Performance Aspects of Linear Multi-User Receivers

Brett Ninness* Steven Weller†

Abstract

Recent works on the analysis of linear multiuser receivers for DS-CDMA applications have led to capacity-relevant expressions that are the solution of an integral equation involving Stieltjes transforms of the distribution of transmit powers. The key tools employed in these works are new results on the asymptotic eigenvalue distributions of random matrices. Unfortunately, it is only in the particular case of the transmit powers being equal that the integral equation has a closed form solution. This paper addresses the same problems pioneered in the afore-mentioned works, but demonstrates how an alternative solution is available that, while appealing to simpler mathematical ideas (principally, the law of large numbers), also offers flexibility in that the results obtained apply for arbitrary (as opposed to strictly constant) received powers. An additional advantage is that the convergence rate of the approximation used here can also be quantified.

Technical Report EE2001002, Department of Electrical and Computer Engineering, University of Newcastle,AUSTRALIA

1 Introduction

The problem considered here is that of linear demodulation of a Direct-Sequence, Code-Division Multiple-Access (DS-CDMA) signal, with a focus on providing a simple approximation that quantifies the interfering effect (on a given user) of other users and background noise.

The nature of study pursued here is inspired by the recent works [9, 7] where attention is targeted on quantifying the so-called 'signal-to-interference ratio' (SIR) in such a manner that the contributions of the various effects of processing gain, spreading sequence length, user transmit powers, and background noise are clearly exposed.

There are two key tools employed in [7, 9]. The first is to model the spreading sequences of individual users as realisations of zero mean, unit variance i.i.d. stochastic processes (this strategy

*This work was supported by the Centre for Integrated Dynamics and Control. This author is with the Department of Electrical and Computer Engineering, University of Newcastle, Australia and can be contacted at email:brett@ee.newcastle.edu.au or FAX: +61 2 49 21 69 93

†This author is with the Department of Electrical and Computer Engineering, University of Newcastle, Australia and can be contacted at email:steve@ee.newcastle.edu.au or FAX: +61 2 49 21 69 93

has been employed in other works [10] as well). The second, is to recognise that the SIR, being then a quadratic form in a random matrix, depends very explicitly on the eigenvalue distribution of that random matrix. Therefore recent results [5] from the mathematical statistics literature providing a characterisation of this distribution may be employed to provide engineering-relevant expressions.

This approach then leads to a key result (Theorem 3.1 of [7]) that provides (for finite coding gains and user numbers) an approximation for the SIR. Unfortunately, in general this expression is only available in an implicit form that is characterised as the solution to a certain integral equation. A consequence of this is that a closed form SIR approximation is then only available for the special case of equal received powers (perfect power control).

As well, the methods used to derive the results of [7] involve relatively sophisticated ideas such as Stieltjes transforms (and hence distributions which must be understood in a measure-theoretic sense), and these may prove to be a limiting factor in the penetration of the results and the understanding of their genesis.

By way of contrast, this paper illustrates how SIR approximations may be derived using only very simple ideas (the Matrix Inversion Lemma and the Law of Large Numbers) and in such a way that a closed form expression is provided for arbitrary distributions of received powers.

2 Problem Description

Of interest is the following chip-sampled discrete time model for a symbol-synchronous DS-CDMA system

$$Y = \sum_{i=1}^{K} x_i S_i + W. \tag{1}$$

Here $x_i \in \mathbf{R}$ is the symbol transmitted by the i'th user who possesses spreading sequence $S_i \in \mathbf{R}^N$. The length of the signature sequence S_i is thus N, which is also termed the 'processing gain'. Therefore, via (1) the received signal vector $Y \in \mathbf{R}^N$ consists of the linear superposition of the signals sent by all K users together with the additive noise vector $W \in \mathbf{R}^N$. Here, as in [7], this noise will be modelled as white and Gaussian so that $W \sim \mathcal{N}(0, \sigma^2 I)$.

Additionally, the symbols $\{x_i\}$ will also be modelled as random variables, all independent from one another, and such that their mean and variance satisfy

$$\mathbf{E}\{x_i\} = 0, \qquad \mathbf{E}\{x_i^2\} = p_i$$

so that p_i is the received power of user number i.

For the purposes of demodulating the signal sent by user $i = 1$, it is then useful to think of the received signal as

$$Y = x_1 S_1 + Z \tag{2}$$

where now $Z \in \mathbf{R}^N$ represents a composite disturbance to the reception of the signal from user 1, and consists of the received signals of all other users together with the background noise:

$$Z \triangleq \sum_{i=2}^{K} x_i S_i + W.$$

Therefore, the variance of Z conditional upon the receiver knowing the signature sequences $\{s_2, \cdots, s_K\}$ is

$$P_z \triangleq \mathbf{E}\left\{ZZ^T \mid S\right\} = SDS^T + \sigma^2 I$$

where

$$S \triangleq [S_2, S_3, \cdots, S_K], \qquad D = \begin{bmatrix} p_2 & 0 & \cdots & 0 \\ 0 & p_3 & & \\ \vdots & & \ddots & \\ 0 & \cdots & \cdots & p_K \end{bmatrix}.$$

