Stochastic Power Control for Cellular Radio Systems *

Sennur Ulukus **Roy D. Yates**

July 11, 1997

Department of Electrical and Computer Engineering
Wireless Information Networks Laboratory (WINLAB)
Rutgers University, PO Box 909, Piscataway NJ 08855-0909

Abstract

For wireless communication systems, iterative power control algorithms have been proposed to minimize transmitter powers while maintaining reliable communication between mobiles and base stations. To derive deterministic convergence results, these algorithms require perfect measurements of one or more of the following parameters: (i) the mobile's signal to interference ratio (SIR) at the receiver, (ii) the interference experienced by the mobile, and (iii) the bit error rate. However, these quantities are often difficult to measure and deterministic convergence results neglect the effect of stochastic measurements. In this work, we develop distributed iterative power control algorithms that use readily available measurements. Two classes of power control algorithms are proposed. Since the measurements are random, the proposed algorithms evolve stochastically and we define the convergence in terms of the mean squared error (MSE) of the power vector from the optimal power vector that is the solution of a feasible deterministic power control problem. For the first class of power control algorithms using fixed step size sequences, we obtain finite lower and upper bounds for the MSE by appropriate selection of the step size. We also show that these bounds go to zero, implying convergence in the MSE sense, as the step size goes to zero. For the second class of power control algorithms, which are based on the stochastic approximations method and use time varying step size sequences, we prove that the MSE goes to zero. Both classes of algorithms are distributed in the sense that each user needs only to know its own channel gain to its assigned base station and its own matched filter output at its assigned base station to update its power.

---

*This work was supported by NSF grant NCRI 95-06505 and was presented in part at the 1996 Allerton Conference.

1 Introduction

In cellular wireless communication systems, the aim of power control is to assign each user a transmitter power level such that all users satisfy their quality of service (QoS) requirements. The power control algorithms that have been developed to date may be classified as centralized or distributed, synchronous or asynchronous, iterative or non-iterative, constrained or unconstrained. Earlier work [1–4] identified the power control problem as an eigenvalue problem for non-negative matrices. The optimal power vector was found by inversion of a matrix which was composed of channel gains of all users. Those algorithms were non-iterative, synchronous and centralized in the sense that all the power vector components were found by a matrix inversion. Due to the computational complexity of these centralized power control algorithms, distributed versions have been developed which need only path gains that can be obtained by local measurements [5–9]. Variations of the ordinary power control problem can be found in references [10–13].

The power control algorithms that have been developed to date are *deterministic* in the sense that they require the exact knowledge or perfect estimates of some deterministic quantities such as (i) signal to interference ratio (SIR), (ii) received interference power, or (iii) bit error rate. Unfortunately, none of those quantities is easy to estimate perfectly, and deterministic convergence results are no longer valid when these deterministic variables are replaced with their random estimates. This observation highlights the need for the study of new power control algorithms that make use of available measurements, evolve stochastically, and converge in a stochastic sense.

In CDMA systems, conventional receivers consist of matched filters that are matched to the signature sequences of the users in the system. Squares of the matched filter outputs are unbiased estimates for the received energies, in the sense that the expected value of the square of a matched filter output is equal to the received energy through that matched filter. The randomness over which the expectation is taken is due to the randomness of the information bits transmitted by the users (i.e., multi access interference) and that of the ambient Gaussian channel noise. The deterministic power control approach assumes that this expectation is taken and a perfect estimate for the interference is available at each power control update. Although the expectation can be approximated by a sample average measurement of the outputs, perfect estimates require an average over an infinite number of bits between power updates. When the measurement is done only over a finite number of bits, the estimates are still random quantities and the deterministic convergence results obtained with perfect estimate assumptions are no longer valid.

Starting from a simple extension of the deterministic power control algorithms given in [7,9], this paper introduces a class of stochastic power control algorithms that are based on the observation of the matched filter outputs. Since the matched filter outputs are random, the proposed algorithms evolve in a stochastic fashion. The convergence of the algorithms is defined in terms of the mean squared error (MSE) of the power vector at any iteration from the optimal power vector. The optimal power vector is the fixed point of the deterministic power control problem where each user transmits with as little as possible power while all the users are satisfying their SIR based QoS requirements. Conditions for obtaining lower and upper bounds on the MSE are identified.

In this paper, two types of stochastic power control algorithms are proposed. These algorithms differ only in terms of the step size sequence being used. As we shall see, the step size scales the random correction term added to the current power to obtain the next power. If a fixed step size sequence is used, we prove that the fixed step size value can always be chosen small enough to have finite lower and upper bounds on the MSE given that the deterministic power control problem is feasible. We also show that these lower and upper bounds on the MSE go to zero, implying an exact convergence in the MSE sense, as the step size value goes to zero. If the step size sequence is allowed to depend on the iteration index, we show that for a particular class of step size sequences the measurement based power control algorithm converges in the MSE, i.e., MSE goes to zero, again conditioned on deterministic power control problem being feasible.

