of frames. Therefore, we adopt the following probabilistic model for the users' QoS requirements ([11, 24]):

$$P\{\mathrm{SIR}_k \geq \gamma | \hat{g}_k = h_m\} \geq p_a, \qquad m = 1, \cdots, M, \tag{22}$$

where γ is the *target SIR*, and $p_a \in [0, 1)$, $k = 1, \cdots, K$.

In what follows we calculate $P\{\mathrm{SIR}_k \geq \gamma | \hat{g}_k = h_m\}$. Because closed-form solutions seem unattainable when each user's small-scale fading effects are correlated across antennas and $\eta > 0$, and also when $\eta = 0$ using MRC leads to the same SIR as in the independent case, we focus on the independent case in the following. We treat systems with MRC and SC separately. To facilitate the notation, in the following we use g to denote the large-scale fading of user 1 and $\hat{g}$ its estimate.

Without loss of generality, suppose $\hat{g} = h_m$, that is, the channel estimate of the large-scale fading indicates that the desired user is in state m, and hence the transmission power is T_m. In the MRC method, the SIR is $T_m g \beta_0 X$, where $X \sim \mathrm{Gamma}(L, 1)$. Therefore, we have that

$$\begin{aligned} P_m &= P\left\{ T_m g \beta_0 X \geq \gamma | \, \hat{g} = h_m \right\} \\ &= P\left\{ X \geq \frac{\gamma}{T_m g \beta_0} \middle| \, \hat{g} = h_m \right\} \\ &= P\left\{ \frac{1}{2}\mathcal{X}^2_{2L} \geq \frac{\gamma}{T_m g \beta_0} \middle| \, \hat{g} = h_m \right\}, \end{aligned} \tag{23}$$

where $\mathcal{X}^2_{2L}$ is a Chi-squared random variable with $2L$ degrees of freedom. On the other hand, in the SC method, the SIR is $T_m g \beta_0 a_{(L)}$. It then follows that

$$P_m = P\left\{ T_m g \beta_0 a_{(L)} \geq \gamma | \, \hat{g} = h_m \right\}$$

$$
\begin{aligned}
&= P\left\{a_{(L)} \geq \frac{\gamma}{T_m g \beta_0} \middle| \hat{g} = h_m\right\} \\
&= 1 - \left(1 - \exp(-\frac{\gamma}{T_m g \beta_0})\right)^L. \qquad (24)
\end{aligned}
$$

We study two cases, one with perfect side information, the other with noisy side information, although the latter is more general than the former. The main reason for presenting them separately is that the case with perfect side information provides a good contrast to the case with noisy side information and helps build up the results.

5.1 Perfect Side Information of Large-Scale Fading

In this subsection, we assume that the estimation of large-scale fading is perfect, that is, $\hat{g} = g$ always.

5.1.1 Maximum Ratio Combining

Based on (23), to keep all the users' SIR requirements satisfied it suffices to have

$$
T_m h_m \beta_0 \geq \frac{2\gamma}{\mathcal{X}^2_{2L,p_a}}, \qquad m = 1, \cdots, M, \qquad (25)
$$

where $\mathcal{X}^2_{2L,p_a}$ is the cutoff point for $\mathcal{X}^2_{2L}$ with right-hand tail probability p_a.

Using (7), we have that for $m = 1, \cdots, M$,

$$
T_m h_m \beta_0 = \frac{T_m h_m}{\eta + \alpha \sum_{j=1}^{M} \pi_j \frac{T_j h_j}{1 + T_j h_j \beta_0}}. \qquad (26)
$$

Then it follows that the Pareto-optimal solution (i.e., optimal in the sense of component-wise minimization of the T_m's) is [20, 25]:

$$
T_m = \frac{\frac{2\gamma\eta}{h_m \mathcal{X}^2_{2L,p_a}}}{1 - \alpha \frac{2\gamma}{\mathcal{X}^2_{2L,p_a} + 2\gamma}}, \qquad m = 1, \cdots, M. \qquad (27)
$$

We note that the received powers at all the states are the same, that is, $T_1 h_1 = T_2 h_2 = \cdots = T_M h_M$.