In order to form the demodulated estimate $\widehat{x}_1$ of the transmitted symbol x_1, then (motivated by issues of computational efficiency) the class of linear receivers of the form

$$\widehat{x}_1 = C^T Y, \qquad C \in \mathbf{R}^N$$

is of interest. As established in [3], the choice

$$C = (p_1 S_1 S_1^T + P_z)^{-1} p_1 S_1 = \frac{p_1 P_Z^{-1} S_1}{1 + p_1 S_1^T P_Z^{-1} S_1} \tag{3}$$

provides the minimum mean square error solution that minimises $\mathbf{E}\{(x_1 - C^T Y)^2 \mid S\}$ while simultaneously maximising

$$\beta_N \triangleq \frac{p_1 (C^T S_1)^2}{C^T P_Z C} \tag{4}$$

over the class of all linear receivers. This latter quantity β_N is the so-called 'signal-to-interference ratio' (SIR) for the reception of a particular user (in this case user number 1), and (for a variety of reasons [3, 7]) it is a useful figure of merit when considering the performance of a DS-CDMA system.

Clearly, substituting (3) into (4) implies that the optimal SIR for the class of linear receivers is

$$\beta_N = p_1 S_1^T P_Z^{-1} S_1 = p_1 S_1^T (SDS^T + \sigma^2 I)^{-1} S_1. \tag{5}$$

It is then of interest to study this expression in order to gain insight into how this optimal SIR is affected by such thing as the distribution of transmit powers (according to the diagonal entries of D), and the ratio $\alpha = K/N$ of number of users K to processing gain (spreading sequence length) N.

3 Previous Work

Previous work has tackled this question of gaining insight into (5), and in particular the recent studies [7, 9] have recognised that (5) is expressible as

$$\beta_N \triangleq p_1 \sum_{k=1}^{N} \frac{(u_k)^2}{\lambda_k + \sigma^2}$$

where the numbers $\{\lambda_k\}$ are the eigenvalues of SDS^T and the numbers u_k are the projections of the elements of S_1 onto the unit length eigenvectors of SDS^T (associated with the eigenvalues $\{\lambda_k\}$ in turn). If $g(\lambda_k)$ is then the number of eigenvalues in a region Δ_k centred on λ_k, then [7] argues that in some sense $u_k^2 \approx g(\lambda_k)/\Delta_k$ and hence in the limit as the signature sequence length $N \to \infty$ and the regions Δ_k shrink

$$\lim_{N\to\infty} \beta_N = \beta = p_1 \int_0^\infty \frac{g(\lambda)}{\lambda + \sigma^2}\, \mathrm{d}\lambda. \tag{6}$$

The issue of understanding the SIR β_N then pivots on quantifying the eigenvalue density $g(\lambda)$, and a core contribution of [7, 9] is to recognise that certain eigen-structure analyses such as [5, 1] may be used for this purpose.

To be specific, firstly supposed that the spreading sequences $\{S_k\}$ are modelled as

$$S_k = \frac{1}{\sqrt{N}}[V_k(1), \cdots, V_k(N)]^T \tag{7}$$

where the elements $\{V_k(j)\}$ are identically distributed random variables that are all independent for different k or j, and such that $\mathbf{E}\{V_k(j)^2\} = 1$ for all k and j. The justification for such a description on physical grounds is discussed in [9, 8, 7].

In this case, if instead of considering $g(\lambda)$ directly, we address its so-called ‘Stieltjes Transform’ $m(z)$ defined for $z \in \mathbf{C}$ as

$$m(z) \triangleq \int_0^\infty \frac{g(\lambda)}{\lambda - z}\, \mathrm{d}\lambda \tag{8}$$

then via the results of [5] $m(z)$ satisfies the following integral equation for all $z \in \mathbf{C}, \mathrm{Im}(z) > 0$:

$$m(z) = \left(\alpha \int_0^\infty \frac{\tau f(\tau)}{1 + \tau m(z)}\, \mathrm{d}\tau - z \right)^{-1}. \tag{9}$$

Here $\alpha = K/N$ and $f(\tau)$ is the limiting density (ie. limiting shape of the histogram) of the transmit powers making up the diagonal entries of D, and

In order to clarify issues, a simplification has been made here by assuming that the densities g and f exist (otherwise Stieltjes integrals involving distributions are required), but a key point is that the above eigenvalue characterisation $g(\lambda)$ is implicit since the same term $m(z)$ appears on both the left and right hand sides of (9).