The variable step size stochastic power control algorithm is based on the idea of stochastic approximations, a method first introduced by Robbins and Monro [14], who solved a deterministic problem with unknown parameters by observing the random outputs of a controllable experiment. They defined a problem in which $M(x)$ was the expected value of a certain experiment conducted at input level x. For a fixed target α, their aim was to find that value of x for which $M(x) = \alpha$. They assumed that the exact functional form of $M(x)$ was not known by the experimenter, but the input level x at which the experiment was run could be controlled. They proposed an iterative method of changing the input level of the experiment by only observing the outputs of the experiment. They showed that the sequence of x values generated by the iterative algorithm converged to the solution of $M(x) = \alpha$ in the mean square sense. Generalizations of the Robbins-Monro method can be found in [15–20].

In the application of stochastic approximation methods to the power control problem, the random experiments are the transmissions of information bits from users to the base stations. Those experiments are stochastic due to the randomness of the transmitted bits and to the existence of additive white Gaussian noise. The experiments are controlled by the iterative update of the transmitter powers of the active users.

The proposed power control algorithms are distributed in the sense that users need to know observations only related to themselves. In particular, a user needs to know only two parameters to update its power: (1) the output of its own receiver filter (matched filter) at its assigned base station and (2) its own channel gain to its assigned base station.

2 System Model

We consider the uplink of a wireless multi-cell CDMA system with a BPSK modulation scheme. We assume that the users are already assigned to their base stations and do not consider the base station assignment problem. The number of users and the number of base stations are represented by N and M respectively. For each user i, we use p_i to denote its transmitted power. The channel gain of user j to the assigned base station of user i is represented by h_{ij}.

Users have pre-assigned, unique signature sequences which they use to modulate their information bits. The signature waveform of user i is denoted by $s_i(t)$ which is non-zero only in the bit interval $[0, T_b]$ and is normalized to unit energy, i.e., $\int_0^{T_b} s_i^2(t)dt = 1$. The receiver consists of a set of matched filters that are matched to the signature waveforms of the users. The only synchronism assumed is between each user and its assigned base station, i.e., the matched filter of user i is synchronized to the arrival delay of user i. For each user i, all other users in the same cell and the users in other cells create interference asynchronously. The relative delays of the users, which can have any value not exceeding the bit duration T_b, do not change with time. For the lth bit of a given user i, an interfering user creates interference by either bits $(l-1)$ and l or bits l and $(l+1)$, depending on whether the interfering user has a positive or negative delay relative to user i. In Figure 1, two possible situations are depicted. The delay of user j relative to the matched filter of user i is represented by d_{ij}. In Figure 1 user j has a positive delay relative to user i and creates interference to the lth bit of user i with bits $(l-1)$ and l. Similarly, user k has a negative relative delay with respect to user i and creates interference to the lth bit of user i with bits l and $(l+1)$. In order to express left, right and same-bit interferences, we define three types of cross correlations between the signature sequences of any two users i and j: $\bar{\Gamma}_{ij}$, Γ_{ij} and $\tilde{\Gamma}_{ij}$. If $d_{ij} \geq 0$ (case of user j in Figure 1), then

$$\bar{\Gamma}_{ij} = \int_0^{d_{ij}} s_i(t)s_j(T_b - d_{ij} + t)dt, \qquad \Gamma_{ij} = \int_{d_{ij}}^{T_b} s_i(t)s_j(t - d_{ij})dt, \qquad \tilde{\Gamma}_{ij} = 0 \tag{1}$$

and if $d_{ij} \leq 0$ (case of user k in Figure 1),

$$\bar{\Gamma}_{ij} = 0, \qquad \Gamma_{ij} = \int_0^{T_b + d_{ij}} s_i(t)s_j(t - d_{ij})dt, \qquad \tilde{\Gamma}_{ij} = \int_{T_b + d_{ij}}^{T_b} s_i(t)s_j(t - T_b - d_{ij})dt \tag{2}$$

Note from (1) and (2) that for each user either $\bar{\Gamma}_{ij}$ or $\tilde{\Gamma}_{ij}$ is equal to zero implying $\bar{\Gamma}_{ij}\tilde{\Gamma}_{ij} = 0$

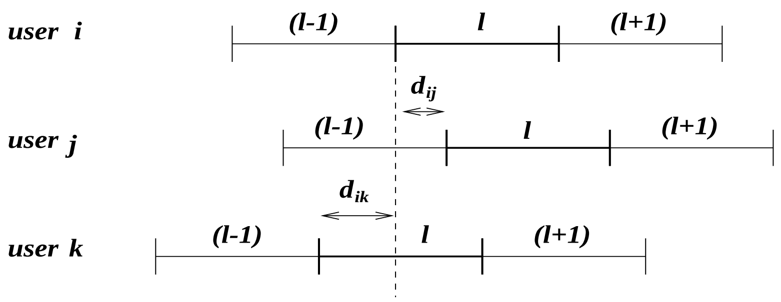

Figure 1: Asynchronous CDMA model.