5.1.2 Selection Combining

To fulfill the SIR requirements in the SC method, we use (24) to obtain

$$
T_m h_m \beta_0 \geq \frac{\gamma}{-\ln\left(1 - (1 - P_a)^{1/L}\right)}, \qquad m = 1, \cdots, M. \qquad (28)
$$

Along the same lines as in the MRC method, it can be shown that the Pareto-optimal solution for the T_m's is

$$
T_m = \frac{\frac{\gamma\eta}{-h_m \ln\left(1 - (1 - P_a)^{1/L}\right)}}{1 - \alpha \frac{\gamma}{\gamma - \ln\left(1 - (1 - P_a)^{1/L}\right)}}, \qquad m = 1, \cdots, M. \qquad (29)
$$

5.2 Noisy Side Information of Large-Scale Fading

Recall that in a more realistic communication system, due to time-varying channel conditions, feedback errors and delays are often inevitable. In this case, the side information about large-scale fading might be noisy. In this subsection, we study the problems of transmission power allocation under these conditions.

5.2.1 Maximum Ratio Combining

Based on (23), it is desirable to allocate the transmission power for each state such that $\gamma/T_m g\beta_0$ is "close" to $\frac{1}{2}\mathcal{X}^2_{2L,p_a}$. Note from (25) that $2\gamma/\mathcal{X}^2_{2L,p_a}$ can be interpreted as the *effective target SIR* in the perfect side information case. (In practice, SIR is usually expressed in decibels (dB).) When the side information is noisy, however, a closed-form solution for p_m seems difficult (if not impossible) to obtain so we cannot use p_m as a basis for the allocation of the T_m's. Observing that large-scale fading may vary over a range more than 60 dB [8, Chapter 2], we choose the minimum-mean-square-error criterion instead; that is, the T_m's are chosen to minimize

$$\mathcal{J}_m = \mathbb{E}\left[\left(10\log\frac{2\gamma}{\mathcal{X}^2_{2L,p_a}} - 10\log_{10}(T_m g\beta_0)\right)^2 \middle| \hat{g} = h_m\right], \qquad m = 1, \cdots, M. \tag{30}$$

That is, the T_m's are chosen to minimize the mean square difference between the effective target SIR and the SIR achieved by using the T_m's (in dB). We note that in the case where the side-information about the large-scale fading is perfect, we can make $\mathcal{J}_m$ precisely zero by allotting the T_m's according to (27).

After some algebra, we obtain that the T_m's minimizing the $\mathcal{J}_m$'s are the solution to the following set of equations:

$$T_m h_m \beta_0 = \frac{2\gamma}{\mathcal{X}^2_{2L,p_a}} \frac{h_m}{\mathbb{E}[g|\hat{g} = h_m]}, \qquad m = 1, \cdots, M. \tag{31}$$

When the side information about the large-scale fading is noisy, β_0 is the unique positive solution to the following fixed point equation:

$$\beta_0 = \frac{1}{\eta + \alpha\sum_{m=1}^M \sum_{n=1}^M \frac{\pi_n P_e^{(m,n)} T_m h_n}{1 + T_m h_n \beta_0}}.$$

Define for $m = 1, \cdots, M$,

$$\beta_{m,\mathrm{mrc}} \triangleq \frac{2\gamma}{\mathcal{X}^2_{2L,p_a}} \frac{h_m}{\mathbb{E}[g|\hat{g} = h_m]}.$$