As a result of this latter issue, while the work [7] is able to provide many insights into the nature of power control in DS-CDMA systems, it is only able to provide an explicit expression for the limiting SIR when all the received powers are equal. In this situation, $p_k = p$ for all k and hence the density $f(\lambda)$ becomes a Dirac delta $f(\lambda) = \delta(\lambda - p)$ so that

$$\int_0^\infty \frac{\tau f(\tau)}{1 + \tau m(z)}\, \mathrm{d}\tau = \frac{p}{1 + p m(z)}. \tag{10}$$

Therefore, noting that (6) and (8) imply $m(-\sigma^2) = \beta/p$, and substituting this and (10) into the limit of (9) as $z \to -\sigma^2$ provides

$$\frac{\beta}{p} = \left(\frac{\alpha p}{1+\beta} + \sigma^2 \right)^{-1}.$$

This is a quadratic equation in β with positive solution

$$\beta = \frac{p}{2\sigma^2}(1-\alpha) - \frac{1}{2} + \frac{1}{2\sigma^2}\sqrt{p^2(1-\alpha)^2 + 2p\sigma^2(1+\alpha) + \sigma^4}. \tag{11}$$

Unfortunately, it appears that this constant power case, by virtue of the Dirac density it implies for $f(\lambda)$, is the only situation in which the characterisation (9) may be applied in order to provide a 'closed form' approximation to the limiting SIR β.

The purpose of this paper is to illustrate how an alternative strategy for analysing (5), by means of avoiding direct characterisation of the spectral distribution of the matrices involved, is able to provide an expression for β which applies for arbitrary received power distributions. As an ancillary benefit, the arguments used here also call upon less sophisticated mathematical ideas than the pre-existing approach just outlined.

4 A New Analysis

There is one essential tool that will be employed here in the analysis of (5), and (as well known) it is a fundamental result of probability theory. Specifically, if $\{X_k\}$ is a sequence of independent random variables for which certain regularity conditions apply (see Theorem A.1 in the appendix), then the (strong) Law of Large Numbers asserts that

$$\lim_{N\to\infty} \frac{1}{N} \sum_{k=1}^{N} (X_k - \mathbf{E}\{X_k\}) = 0, \qquad \text{w.p.1} \tag{12}$$

where the 'with probability one' (w.p.1) epithet indicates that the above limit may only fail to hold on a subset $\Omega' \subset \Omega$ (of the underlying probability space $\{\Omega, \mathcal{F}, \mathbf{P}\}$ that the $\{X_k\}$ are defined on) for which $\mathbf{P}(\Omega') = 0$. The intuitive understanding of this law is that if $\mathbf{E}\{X_k\} = \overline{X}$ for all k, then for 'large' N, the approximation

$$\frac{1}{N}\sum_{k=1}^{N} X_k \approx \frac{1}{N}\sum_{k=1}^{N} \mathbf{E}\{X_k\} = \overline{X}$$

is likely to be accurate.

The purpose of this section is to show how this single principle can be used to provide insight into (5), and perhaps the first temptation would be to apply it to the SDS^T term in (5) to try to approximate it by some diagonal matrix.

This approach would be fraught with difficulty since, by construction SDS^T is rank deficient (for the most common case of $K < N$) so that no diagonal expression is likely to be an accurate approximation; indeed, one view of the pre-existing work [7, 9] is that it deals with exactly this difficulty by means of the previously discussed eigenvalue distribution calculations.

In recognition of these pitfalls, this paper takes an alternative approach by using the Matrix Inversion Lemma [11] which states that for arbitrary matrices A, B, C, D of compatible dimensions and such that the indicated inverses exist:

$$[A + BCD]^{-1} = A^{-1} - A^{-1}B[C^{-1} + DA^{-1}B]^{-1}DA^{-1}$$

so that the formulation (5) may be re-expressed as

$$\beta_N = \frac{p_1}{\sigma^2} S_1^T \left[I - S(\sigma^2 D^{-1} + S^T S)^{-1} S^T \right] S_1. \tag{13}$$

This introduces the matrix $S^T S$ which is generically full rank, thus obviating the need to compute eigenvalue distributions.

Now, using the notation $[A]_{m,n}$ to denote the m, n'th element of an arbitrary matrix A, then under the random variable model (7)

$$\left[S^T S\right]_{m,n} = \frac{1}{N} \sum_{k=1}^{N} V_{m+1}(k) V_{n+1}(k) \tag{14}$$

(assuming matrix indexing begins at $m, n = 1$) and therefore, using $X_k = V_{m+1}(k)V_{n+1}(k)$ in the strong law of large numbers result (12) and after noting that by the independence assumptions $\mathbf{E}\{X_k\} = \delta(m-n)$ (this is the Kronecker delta) then the following approximation (valid for large N) follows

$$S^T S \approx I$$

so that for large N

$$\beta_N \approx \frac{p_1}{\sigma^2} \left[S_1^T S_1 - S_1^T S \Sigma S^T S_1 \right] \tag{15}$$

where

$$\Sigma \triangleq \begin{bmatrix} (\sigma^2/p_2 + 1)^{-1} & 0 & \cdots & 0 \\ 0 & (\sigma^2/p_3 + 1)^{-1} & & \\ \vdots & & \ddots & \\ 0 & \cdots & \cdots & (\sigma^2/p_K + 1)^{-1} \end{bmatrix}. \tag{16}$$

Now, again by using (7) and (12)

$$S_1^T S_1 = \frac{1}{N} \sum_{k=1}^{N} V_1^2(k) \approx \frac{1}{N} \sum_{k=1}^{N} \mathbf{E} \left\{ V_1^2(k) \right\} = 1. \tag{17}$$