for all i, j. Note also that $\Gamma_{ii} = 1$ and $\bar{\Gamma}_{ii} = \tilde{\Gamma}_{ii} = 0$. The matched filter output for the bit l of user i is

$$y_i(l) = \sum_{j=1}^{N} \sqrt{p_j}\sqrt{h_{ij}} \left(\bar{\Gamma}_{ij} b_j(l-1) + \Gamma_{ij} b_j(l) + \tilde{\Gamma}_{ij} b_j(l+1) \right) + n_i(l) \tag{3}$$

where $b_j(l)$ is the information bit of user j (± 1 equiprobably) in the lth bit interval and n_i is a sample of additive white Gaussian noise (AWGN) having zero mean and σ^2 variance. Defining

$$\hat{b}_{ij}(l) = \bar{\Gamma}_{ij} b_j(l-1) + \Gamma_{ij} b_j(l) + \tilde{\Gamma}_{ij} b_j(l+1) \tag{4}$$

we can write (3) as

$$y_i(l) = \sum_{j=1}^{N} \sqrt{p_j}\sqrt{h_{ij}} \hat{b}_{ij}(l) + n_i(l) \tag{5}$$

Note that although $b_j(l)$ and $b_j(l+k)$ are independent for $k \neq 0$, $\hat{b}_{ij}(l)$ and $\hat{b}_{ij}(l+k)$ are not for $k = \pm 1$.

3 Deterministic Power Control

The aim of a power control algorithm is to minimize the users' transmitted powers while maintaining a certain QoS for each user. Typically, QoS is defined in terms of the probability of bit error which in turn is assumed to be a monotonically decreasing function of signal to interference ratio (SIR). Therefore the QoS requirement directly translates to the SIR being larger than a *target* SIR. Let γ_i^* denote the SIR target of user i. The power control problem can be stated as follows:

$$\begin{array}{ll} \min & \sum_{i=1}^{N} p_i \\ \text{s.t.} & \frac{p_i h_{ii}}{\sum_{j \neq i} p_j h_{ij} \left(\bar{\Gamma}_{ij}^2 + \Gamma_{ij}^2 + \tilde{\Gamma}_{ij}^2 \right) + \sigma^2} \geq \gamma_i^* \qquad \text{for } i = 1, \ldots, N \end{array} \tag{6}$$

Defining the diagonal matrix $\boldsymbol{\Upsilon}$ with ith diagonal element $\Upsilon_{ii} = \gamma_i^*$, the transmitted power vector $\boldsymbol{p}$ with ith element p_i, and non-negative matrices $\boldsymbol{A} = [A_{ij}]_{N\times N}$ and $\boldsymbol{H} = [H_{ij}]_{N\times N}$ as

$$A_{ij} = \begin{cases} 0 & i = j \\ h_{ij}\left(\bar{\Gamma}_{ij}^2 + \Gamma_{ij}^2 + \tilde{\Gamma}_{ij}^2\right) & i \neq j \end{cases} \qquad \text{and} \qquad H_{ij} = \begin{cases} h_{ii} & i = j \\ 0 & i \neq j \end{cases} \tag{7}$$

the QoS requirements in (6) can be written as the vector inequality

$$\boldsymbol{p} \geq \boldsymbol{\Upsilon}\boldsymbol{H}^{-1}(\boldsymbol{A}\boldsymbol{p} + \sigma^2\boldsymbol{u}) \tag{8}$$

where $\mathbf{u} = [1\cdots 1]^\top$. We say that the set of SIR targets γ_i^*, $i = 1, \cdots, N$ are feasible if there is a non-negative finite vector $\boldsymbol{p}$ that satisfies (8).

It is not difficult to show that if the SIR targets (γ_i^*'s) are feasible then the power vector which satisfies all N inequalities with equality in (6) minimizes the sum of the transmitted powers (see [9, 10] or the appendix of [21]). Therefore, if γ_i^*, $i = 1, \cdots, N$ are feasible, a power control algorithm finds (in the iterative case converges to) the solution of

$$\boldsymbol{p} = \boldsymbol{\Upsilon}\boldsymbol{H}^{-1}(\boldsymbol{A}\boldsymbol{p} + \sigma^2\boldsymbol{u}) \tag{9}$$

The existence of a non-negative solution to (9) is specified by the following two theorems on eigenvalues of non-negative matrices by Perron and Frobenius [22].

Theorem 1 *If $\boldsymbol{S}$ is a square non-negative matrix, there exists an eigenvalue ρ called the Perron-Frobenius eigenvalue of $\boldsymbol{S}$, such that: (a) ρ is real and non-negative, (b) with ρ can be associated non-negative left and right eigenvectors, (c) $\rho \geq |\lambda|$ for any eigenvalue λ of $\boldsymbol{S}$.*

Theorem 2 *For an irreducible non-negative matrix $\boldsymbol{S}$, $\boldsymbol{x} = \boldsymbol{S}\boldsymbol{x} + \boldsymbol{c}$ has a non-negative solution $\boldsymbol{x}$ for any non-negative, nonzero $\boldsymbol{c}$ iff the Perron-Frobenius eigenvalue ρ of $\boldsymbol{S}$ satisfies $\rho < 1$.*

Thus, Equation (9) and Theorem 2 imply the following result.