It is worth noting that the $\beta_{m,\mathrm{mrc}}$'s can be interpreted as the *effective target SIR* in the noisy side information case. Since

$$\begin{aligned} P\{g = h_n|\hat{g} = h_m\} &= \frac{P\{g = h_n, \hat{g} = h_m\}}{P\{\hat{g} = h_m\}} \\ &= \frac{\pi_n P_e^{(m,n)}}{\sum_{n=1}^M \pi_n P_e^{(m,n)}}, \end{aligned}$$

it then follows that $\beta_{m,\mathrm{mrc}}$ can be expressed (in dB) as

$$\beta_{m,\mathrm{mrc}} = 10\log_{10}\frac{2\gamma}{\mathcal{X}^2_{2L,p_a}} + 10\log_{10} h_m - 10\log_{10}\frac{\sum_{n=1}^M h_n\pi_n P_e^{(m,n)}}{\sum_{n=1}^M \pi_n P_e^{(m,n)}}. \tag{32}$$

Note that for state m, the effect of the estimation error of large-scale fading is quantified by $10\log h_m - 10\log\left(\sum_{n=1}^M h_n\pi_n P_e^{(m,n)} / \sum_{n=1}^M \pi_n P_e^{(m,n)}\right)$.

It can be shown that the desired solution for the T_m's is the Pareto-optimal solution to the following set of inequalities:

$$\beta_{m,\mathrm{mrc}} \le \frac{T_m h_m}{\eta + \alpha\sum_{j=1}^M\sum_{n=1}^M \frac{\pi_n P_e^{(j,n)} T_j h_n}{1+\frac{h_n}{h_j}\beta_{j,\mathrm{mrc}}}}, \qquad m = 1,\cdots,M.$$

The optimal allocation for the T_m's is therefore

$$T_m = \frac{\frac{\eta\beta_{m,\mathrm{mrc}}}{h_m}}{1-\alpha\sum_{m=1}^M\sum_{n=1}^M \frac{\pi_n P_e^{(m,n)} h_n\beta_{m,\mathrm{mrc}}}{h_m+h_n\beta_{m,\mathrm{mrc}}}}, \qquad m = 1,\cdots,M. \tag{33}$$

5.2.2 Selection Combining

Similar to the argument in the preceding subsection, we choose the T_m's to minimize

$$\mathcal{J}_m = \mathbb{E}\left[\left(10\log_{10}\frac{\gamma}{-\ln\left(1-(1-P_a)^{1/L}\right)} - 10\log_{10}\left(T_m g\beta_0\right)\right)^2 \middle| \hat{g} = h_m\right], \quad m = 1,\cdots,M. \tag{34}$$

Define, for $m = 1,\cdots,M$,

$$\beta_{m,\mathrm{sc}} \triangleq \frac{\gamma}{-\ln\left(1-(1-P_a)^{1/L}\right)}\frac{h_m}{\mathbb{E}[g|\hat{g}=h_m]}.$$

Then it can be shown that

$$\beta_{m,\mathrm{sc}} = 10\log_{10}\frac{\gamma}{-\ln\left(1-(1-P_a)^{1/L}\right)} + 10\log_{10} h_m - 10\log_{10}\frac{\sum_{n=1}^M h_n\pi_n P_e^{(m,n)}}{\sum_{n=1}^M \pi_n P_e^{(m,n)}},$$

and the Pareto-optimal allocation for the T_m's is given by

$$T_m = \frac{\frac{\eta\beta_{m,\mathrm{sc}}}{h_m}}{1-\alpha\sum_{m=1}^M\sum_{n=1}^M \frac{\pi_n P_e^{(m,n)} h_n\beta_{m,\mathrm{sc}}}{h_m+h_n\beta_{m,\mathrm{sc}}}}, \qquad m = 1,\cdots,M. \tag{35}$$

6 Numerical Results

In this section, we provide a numerical example to illustrate our results. Our objectives are threefold. First, we use Figures 3 and 4 to illustrate the significant gains for the transmission power