Finally, denoting

$$\eta_k \triangleq \frac{1}{N} \sum_{n=1}^{N} V_1(n) V_k(n) \tag{18}$$

then

$$S_1^T S\Sigma S^T S_1 = \sum_{k=2}^{K} \frac{\eta_k^2}{\sigma^2/p_k + 1} \tag{19}$$

and therefore, again by (12) and now for large K

$$\frac{1}{K-1} S_1^T S\Sigma S^T S_1 \approx \frac{1}{K-1} \sum_{k=2}^{K} \frac{\mathbf{E}\{\eta_k^2\}}{\sigma^2/p_k + 1}.$$

Furthermore, by the definition (18) and the independence and unit variance assumptions on $\{V_k(n)\}$

$$\mathbf{E}\{\eta_k^2\} = \frac{1}{N^2} \sum_{n=1}^{N} \sum_{m=1}^{N} \mathbf{E}\{V_1(n)V_k(n)V_1(m)V_k(m)\} = \frac{1}{N^2} \sum_{n=1}^{N} \mathbf{E}\{V_1^2(n)V_k^2(n)\} = \frac{1}{N}$$

so that for large K and N

$$S_1^T S\Sigma S^T S_1 \approx \frac{1}{N} \sum_{k=2}^{K} \frac{1}{\sigma^2/p_k + 1}. \tag{20}$$

Substituting (20) and (17) into (15) then provides the large N and K approximation for SIR of

$$\beta_N \approx \frac{p_1}{\sigma^2} \left[1 - \frac{1}{N} \sum_{k=2}^{K} \frac{1}{\sigma^2/p_k + 1} \right]. \tag{21}$$

Clearly, this approximation holds for arbitrary distribution of the received powers $\{p_k\}$, but it is interesting to reconcile it with the pre-existing approximation (11) that applies only for equal powers. Specifically, note that if $p_k = p$ for all k, then if the SNR is high enough such that $\sigma^2/p \ll 1$ the approximation (21) becomes

$$\beta_N \approx \frac{p_1}{\sigma^2} [1 - \alpha] \tag{22}$$

where $\alpha = K/N$. At the same time, under the same high SNR assumption, then the square root in (11) is dominated by the first term so that it also implies (22) after it is recognised the $-1/2$ term is negligible.

These somwehat heuristic arguments leading to (21) are now presented more formally in the following Theorem which is the main result of the paper.

Theorem 4.1. *With β_N defined via (5), and under the random spreading sequence model (7) with $\mathbf{E}\{|V_k(n)|^4\} = \kappa < \infty$, then for some $C < \infty$ and δ, ϵ arbitrarily close (but not equal) to zero*

$$\left| \beta_N - \frac{p_1}{\sigma^2} \left[1 - \frac{1}{N} \sum_{k=2}^{K} \frac{1}{\sigma^2/p_k + 1} \right] \right| \le C \left[\left(\frac{K^{1+\delta}}{N^{3/4-\epsilon}} \right)^2 + \frac{K^{1+\delta}}{N} \right]$$

with probability one.

Proof. See Appendix B. □

Notice that a key point about this result, which discriminates it from pre-existing constant power ones [7, 9], is that it establishes a convergence-rate bound on the approximation for β_N that applies with probability one. Balancing this advantage, is the fact that the bound indicates low error in the approximation (21) only for the small α case in which $K \ll N$; although the simulation results to follow indicate that (21) is, in fact, still quite accurate for α near 1.

This feature of the accuracy depending on $\alpha = K/N$ is fundamental to the approximation strategy used here, which hinged on employing the Matrix Inversion Lemma to replace consideration of the rank deficient $SDS^T \in \mathbf{R}^{N\times N}$ with that of the full rank $S^TS \in \mathbf{R}^{(K-1)\times(K-1)}$. However, since $S \in \mathbf{R}^{N\times(K-1)}$ then S^TS is, in fact, only generically full rank if $\alpha = K/N \ll 1$, hence the approximation accuracy dependence on α.

5 Simulation Results

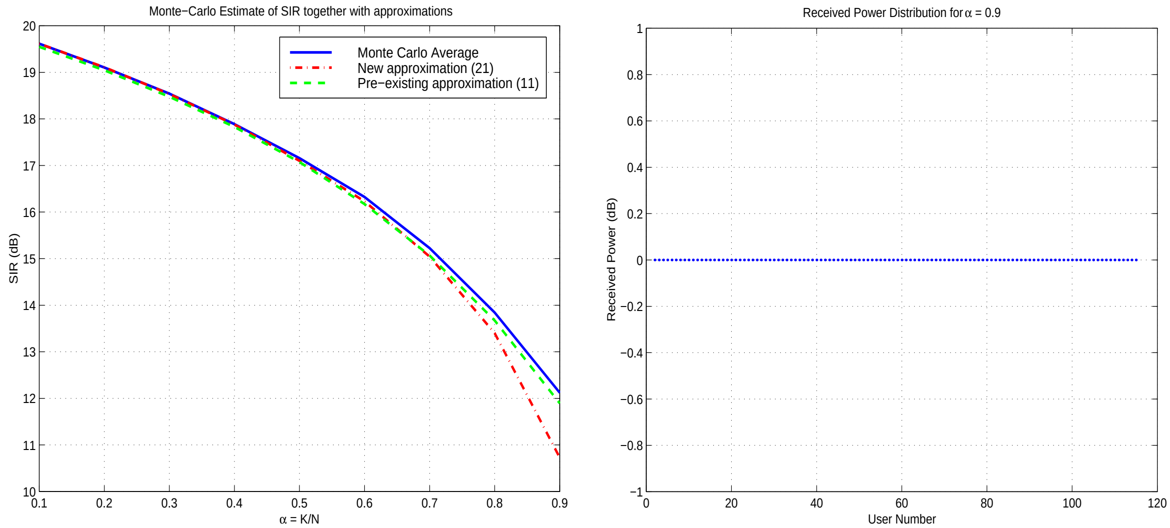
In order to assess the utility of the new SIR approximation (21) that has been derived here, this section profiles a simulation study in which (21) is compared to a Monte–Carlo estimate of SIR and also (when applicable) to the pre-existing approximation (11).