Lemma 1 *The SIR targets γ_i^*, $i = 1, \cdots, N$ are feasible iff $\rho < 1$ where ρ is the Perron-Frobenius eigenvalues of matrix $\boldsymbol{\Upsilon}\boldsymbol{H}^{-1}\boldsymbol{A}$.*

For the remainder of this paper, we will assume that the deterministic power control problem is feasible. From (9), we obtain

$$\left(\boldsymbol{I} - \boldsymbol{\Upsilon}\boldsymbol{H}^{-1}\boldsymbol{A}\right)\boldsymbol{p} = \sigma^2\boldsymbol{\Upsilon}\boldsymbol{H}^{-1}\boldsymbol{u} \tag{10}$$

Note that, if the relative delays, correlation coefficients and channel gains are known, then from (10), the optimal power vector can be found as

$$\bar{\boldsymbol{p}} = \sigma^2 \boldsymbol{B}^{-1} \boldsymbol{\Upsilon} \boldsymbol{H}^{-1} \boldsymbol{u} \tag{11}$$

where

$$\boldsymbol{B} = \boldsymbol{I} - \boldsymbol{\Upsilon} \boldsymbol{H}^{-1} \boldsymbol{A} \tag{12}$$

with $\boldsymbol{I}$ the $N \times N$ identity matrix. However, using (11) to solve the power control problem is a centralized approach that requires exact knowledge of all channel gains, relative delays and corresponding correlation coefficients of all users in the system.

In [9], iterative, distributed, deterministic power control algorithms were defined in general as,

$$\boldsymbol{p}(n+1) = \boldsymbol{T}(\boldsymbol{p}(n)) \tag{13}$$

When the current power vector is $\boldsymbol{p}(n)$, $T_j(\boldsymbol{p}(n))$, the jth component of $\boldsymbol{T}(\boldsymbol{p}(n))$, is the interference that user j is required to overcome. If the SIR targets γ_i^*, $i = 1, \cdots, N$ are feasible, then the algorithm (13) will converge to the optimal solution $\bar{\boldsymbol{p}} = \boldsymbol{T}(\bar{\boldsymbol{p}})$ if $\boldsymbol{T}(\boldsymbol{p})$ is a *standard* interference function; see [9]. In the notation of the present paper, the interference function is

$$\boldsymbol{T}(\boldsymbol{p}) = \boldsymbol{\Upsilon} \boldsymbol{H}^{-1} (\boldsymbol{A}\boldsymbol{p} + \sigma^2 \boldsymbol{u}) \tag{14}$$

When $\boldsymbol{T}(\boldsymbol{p})$ is of the form of (14), the iteration (13) is the power control algorithm of Foschini and Miljanic [7, Eqn. (18)]. Implicit in the iteration (13), is that the normalized interference $T_j(\boldsymbol{p})$ that user j must overcome is measured perfectly. Hence the convergence of the iteration proceeds deterministically to the fixed point $\bar{\boldsymbol{p}}$. In the next section, we introduce stochastic power control by using the random squared matched filter outputs to measure the interference function $\boldsymbol{T}(\boldsymbol{p})$.

4 Stochastic Power Control

We consider an L-bit interval (window) in which users keep their transmitter powers fixed. We define $v_i(n, l)$ to be the squared value of the matched filter output of user i at its assigned base station at the end of lth bit interval in the nth L-bit window, $y_i(n, l)$. From (3), squaring the matched filter output $y_i(n, l)$ yields

$$v_i(n, l) = y_i^2(n, l) = \sum_{j=1}^{N} \left(\bar{\Gamma}_{ij}^2 + \Gamma_{ij}^2 + \tilde{\Gamma}_{ij}^2 \right) h_{ij} p_j(n) + \sigma^2 + w_i(n, l) \tag{15}$$

The contribution of the noise and the transmitted bits to the squared correlator output is

$$\begin{aligned} w_i(n,l) &= \sum_{j=1}^{N} \sqrt{p_j}\sqrt{h_{ij}}\Gamma_{ij}b_j(l)\left(\bar{\Gamma}_{ij}b_j(l-1)+\tilde{\Gamma}_{ij}b_j(l+1)\right) \\ &\quad + \sum_{j=1}^{N}\sum_{k\neq j}\sqrt{p_j}\sqrt{p_k}\sqrt{h_{ij}}\sqrt{h_{ik}}\hat{b}_{ij}(l)\hat{b}_{ik}(l) + 2n_i(l)\sum_{j=1}^{N}\sqrt{p_j}\sqrt{h_{ij}}\hat{b}_j(l) + \left(n_i^2(l)-\sigma^2\right) \quad (16) \end{aligned}$$

where we dropped the common index n from the terms in (16) for convenience. In a more careful statement of (16) p_j, $b_j(l)$, $\hat{b}_{ij}(l)$ and $n_i(l)$ should be written as $p_j(n)$, $b_j(n,l)$, $\hat{b}_{ij}(n,l)$ and $n_i(n,l)$, respectively. We define the average of the squared matched filter outputs of user i at its assigned base station at the end of nth L-bit interval as