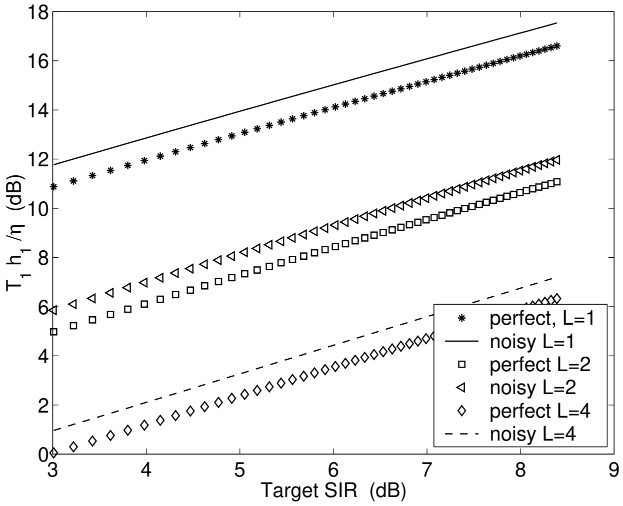

Figure 3: $\frac{T_1 h_1}{\eta}$ vs. target SIR in the MRC method, $\alpha = 0.5$.

allocation by using space diversity. Recall that the h_m's span over a wide region but the $T_m h_m$'s have the same order of magnitude. Therefore, we plot $T_1 h_1/\eta$ (instead of T_1) as a function of γ. Our next objective is to show the impact of space diversity on the system feasibility, as in Figures 5 and 6. Finally, the impact of noisy side information of large-scale fading on the transmission power allocation is illustrated in Figure 7.

In our numerical example, the model for the large-scale fading is based on a histogram of excess path loss measured in New Providence, New Jersey [8, p. 119]. (We note that the measurements were taken at 11.2 GHZ, which indicates that fading there was more severe than in current CDMA systems.) For simplicity, we quantize the large-scale fading into six levels as follows:

$$h_1 = -27dB, \quad h_2 = -30dB, \quad h_3 = -33dB,$$
$$h_4 = -36dB, \quad h_5 = -39dB, \quad h_6 = -40dB,$$

and the corresponding distribution is

$$P\{g = h_1\} = 0.1, \quad P\{g = h_2\} = 0.2, \quad P\{g = h_3\} = 0.2,$$
$$P\{g = h_4\} = 0.2, \quad P\{g = h_5\} = 0.2, \quad P\{g = h_6\} = 0.1.$$

We assume that the estimation error probability is given by

$$P_e^{(m,n)} = \begin{cases} 0.1 & |m-n| = 1 \\ 0 & |m-n| > 1. \end{cases}$$

Figures 3 and 4 show plots of $T_1 h_1/\eta$ as a function of the target SIR γ and the load α (the number of users per unit processing gain), respectively. Several observations are worth noting. First, there is nearly a 6dB gain for the transmission power allocation with four versus two receiver

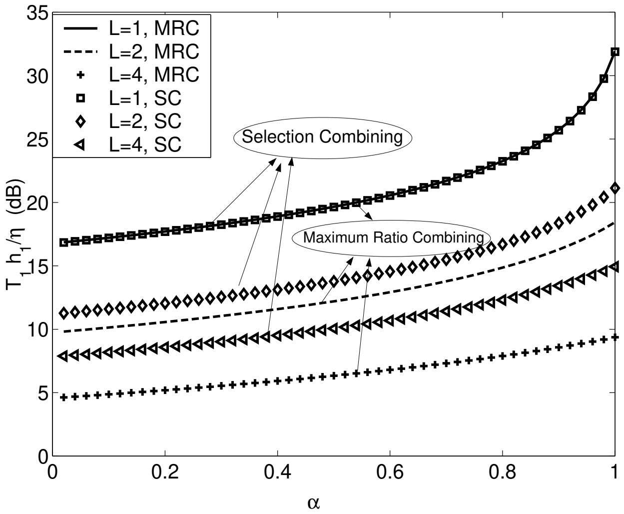

Figure 4: $\frac{T_1 h_1}{\eta}$ vs. α in both the MRC and SC methods, $p_a = 0.75$.