To begin with, we first address the scenario considered in [7] in which, for an SNR of $p_1^2/\sigma^2 = 20$dB and perfect power control so that the received powers $p_k = p_1$ for all k (see figure 1(b)), then the true SIR is estimated in a Monte–Carlo fashion by averaging over 1000 realisations of length $N = 128$ signature sequences $S_1, \cdots, S_K$ (generated according to (7)) and for various ratios $\alpha = K/N$ in the range $\alpha \in [0.1, 0.9]$.

The results of this Monte–Carlo estimation are shown as the solid line in figure 1(a). Also shown there as the dashed line is the pre-existing approximation (11) which applies in this special case of perfect power control. Finally, the new approximation (21) of this paper is shown as a dash-dot line in figure 1. Clearly both approximations (11) and (21) provide an excellent quantification of the true SIR, but the pre-existing one (11) is superior to that of (21) for high ratios α. This degradation in the quality of (21) is to be expected via the results of Theorem 4.1 which provide an error bound that, because of the K/N-type terms, only guarantees good accuracy for low ratios of $\alpha = K/N$. Nevertheless, even within the region of high α, the new approximation (21) is still highly informative.

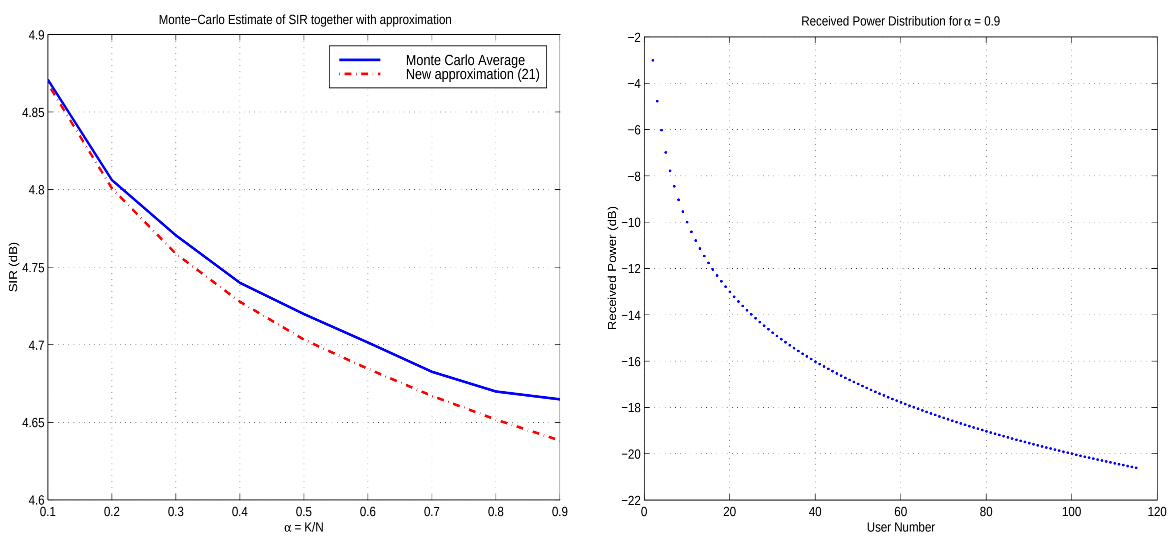
However, the chief advantage of (21) arises in scenarios of imperfect power control, for which it still applies, whereas pre-existing methods [7, 9] cannot provide an equivalent closed form expression.

In order to assess this situation, a lower SNR for user 1 of $p_1^2/\sigma^2 = 5$dB was considered in conjunction with the received powers of the other users being the non-constant distribution illustrated in figure 2(b). Shown as the solid line in figure 2(a) is the 1000 realisation Monte–Carlo estimate of SIR that corresponds to this received power distribution and various values of α. The dash-dot line is the approximation (21) for this imperfect power control situation, and considering its close agreement with the solid line, it appears to remain a highly informative approximation in this scenario. As before, it is most accurate for low values of α.