$$v_i(n) = \frac{1}{L}\sum_{l=1}^{L} v_i(n,l) = \sum_{j=1}^{N}\left(\bar{\Gamma}_{ij}^2+\Gamma_{ij}^2+\tilde{\Gamma}_{ij}^2\right)h_{ij}p_j(n) + \sigma^2 + w_i(n) \quad (17)$$

where

$$w_i(n) = \frac{1}{L}\sum_{l=1}^{L} w_i(n,l) \quad (18)$$

Note that in deriving (17) we assumed that the power levels of the interfering users do not change in the L-bit window of the ith user. Since the relative delay of any interfering user is at most one bit, this assumption is valid only if users have one *idle bit*, an information bit that is unused for power control measurements, between each of their L-bit measurement windows. That is, all users keep their transmitter powers fixed for one more bit after every L-bit power control window. For practical values of L, this idle bit will not affect the performance of the overall system significantly.

Defining vectors $\boldsymbol{v}$ and $\boldsymbol{w}$ with v_i and w_i as their ith elements, and using power vector $\boldsymbol{p}$, (17) becomes

$$\boldsymbol{v} = (\boldsymbol{A}+\boldsymbol{H})\boldsymbol{p} + \sigma^2\boldsymbol{u} + \boldsymbol{w} \quad (19)$$

From (16), it follows that $E[w_i(l)] = 0$ since the transmitted bits ($b_j(l)$) are independent and equiprobably ± 1, and the zero mean noise n_i is independent of the bits and has variance σ^2. Hence,

$$\begin{aligned} E[\boldsymbol{w}] &= \boldsymbol{0} \quad (20) \\ E[\boldsymbol{v}] &= (\boldsymbol{A}+\boldsymbol{H})\boldsymbol{p} + \sigma^2\boldsymbol{u} \quad (21) \end{aligned}$$

From (21) we see that $E[v_i]$ equals the total received power through the matched filter for user i at its assigned base station. By applying (21) to (14) we can express the interference $\boldsymbol{T}(\boldsymbol{p})$ in terms of $E[\boldsymbol{v}]$ as,

$$\boldsymbol{T}(\boldsymbol{p}) = -\boldsymbol{\Upsilon}\boldsymbol{p} + \boldsymbol{\Upsilon}\boldsymbol{H}^{-1}E[\boldsymbol{v}] \tag{22}$$

Equation (22) effectively subtracts the signal component from the total received power to obtain the interference. In [9], the following deterministic power control algorithm was given.

$$\boldsymbol{p}(n+1) = (1-\epsilon)\boldsymbol{p}(n) + \epsilon\boldsymbol{T}(\boldsymbol{p}(n)) \tag{23}$$

The algorithm (23) is called interference averaging because the required power $\boldsymbol{T}(\boldsymbol{p}(n))$ at iteration n is averaged with the current power $\boldsymbol{p}(n)$ to yield the new power vector $\boldsymbol{p}(n+1)$. The motivation given in [9] for interference averaging is that in real systems, $\boldsymbol{T}(\boldsymbol{p}(n))$ must be measured and if that measurement is not accurate, then it may be desirable to make only a small adjustment in the transmitter power. By inserting (22) into (23), the interference averaging algorithm becomes

$$\boldsymbol{p}(n+1) = (1-\epsilon)\boldsymbol{p}(n) + \epsilon\left(-\boldsymbol{\Upsilon}\boldsymbol{p}(n) + \boldsymbol{\Upsilon}\boldsymbol{H}^{-1}E[\boldsymbol{v}(n)]\right) \tag{24}$$

The deterministic iteration (24) requires *perfect* knowledge of the total received power $E[\boldsymbol{v}]$. In practice we must use estimates of $E[\boldsymbol{v}]$. Thus we propose the following stochastic power control algorithm in which we replace $E[\boldsymbol{v}]$ in (24) by the unbiased estimate $\boldsymbol{v}$.

$$\boldsymbol{p}(n+1) = (1-\epsilon)\boldsymbol{p}(n) + \epsilon\left(-\boldsymbol{\Upsilon}\boldsymbol{p}(n) + \boldsymbol{\Upsilon}\boldsymbol{H}^{-1}\boldsymbol{v}(n)\right) \tag{25}$$

We note that (25) is a special case of the following more general algorithm

$$\boldsymbol{p}(n+1) = (1-a_n)\boldsymbol{p}(n) + a_n\left(-\boldsymbol{\Upsilon}\boldsymbol{p}(n) + \boldsymbol{\Upsilon}\boldsymbol{H}^{-1}\boldsymbol{v}(n)\right) \tag{26}$$

where the fixed step size ϵ is replaced with a variable step size sequence a_n that may be a function of the iteration index n. Therefore, (25) corresponds to a special case of the algorithm (26) when $a_n = \epsilon$ for all n. The iteration (26) can be re-written in the form

$$\boldsymbol{p}(n+1) = \left[1 - a_n\left(\boldsymbol{I} + \boldsymbol{\Upsilon}\right)\right]\boldsymbol{p}(n) + a_n\boldsymbol{\Upsilon}\boldsymbol{H}^{-1}\boldsymbol{v}(n) \tag{27}$$