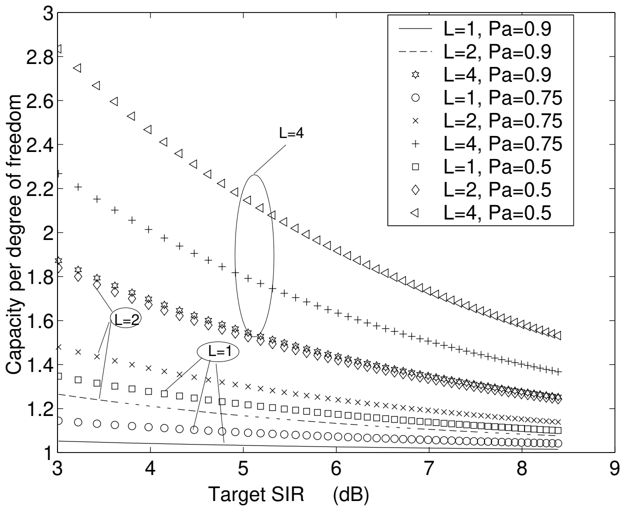

Figure 5: Network Capacity vs. target SIR in the MRC method.

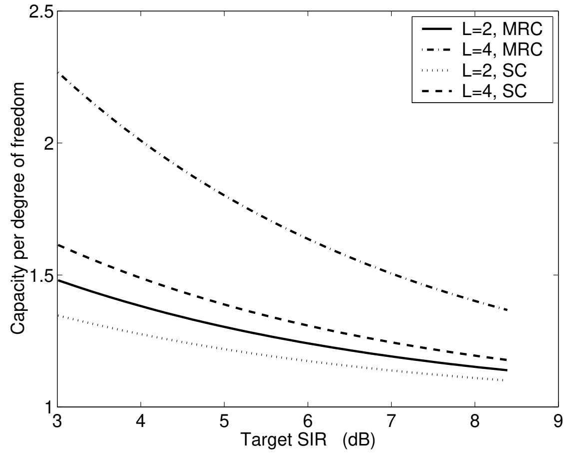

Figure 6: Network Capacity vs. target SIR in both the MRC and SC methods, $P_a = 0.75$.

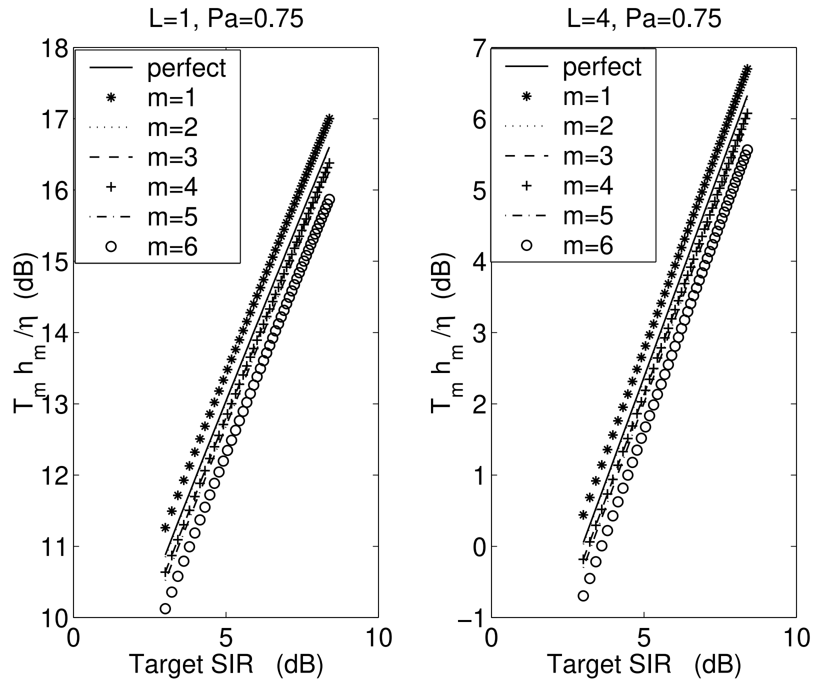

Figure 7: $\frac{T_m h_m}{\eta}$ vs. target SIR in the MRC method, $m = 1, \cdots, 6$.

antennas when the MRC method is used and the load α is moderate; the gain in the SC method is still pronounced although smaller. Second, when the system load is high, the gain for transmission power allocation is even higher. Third, the gain with four versus two receiver antennas is higher than that with two versus one receiver antennas.