(a) SIR (Monte-Carlo and Approx) (b) Received Power Distribution

Figure 1: *The left figure shows the Monte–Carlo estimate of SIR versus* $\alpha = K/N$ *as the solid line, and it shows the pre-existing approximation (11) as the dashed line, with the new approximation (21) as the dash-dot line. The right hand figure shows the received power distribution used for the Monte–Carlo simulation.*



(a) SIR (Monte-Carlo and Approx) (b) Received Power Distribution

Figure 2: *The left figure shows the Monte–Carlo estimate of SIR versus* $\alpha = K/N$ *as the solid line, and it shows the new approximation (21) as the dash-dot line. The right hand figure shows the received power distribution used for the Monte–Carlo simulation.*

6 Conclusions

Previous authors [7, 9, 3] have established the utility of considering SIR, and asymptotically derived approximations to it, as a performance-measure of DS-CDMA systems. The work here has built upon these ideas in three ways:

1. An alternative derivation method that relies on relatively simple principles has been established;
2. Because of the directness of the methods employed here, an extension beyond previous works to the case of arbitrary received power distributions is possible;
3. The convergence rate of the approximation to the true SIR can be quantified.

The chief deficiency of the new results presented here is that, as a price paid for their simplicity of derivation, they are not as accurate as the perfect power control expressions of previous authors when $\alpha = K/N$ is close to one. Their utility in comparison to previous results is thus application dependent.

A Technical Results and Proofs

Theorem A.1. *[A Strong Law of Large Numbers] Suppose $\{X_k\}$ is a sequence of random variables, not necessarily zero mean, and with arbitrary correlation structure (not necessarily stationary) that is characterised by the existence of a $C < \infty$, $1 < \beta < \infty$ such that*

$$\sum_{k=1}^{N} \sum_{\ell=1}^{N} \boldsymbol{E} \{X_k X_\ell\} \leq C N^{\beta}.$$

Then for any $\alpha > \beta/2$

$$\frac{1}{N^\alpha} \sum_{K=1}^{N} X_k \xrightarrow{a.s.} 0 \qquad \text{as } N \to \infty.$$

Proof. See [4] for a proof of the Theorem as stated, or see Theorem 3.7.2 [6] for a slightly weaker result that is still adequate for the purposes of this paper. □

B Proof of Theorem 4.1

Proof. In what follows, C will denote an unspecified but guaranteed finite quantity that may be different in different parts of the proof. Returning to the formulation (13)

$$\beta_N = \frac{p_1}{\sigma^2} S_1^T \left[I - S(\sigma^2 D^{-1} + S^T S)^{-1} S^T \right] S_1, \tag{B.23}$$

then initially focusing on (14), which states that

$$\left[S^T S\right]_{m,n} = \frac{1}{N}\sum_{k=1}^{N} V_{m+1}(k)V_{n+1}(k), \tag{B.24}$$

then use of the assumptions on $\{V_n(k)\}$ provides

$$\mathbf{E}\left\{V_{m+1}(k)V_{n+1}(k)\right\} = \delta(m-n)$$

(this is Kronecker delta) and

$$\mathbf{E}\left\{V_{m+1}(k)V_{n+1}(k)V_{m+1}(j)V_{n+1}(j)\right\} = \begin{cases} \delta(k-j) & ; m \neq n \\ \begin{cases} 1 & ; k \neq j \\ \kappa & ; k = j \end{cases} & ; m = n \end{cases}$$

so that

$$\mathbf{E}\left\{\left|\sum_{k=1}^{N} V_{m+1}(k)V_{n+1}(k) - N\delta(m-n)\right|^2\right\} \leq CN.$$

Therefore, by Theorem A.1 and for some $\delta > 0$

$$\lim_{N\to\infty} \frac{1}{N^{1/2+\delta}}\left|\sum_{k=1}^{N} V_{m+1}(k)V_{n+1}(k) - N\delta(m-n)\right| = 0, \qquad \text{w.p.1}$$

so that by (B.24)

$$\left|\left[S^T S\right]_{m,n} - \delta(m-n)\right| \leq \frac{C}{N^{1/2-\delta}}, \qquad \text{w.p.1.}$$

For an arbitrary square matrix $A \in \mathbf{R}^{n\times n}$, define the norm $\|A\|$ to be the 'spectral norm' $\|A\| = \sup_{x\in\mathbf{R}^n} x^T Ax/x^T x$. Then via the above result combined with (2.3.8) of [2] and with probability one

$$S^T S = I + \Delta_1, \qquad \|\Delta_1\| \leq C\frac{K}{N^{1/2-\delta}}. \tag{B.25}$$

Using an identical argument together with the formulation (17) and Theorem A.1

$$S_1^T S_1 = 1 + \Delta_2, \qquad |\Delta_2| \leq \frac{C}{N^{1/2-\delta}}, \qquad \text{w.p.1.} \tag{B.26}$$