Before investigating the convergence properties of the stochastic power control algorithm (27), we write it componentwise by using the definitions of $\boldsymbol{H}$, $\boldsymbol{\Upsilon}$ and $\boldsymbol{v}(n)$ as

$$p_i(n+1) = \left[1 - a_n(1+\gamma_i^*)\right]p_i(n) + a_n\frac{\gamma_i^*}{h_{ii}}\left(\frac{1}{L}\sum_{l=1}^{L} y_i^2(n,l)\right) \tag{28}$$

Equation (28) defines the power update rule for user i. As seen from (28), the stochastic power control algorithm is *distributed* in the sense that in order to update its power level at iteration $(n+1)$ user i needs only to know the average of the squares of its own matched filter output at its assigned base station, $1/L \sum_{l=1}^{L} y_i^2(n,l)$, and its own channel gain to its assigned base station, h_{ii}. Note that there are a total of N matched filter outputs in the system, one corresponding to each user, and a user needs to know only its own matched filter output at its assigned base station to update its power. Also note that for a single user i, there are M associated channel gains to each of M base stations, but user i needs to know only the gain to its assigned base station. The remaining three parameters of (28): the users power value in the previous iteration $p_i(n)$, its SIR target value γ_i^*, and step size sequence a_n are trivially known by the user.

Also seen from (28) is the fact that the base station of each user needs to transmit the average value of the user's matched filter outputs back to the user every L bits. Each user keeps its transmitter power level fixed until this feedback from its base station arrives and then updates its transmitter power according to (28). As we shall see, the convergence proof for (28) will be valid for any value of L. But the selection of an appropriate L will have a significant impact on the system performance. If a small L is chosen, the power control updates will be more frequent and thus the convergence will be faster. However, frequent transmission of the feedback on the downlink will effectively decrease the capacity of the system since more system resources (bandwidth) will have to be used for power control.

Since the matched filter outputs, $y_i(n,l)$'s depend on the transmitted bits and the Gaussian channel noise, the convergence of (28) will be stochastic and will be specified in terms of the mean squared error at iteration n,

$$\mathrm{MSE}(n) = E\left[\|\boldsymbol{p}(n) - \bar{\boldsymbol{p}}\|^2\right] \tag{29}$$

We will prove that under certain conditions on a_n, the sequence $\boldsymbol{p}(n)$ converges to the optimal power vector $\bar{\boldsymbol{p}}$ in the mean square sense. In particular, we will prove that

(i) If $a_n = \epsilon$ and if ϵ is chosen sufficiently small, then we will have finite lower and upper bounds on MSE as the number of iterations grows. In the limiting case as $\epsilon \to 0$ both lower and upper bounds on the limiting MSE as well as the limiting MSE itself go to zero.

(ii) If $a_n = \epsilon$ but ϵ is chosen too large, then the MSE may diverge even if the deterministic power control algorithm would converge.

(iii) If $a_n = a/n$, then the algorithm converges to the optimal power vector $\bar{\boldsymbol{p}}$ in the sense that $\lim_{n\to\infty} \mathrm{MSE}(n) = 0$, irrespective of other system parameters.

5 Stochastic Convergence Results

In this section, we will derive mean squared convergence results starting with the most general form of the stochastic power control iteration (27) (equivalent componentwise representation was given in (28)). Equation (27) can equivalently be written as

$$\boldsymbol{p}(n+1) = \boldsymbol{p}(n) - a_n\Big[\boldsymbol{p}(n) + \boldsymbol{\Upsilon}\boldsymbol{p}(n) - \boldsymbol{\Upsilon}\boldsymbol{H}^{-1}\boldsymbol{v}(n)\Big] \tag{30}$$

From (19), at time n, we have

$$\boldsymbol{v}(n) = (\boldsymbol{A} + \boldsymbol{H})\boldsymbol{p}(n) + \sigma^2\boldsymbol{u} + \boldsymbol{w}(n) \tag{31}$$

Applying (31) to (30), we obtain

$$\boldsymbol{p}(n+1) = \boldsymbol{p}(n) - a_n\Big[\boldsymbol{B}\boldsymbol{p}(n) - \sigma^2\boldsymbol{\Upsilon}\boldsymbol{H}^{-1} - \boldsymbol{\Upsilon}\boldsymbol{H}^{-1}\boldsymbol{w}(n)\Big] \tag{32}$$

where $\boldsymbol{B}$ was defined in (12). It will be mathematically convenient to define

$$\boldsymbol{\eta}(n) = \boldsymbol{\Upsilon}\boldsymbol{H}^{-1}\boldsymbol{w}(n) \tag{33}$$

Note that $\boldsymbol{\eta}(n)$ represents a normalized form of the random component of the noise contribution $\boldsymbol{w}(n)$. Also note from (20) that $E[\boldsymbol{\eta}(n)] = 0$. Inserting (33) to (32) and observing from (11) that $\boldsymbol{B}\bar{\boldsymbol{p}} = \sigma^2\boldsymbol{\Upsilon}\boldsymbol{H}^{-1}\boldsymbol{u}$, we obtain

$$\boldsymbol{p}(n+1) = \boldsymbol{p}(n) - a_n[\boldsymbol{B}\left(\boldsymbol{p}(n) - \bar{\boldsymbol{p}}\right) - \boldsymbol{\eta}(n)] \tag{34}$$