To illustrate the impact of diversity combining on the system feasibility, we consider the case where the side information about the large-scale fading is perfect. Following [20, 24], we can obtain that the network capacities in the MRC and SC methods are

(MRC) $\alpha_{\mathrm{mrc}} = 1 + \mathcal{X}_{2L}^2/2\gamma$ users per unit processing gain;

(SC) $\alpha_{\mathrm{sc}} = 1 - \ln\left(1 - (1 - P_a)^{1/L}\right)$ users per unit processing gain.

As is evident in Figures 5 and 6, the increase of network capacity is significant, especially when the users' SIR requirements are not stringent. Moreover, the increase of the network capacity in the MRC method is more significant than that in the SC method. When users' SIR requirements become more stringent, the increase of the network capacity in both MRC and SC methods decreases.

Figure 7 is used to show the impact of noisy side information of large-scale fading on the transmission power allocation. Again, we plot $T_m h_m/\eta$ (instead of T_m) against γ because h_m/η is a constant. We observe that in the case of perfect side information the values of $T_m h_m/\eta$ are equal for all m, and in the case of noisy side information, the values of $T_m h_m/\eta$ are spread out around that in the perfect side information case. The spread is small and is about 1dB, which indicates that adopting the minimum-mean-square-error criterion leads to desirable power allocation. We also note that the more severe the large-scale fading, the smaller the values of $T_m h_m/\eta$.

7 Conclusions

In this work we focused primarily on the performance of large CDMA systems in a multiple time-scale flat fading environment. In particular, we established that the interferences across the antennas are jointly Gaussian in a large system, and characterized the SIR for both MRC and SC methods. Still, we believe that our study can be extended to models with frequency-selective small-scale fading. For the adaptive implementation of diversity combining, there is a large literature covering this topic in the context of narrowband systems (see, e.g., [3, 15]). Assuming fixed channel gains (fading effects), a recent work [23] has addressed the problem of joint power control, multiuser detection, and diversity combining in a CDMA system. The development of adaptive algorithms for joint power control, multiuser detection, and diversity combining in a (multiple time-scale) fading environment remains open. Another important issue is the feasibility of power control in a multi-cell setting where the side information about the fading is noisy, as is typically the case in a practical system. We are currently exploiting a recent result [4] on perturbations of the Perron-Frobenius eigenvalues to study this problem.

Acknowledgments

The authors would like to thank the anonymous reviewers for their helpful comments that improved the presentation of the paper.

Appendices

A Proof of Lemma 3.2

Part (i) follows from [25, Lemma 4.3]. In what follows, we prove Parts (ii) and (iii).

Proof of part (ii):

Because $a_{k\ell_1}$ and $a_{k\ell_2}$ are independent for any $\ell_1 \neq \ell_2$, it is clear that

$$\mathbb{E}\left[s_1^H M_I^{-1} \left[\sum_{k=2}^{K} a_{k\ell_1} a_{k\ell_2}^* T(\hat{g}_k) g_k s_k s_k^H \right] M_I^{-1} s_1 \right] = 0.$$

Using Chebyshev's Inequality, it suffices to show that as $N \to \infty$,

$$\mathbb{E}\left\{ \left| s_1^H M_I^{-1} \left[\sum_{k=2}^{K} a_{k\ell_1} a_{k\ell_2}^* T(\hat{g}_k) g_k s_k s_k^H \right] M_I^{-1} s_1 \right|^2 \right\} \to 0.$$