Substituting (B.25) and (B.26) into (13) then provides

$$\begin{aligned} \beta_N &= \frac{p_1}{\sigma^2}\left(1 - S_1^T S(\Sigma^{-1}+\Delta_1)^{-1}S^T S_1\right) + \Delta_2\frac{p_1}{\sigma^2} \\ &= \frac{p_1}{\sigma^2}\left(1 - S_1^T S\Sigma S^T S_1 + S_1^T S(\Sigma^{-1}+\Delta_1)^{-1}\Delta_1\Sigma S^T S_1\right) + \Delta_2\frac{p_1}{\sigma^2} \end{aligned} \tag{B.27}$$

where Σ is given by (16) and the Matrix Inversion Lemma has been used in progressing to the last line. Now, with the definition (18) of

$$\eta_k \triangleq \frac{1}{N}\sum_{n=1}^{N} V_1(n)V_k(n)$$

then via (19)

$$S_1^T S\Sigma S^T S_1 = \sum_{k=2}^{K} \frac{\eta_k^2}{\sigma^2/p_k + 1}$$

and under the independence, unit variance, and finite 4'th moment assumptions on $\{V_k(n)\}$

$$\mathsf{E}\left\{\eta_k^2\right\} = \frac{1}{N^2}\sum_{n=1}^{N}\sum_{m=1}^{N}\mathsf{E}\left\{V_1(n)V_1(m)V_k(n)V_k(m)\right\} = \frac{1}{N^2}\sum_{n=1}^{N}\mathsf{E}\left\{V_1(n)^2\right\}\mathsf{E}\left\{V_k(n)^2\right\} = \frac{1}{N}. \tag{B.28}$$

Furthermore,

$$\mathsf{E}\left\{\eta_k^2\eta_j^2\right\} = \frac{1}{N^4}\sum_{n=1}^{N}\sum_{m=1}^{N}\sum_{\ell=1}^{N}\sum_{r=1}^{N}\mathsf{E}\left\{V_1(n)V_1(m)V_1(\ell)V_1(r)\right\}\mathsf{E}\left\{V_k(m)V_k(n)V_j(\ell)V_j(r)\right\}$$

so that when $k \neq j$

$$\begin{aligned}
\mathsf{E}\left\{\eta_k^2\eta_j^2\right\} &= \frac{1}{N^4}\sum_{n=1}^{N}\sum_{m=1}^{N}\sum_{\ell=1}^{N}\sum_{r=1}^{N}\mathsf{E}\left\{V_1(n)V_1(m)V_1(\ell)V_1(r)\right\}\mathsf{E}\left\{V_k(m)V_k(n)\right\}\mathsf{E}\left\{V_j(\ell)V_j(r)\right\} \\
&= \frac{1}{N^4}\sum_{n=1}^{N}\sum_{m=1}^{N}\sum_{\ell=1}^{N}\sum_{r=1}^{N}\mathsf{E}\left\{V_1(n)V_1(m)V_1(\ell)V_1(r)\right\}\delta(n-m)\delta(\ell-r) \\
&= \frac{1}{N^4}\sum_{n=1}^{N}\sum_{\ell=1}^{N}\mathsf{E}\left\{V_1^2(n)V_1^2(\ell)\right\} \\
&= \frac{1}{N^4}(N\kappa + N(N-1))
\end{aligned}$$

while when $k = j$

$$\begin{aligned}
\mathsf{E}\left\{\eta_k^2\eta_j^2\right\} &= \frac{1}{N^4}\sum_{n=1}^{N}\sum_{m=1}^{N}\sum_{\ell=1}^{N}\sum_{r=1}^{N}\mathsf{E}\left\{V_1(n)V_1(m)V_1(\ell)V_1(r)\right\}\mathsf{E}\left\{V_k(m)V_k(n)V_k(\ell)V_k(r)\right\} \\
&= \frac{1}{N^4}\left[N\kappa^2 + N(3N-1)\right]
\end{aligned}$$

where the last line follows since the expectations will be zero unless there are two matched pairs of indices, and taking the first index n, as it ranges through N values, it can match 3 other indices, each of which can range through N possible values. Therefore, there are $3N^2$ times that the above quadruple summation involves pairs of matched indices, and of these, N occurrences involve all four indices being matched. Therefore

$$\mathbf{E}\left\{\eta_k^2\eta_j^2\right\} = \begin{cases} \dfrac{1}{N^3}\left[\kappa^2 + (3N-1)\right] & ; k = j \\ \dfrac{1}{N^3}\left[\kappa + (N-1)\right] & ; k \neq j \end{cases} \tag{B.29}$$

and hence

$$\begin{aligned} \sum_{k=2}^{K}\sum_{j=2}^{K} \frac{\mathbf{E}\left\{(N\eta_k^2-1)(N\eta_j^2-1)\right\}}{(\sigma^2/p_k+1)(\sigma^2/p_j+1)} &= \sum_{k=2}^{K}\sum_{j=2}^{K} \frac{N^2\mathbf{E}\left\{\eta_k^2\eta_j^2\right\}-1}{(\sigma^2/p_k+1)(\sigma^2/p_j+1)} \\ &= \frac{1}{N}(\kappa^2+3N-1)\sum_{k=2}^{K}\frac{1}{(\sigma^2/p_k+1)^2} + \\ &\quad \frac{1}{N}[\kappa+N-1]\left[\left(\sum_{k=2}^{K}\frac{1}{\sigma^2/p_k+1}\right)^2 - \sum_{k=2}^{K}\frac{1}{(\sigma^2/p_k+1)^2}\right] - \\ &\quad \left(\sum_{k=2}^{K}\frac{1}{\sigma^2/p_k+1}\right)^2 \\ &\leq CK^2. \end{aligned}$$