As stated in the previous section, we will prove the convergence in the mean squared sense. The norm used in (29) is the usual Euclidean norm defined as $\|\boldsymbol{x}\| = (\boldsymbol{x}^\top\boldsymbol{x})^{1/2}$. Although at the end we will prove convergence in terms of the Euclidean norm, the lack of symmetry in the system (in particular, $\boldsymbol{B}$ is not symmetric), dictates that we start our proof with a G-norm $\|\boldsymbol{x}\|_G = \left(\boldsymbol{x}^\top\boldsymbol{G}\boldsymbol{x}\right)^{1/2}$, for a specifically chosen symmetric and positive definite matrix $\boldsymbol{G}$. The necessary results about the matrices and matrix norms, including Lyapunov's result on stability of matrices and the Rayleigh quotient, are summarized in Appendix A.

In Appendix B, we prove the following lemma as a simple consequence of the Rayleigh quotient (see (69)).

Lemma 2 *If* $\lim_{n\to\infty} E\Big[\|\boldsymbol{p}(n) - \bar{\boldsymbol{p}}\|_G^2\Big] = 0$, *then* $\lim_{n\to\infty} E\Big[\|\boldsymbol{p}(n) - \bar{\boldsymbol{p}}\|^2\Big] = 0$.

Lemma 2 verifies that it is sufficient to prove convergence for a G-norm with a suitably chosen symmetric matrix $\boldsymbol{G}$. From (34), we see that convergence will depend on the properties of the matrix $\boldsymbol{B}$. In Appendix B, we verify the following result.

Lemma 3 *Matrix $-\boldsymbol{B}$ is stable iff the deterministic power control problem (6) is feasible.*

To simplify our convergence proofs, we define $\boldsymbol{x}(n) = \boldsymbol{p}(n) - \bar{\boldsymbol{p}}$ and study the convergence of $\boldsymbol{x}(n)$ to the zero vector. Subtracting $\bar{\boldsymbol{p}}$ from both sides of (34) yields

$$\boldsymbol{x}(n+1) = \boldsymbol{x}(n) - a_n[\boldsymbol{B}\boldsymbol{x}(n) - \boldsymbol{\eta}(n)] \tag{35}$$

Taking and squaring the G-norms of both sides of (35) we obtain,

$$\begin{aligned} \boldsymbol{x}(n+1)^\top \boldsymbol{G}\boldsymbol{x}(n+1) &= \boldsymbol{x}(n)^\top \boldsymbol{G}\boldsymbol{x}(n) - a_n \boldsymbol{x}(n)^\top \left[\boldsymbol{G}\boldsymbol{B} + \boldsymbol{B}^\top \boldsymbol{G}\right]\boldsymbol{x}(n) \\ &\quad + 2a_n \boldsymbol{x}(n)^\top \boldsymbol{G}\boldsymbol{\eta}(n) - 2a_n^2 \boldsymbol{x}(n)^\top \boldsymbol{B}^\top \boldsymbol{G}\boldsymbol{\eta}(n) \\ &\quad + a_n^2 \boldsymbol{x}(n)^\top \boldsymbol{B}^\top \boldsymbol{G}\boldsymbol{B}\boldsymbol{x}(n) + a_n^2 \boldsymbol{\eta}(n)^\top \boldsymbol{G}\boldsymbol{\eta}(n) \end{aligned} \tag{36}$$

By taking the conditional expectation of both sides of (36), conditioned on $\boldsymbol{x}(n) = \boldsymbol{x}$, and observing that $E[\boldsymbol{\eta}(n) \mid \boldsymbol{x}(n) = \boldsymbol{x}] = \boldsymbol{0}$, we obtain

$$\begin{aligned} E\left[\|\boldsymbol{x}(n+1)\|_G^2 \mid \boldsymbol{x}(n) = \boldsymbol{x}\right] &= \|\boldsymbol{x}\|_G^2 - a_n \boldsymbol{x}^\top \left[\boldsymbol{G}\boldsymbol{B} + \boldsymbol{B}^\top \boldsymbol{G}\right]\boldsymbol{x} + a_n^2 \boldsymbol{x}^\top \boldsymbol{B}^\top \boldsymbol{G}\boldsymbol{B}\boldsymbol{x} \\ &\quad + a_n^2 E\left[\|\boldsymbol{\eta}(n)\|_G^2 \mid \boldsymbol{x}(n) = \boldsymbol{x}\right] \end{aligned} \tag{37}$$

Our proof of convergence will proceed by bounding the individual terms on the right hand side of (37). As proven in Lemma 3, the feasibility of the deterministic power control problem guarantees the stability of $-\boldsymbol{B}$. By Theorem 3 in Appendix A, stability of matrix $-\boldsymbol{B}$, in turn, guarantees the existence of a symmetric, positive definite matrix $\boldsymbol{G}$ as a solution of

$$\boldsymbol{G}\boldsymbol{B} + \boldsymbol{B}^\top \boldsymbol{G} = \boldsymbol{C} \tag{38}$$

for any symmetric and positive definite $\boldsymbol{C}$. Therefore, selection of the G-norm allows us to develop the following bounds for the second term in (37) by using Lemma 6 in Appendix A and the fact that both $\boldsymbol{G}\boldsymbol{B} + \boldsymbol{B}^\top \boldsymbol{G}$ and $\boldsymbol{B}^\top \boldsymbol{G}\boldsymbol{B}$ are symmetric, positive definite matrices.