For convenience, we define $\hat{T} \triangleq \max_{m,n} T_m h_n$. Observe that[5]

$$\begin{aligned}
& \mathbb{E}\left\{ \left| s_1^H M_I^{-1} \left(\sum_{k=2}^{K} a_{k\ell_1} a_{k\ell_2}^* T(\hat{g}_k) g_k s_k s_k^H \right) M_I^{-1} s_1 \right|^2 \right\} \\
= \ & \mathbb{E}\left\{ \sum_{k_1=2}^{K} a_{k_1\ell_1} a_{k_1\ell_2}^* T(g_{k_1}) g_{k_1} |s_1^H M_{I_1}^{-1} s_{k_1}|^2 \sum_{k_2=2}^{K} a_{k_2\ell_1}^* a_{k_2\ell_2} T(g_{k_2}) g_{k_2} |s_1^H M_{I_2}^{-1} s_{k_2}|^2 \right\} \\
\overset{(a)}{=} \ & \mathbb{E}\left\{ \sum_{k=2}^{K} (T(\hat{g}_k) g_k)^2 |s_1^H M_I^{-1} s_k|^4 \right\} \\
\overset{(b)}{\leq} \ & \sum_{k=2}^{K} (T(\hat{g}_k) g_k)^2 \frac{C_1}{N^2 \eta^4} \\
\leq \ & \frac{K (\hat{T})^2 C_1}{N^2 \eta^4} \\
\to \ & 0,
\end{aligned} \tag{36}$$

where (a) follows from the fact that each of these expectations is non-zero only when $k_1 = k_2$, and (b) results from Part (i).

Proof of part (iii):

[5] The expectation in different lines may be taken over different random elements.

It is easy to show that

$$\mathbb{E}\left[s_1^H M_I^{-1}\left(\sum_{k=2}^{K}\left(|a_{k\ell}|^2-1\right)T(\hat{g}_k)g_k s_k s_k^H\right)M_I^{-1}s_1\right]=0.$$

By the conditional variance formula [16, p. 51], we have that

$$\begin{aligned}
&\operatorname{var}\left(\sum_{k=2}^{K}T(\hat{g}_k)g_k\left(|a_{k\ell}|^2-1\right)|s_1^H M_I^{-1}s_k|^2\right)\\
=\;&\mathbb{E}\left[\operatorname{var}\left(\sum_{k=2}^{K}T(\hat{g}_k)g_k\left(|a_{k\ell}|^2-1\right)|s_1^H M_I^{-1}s_k|^2\middle|S\right)\right]\\
&+\operatorname{var}\left(\mathbb{E}\left[\sum_{k=2}^{K}T(\hat{g}_k)g_k\left(|a_{k\ell}|^2-1\right)|s_1^H M_I^{-1}s_k|^2\middle|S\right]\right)\\
=\;&\sum_{k=2}^{K}(T(\hat{g}_k)g_k)^2|s_1^H M_I^{-1}s_k|^4\\
\to\;&0,
\end{aligned}\tag{37}$$

where the last steps follows from (36). Therefore, as $N\to\infty$,

$$s_1^H M_I^{-1}\left[\sum_{k=2}^{K}\left(|a_{k\ell}|^2-1\right)T(\hat{g}_k)g_k s_k s_k^H\right]M_I^{-1}s_1\xrightarrow{P}0,$$

thus completing the proof.

References

[1] T. A. Azlarov and N. A. Volodin, *Characterization Problems Associated with the Exponential Distribution.* New York: Springer-Verlag, 1986.

[2] P. Billingsley, *Probability and Measure.* John Wiley & Sons, Inc., third ed., 1995.

[3] R. L. Cupo, G. D. Golden, C. C. Martin, K. L. Sherman, N. R. Sollenberger, J. H. Winters, and P. W. Wolniansky, "A four-element adaptive antenna array for IS-136 PCS base stations," in *Proc. VTC'97*, pp. 1577–1581, May 1997.

[4] L. Elsner and I. Koltracht, "On computations of the Perron root," *SIAM Journal on Matrix Analysis and Applications*, vol. 14, no. 2, pp. 456–467, 1993.

[5] S. V. Hanly and D. Tse, "Resource pooling and effective bandwidths in CDMA networks with multiuser receivers and spatial diversity," *IEEE Transactions on Information Theory*, to appear.