Therefore, by Theorem A.1, and for some $\delta > 0$

$$\lim_{K\to\infty}\frac{1}{K^{1+\delta}}\sum_{k=2}^{K}\frac{N\eta_k^2-1}{\sigma^2/p_k+1} = 0, \qquad \text{w.p.1}$$

so that since

$$N\left(S_1^T S\Sigma S^T S_1 - \frac{1}{N}\sum_{k=2}^{K}\frac{1}{\sigma^2/p_k+1}\right) = \sum_{k=2}^{K}\frac{N\eta_k^2-1}{\sigma^2/p_k+1}$$

then for some $C < \infty$ and with probability one

$$S_1^T S\Sigma S^T S_1 = \frac{1}{N}\sum_{k=2}^{K}\frac{1}{\sigma^2/p_k+1} + \Delta_3, \qquad |\Delta_3| \leq C\frac{K^{1+\delta}}{N}. \tag{B.30}$$

Therefore, using (B.27)

$$\left|\beta_N - \frac{p_1}{\sigma^2}\left(1 - \frac{1}{N}\sum_{k=2}^{K}\frac{1}{\sigma^2/p_k+1}\right)\right| = \left|\frac{p_1}{\sigma^2}\left(\Delta_3 + \Delta_2 + S_1^T S(\Sigma^{-1}+\Delta_1)^{-1}\Delta_1\Sigma S^T S_1\right)\right|. \tag{B.31}$$

Finally, by the definition of the Matrix Norm (induced 2-norm) and the Cauchy—Schwarz inequality

$$\left|S_1^T S(\Sigma^{-1}+\Delta_1)^{-1}\Delta_1\Sigma S^T S_1\right| \leq \|\Delta_1\|\cdot\|(\Sigma^{-1}+\Delta_1)^{-1}\|\cdot\|\Sigma\|\left|S_1^T S S^T S_1\right|.$$

However, since Σ is diagonal

$$\|\Sigma\| = \max_k \left|\frac{p_k}{\sigma^2+p_k}\right|$$

which is clearly bounded. Also, since matrix inversion is convex with respect to the induced 2-norm, then

$$\|(\Sigma^{-1}+\Delta_1)^{-1}\| \leq \|\Sigma\| + \|\Delta_1^{-1}\| = \|\Sigma\| + \|\Delta_1\|^{-1}$$

Therefore, by setting $\sigma^2 = 0$ in the right hand side of (B.30)

$$\left|S_1^T S(\Sigma^{-1}+\Delta_1)^{-1}\Delta_1\Sigma S^T S_1\right| \leq C\|\Delta_1\||S_1^T S S^T S_1| = C\|\Delta_1\|\left|\sum_{k=2}^{K}\eta_k^2\right| \leq C\|\Delta_1\|\Delta_3, \qquad \text{w.p.1.}$$

Using this bound in (B.31) together with the expressions (B.30) and (B.26) then completes the proof. □

References

[1] Z. D. Bai and Y. Q. Yin, *Limit of the smallest eigenvalue of a large-dimensional sample covariance matrix*, Ann. Probab., 21 (1993), pp. 1275–1294.

[2] G. Golub and C. V. Loan, *Matrix Computations*, Johns Hopkins University Press, 1989.

[3] U. Madhow and M. L. Honig, *Mmse interference suppresion for direct-sequence spread-spectrum cdma*, IEEE Transactions on Communications, 42 (1994), pp. 3178–3188.

[4] B. Ninness, *Strong laws of large numbers under weak assumptions with applications*, Submitted to IEEE Transactions on Automatic Control. Available Electronically from http://www.ee.newcastle.edu.au/users/staff/brett/, (1999).

[5] J. W. Silverstein, *On the empirical distribution of eigenvalues of a class of large dimensional random matrices*, Journal of Multivariate Analysis, 54 (1995), pp. 175–192.

[6] W. F. Stout, *Almost Sure Convergence*, Academic Press, 1974.

[7] D. N. C. Tse and S. V. Hanly, *Linear multiuser receivers: Effective interference, effective bandwidth and user capacity*, IEEE Transactions on Information Theory, 45 (1999), pp. 641–657.

[8] S. Verdu, *Multiuser Detection*, Cambridge University Press, 1998.

[9] S. VERDÚ AND S. SHAMAI (SCHITZ), *Spectral efficiency of cdma with random spreading*, IEEE Transactions on Information Theory, 45 (1999), pp. 622–640.

[10] P. VISWANATH, V. ANANTHARAM, AND D. N. C. TSE, *Optimal sequences, power control, and user capacity of synchronous CDMA systems with linear MMSE multiuser receivers*, IEEE Trans. Inform. Theory, 45 (1999), pp. 1968–1983.

[11] A. WEINMANN, *Uncertain Models and Robust Control*, Springer-Verlag, New York, 1991.