[6] M. L. Honig and V. Poor, "Adaptive interference suppression," in *Wireless Communications: Signal Processing Perspectives* (H. V. Poor and G. W. Wornell, eds.), pp. 64–128, Upper Saddle River, NJ: Prentice-Hall, 1998.

[7] R. A. Horn and C. A. Johnson, *Matrix Analysis.* Cambridge University Press, 1985.

[8] W. C. Jakes, *Microwave Mobile Communication.* Piscataway, NJ: IEEE Press, second ed., 1994.

[9] U. Madhow and M. L. Honig, "MMSE interference suppression for directed-sequence spread-spectrum CDMA," *IEEE Transactions on Communications*, vol. 42, pp. 3178–3188, Dec. 1994.

[10] D. L. McLeish, "Dependent central limit theorems and invariance principles," *Annals of Probability*, vol. 2, no. 4, pp. 620–628, 1974.

[11] D. Mitra and J. A. Morrison, "A distributed power control algorithm for bursty transmissions on cellular, spread spectrum wireless networks," in *Proc. 5th WINLAB Workshop on Third Generation Wireless Information Networks* (J. M. Holtzman, ed.), pp. 201–212, Kluwer Academic Publishers, 1996.

[12] A. F. Naguib, A. Paulraj, and T. Kailath, "Capacity improvement with base-station antenna arrays in cellular CDMA," *IEEE Transactions on Vehicular Technology*, vol. 43, pp. 691–698, Aug. 1994.

[13] F. D. Nesser and J. L. Massey, "Proper complex random processes with applications to information theory," *IEEE Transactions on Information Theory*, vol. 39, pp. 1293–1302, July 1993.

[14] R. L. Peterson, R. E. Ziemer, and D. E. Borth, *Introduction to Spread Spectrum Communications.* Prentice Hall International, Inc., 1995.

[15] F. Rashid-Farrokhi, L. Tassiulas, and K. J. R. Liu, "Joint optimal power control and beamforming in wireless networks using antenna arrays," *IEEE Transactions on Communications*, vol. 46, no. 10, pp. 1313–1324, Oct. 1998.

[16] S. M. Ross, *Stochastic Processes.* John Wiley & Sons, Inc., 1995.

[17] J. Salz and P. Balabab, "Optimal diversity combining and equalization in digital data transmission with applications to cellular mobile radio – Part I: Theoretical considerations," *IEEE Transactions on Communications*, vol. 40, no. 5, pp. 885–894, 1992.

[18] M. Schwartz, W. R. Bennett, and S. Stein, *Communication Systems and Techniques.* New York: Wiley, 1974.

[19] J. W. Silverstein and Z. D. Bai, "On the empirical distribution of eigenvalues of a class of large dimensional random matrices," *Journal of Multivariate Analysis*, vol. 54, no. 2, pp. 175–192, 1995.

[20] D. Tse and S. V. Hanly, "Linear multiuser receivers: effective interference, effective bandwidth and user capacity," *IEEE Transactions on Information Theory*, vol. 45, pp. 641–657, Mar. 1999.

[21] A. J. Viterbi, *CDMA—Principles of Spread Spectrum Communications.* Addison–Wesley, 1995.

[22] G. W. Wornell, "Linear diversity techniques for fading channels," in *Wireless Communications: Signal Processing Perspectives* (H. V. Poor and G. W. Wornell, eds.), pp. 1–63, Upper Saddle River, NJ: Prentice-Hall, 1998.

[23] A. Yener, R. D. Yates, and S. Ulukus, "Joint power control, multiuser detection and beamforming for CDMA systems," in *Proc. VTC'99*, May 1999.

[24] J. Zhang and E. K. P. Chong, "CDMA systems in fading channels: Admissibility, network capacity, and power control," *IEEE Transactions on Information Theory*, vol. 46, pp. 962–981, May 2000.

[25] J. Zhang, E. K. P. Chong, and D. N. C. Tse, "Output MAI distributions of linear MMSE multiuser receivers in DS-CDMA systems," *IEEE Transactions on Information Theory*, Mar. 2001.