Lemma 4 *There exist positive constants k_0, k_0', $0 < k_0 \le k_0' < \infty$, and k_1, k_1', $0 < k_1 \le k_1' < \infty$, such that*

$$k_0 \|\boldsymbol{x}\|_G^2 \le \boldsymbol{x}^\top \left[\boldsymbol{G}\boldsymbol{B} + \boldsymbol{B}^\top \boldsymbol{G}\right]\boldsymbol{x} \le k_0' \|\boldsymbol{x}\|_G^2 \tag{39}$$

$$k_1 \|\boldsymbol{x}\|_G^2 \le \boldsymbol{x}^\top \boldsymbol{B}^\top \boldsymbol{G}\boldsymbol{B}\boldsymbol{x} \le k_1' \|\boldsymbol{x}\|_G^2 \tag{40}$$

For our convergence proof, we need k_0 and k_0' to be non-negative and if we used the usual Euclidean norm by choosing $\boldsymbol{G} = \boldsymbol{I}$, Lemma 4 would not hold since in general $\boldsymbol{B} + \boldsymbol{B}^\top$ is not a positive definite matrix, even though the real parts of the eigenvalues of $\boldsymbol{B}$ are guaranteed to be positive by the feasibility of the power control problem.

Our next lemma is more difficult to prove because of the indirect way in which the power vector $\boldsymbol{p}(n)$ interacts with the noise vector $\boldsymbol{\eta}(n)$. The proof can be found in Appendix B.

Lemma 5 *There exist positive constants d_0, c_0 and c_1, $0 < d_0 \le c_0 < \infty$, such that*

$$d_0 \le E\Big[\|\boldsymbol{\eta}(n)\|_G^2 \mid \boldsymbol{x}(n) = \boldsymbol{x}\Big] \le c_0 + c_1 \|\boldsymbol{x}\|_G^2 \tag{41}$$

Using Lemmas 4 and 5, we obtain the following upper and lower bounds from (37)

$$E\Big[\|\boldsymbol{x}(n+1)\|_G^2 \mid \boldsymbol{x}(n) = \boldsymbol{x}\Big] \le \|\boldsymbol{x}\|_G^2 - a_n k_0 \|\boldsymbol{x}\|_G^2 + a_n^2\Big\{k_1' \|\boldsymbol{x}\|_G^2 + c_0 + c_1 \|\boldsymbol{x}\|_G^2\Big\} \tag{42}$$

$$E\Big[\|\boldsymbol{x}(n+1)\|_G^2 \mid \boldsymbol{x}(n) = \boldsymbol{x}\Big] \ge \|\boldsymbol{x}\|_G^2 - a_n k_0' \|\boldsymbol{x}\|_G^2 + a_n^2\Big\{k_1 \|\boldsymbol{x}\|_G^2 + d_0\Big\} \tag{43}$$

Taking the expectation of both sides of the final inequalities (42) and (43) with respect to $\boldsymbol{x}(n)$ and letting $b_n = E\Big[\|\boldsymbol{x}(n)\|_G^2\Big]$ we obtain,

$$b_{n+1} \le \Big\{1 - k_0 a_n + (k_1' + c_1)a_n^2\Big\} b_n + a_n^2 c_0 \tag{44}$$

$$b_{n+1} \ge \Big\{1 - k_0'\epsilon + k_1 a_n^2\Big\} b_n + a_n^2 d_0 \tag{45}$$

Note that b_n is the MSE of the power vector at iteration n from the optimal power vector $\bar{\boldsymbol{p}}$. In the following two subsections we will derive the convergence results for two cases: (i) the constant coefficient sequence, $a_n = \epsilon$ and (ii) the iteration index dependent coefficient sequence, a_n which depends on n. In both cases we will start the convergence proof by the bounds on MSE given in (44) and (45).

5.1 Convergence Results for Fixed a_n

We now consider the fixed step size stochastic iteration (25). By defining

$$\alpha_0 = 1 - k_0\epsilon + (k_1' + c_1)\epsilon^2 \qquad \alpha_1 = 1 - k_0'\epsilon + k_1\epsilon^2 \tag{46}$$

we can write the lower and upper bounds (44) and (45) on the non-negative sequence b_n as,

$$\alpha_1 b_n + \epsilon^2 d_0 \le b_{n+1} \le \alpha_0 b_n + \epsilon^2 c_0 \tag{47}$$

Therefore, the non-negative sequence b_n is sandwiched between two sequences generated according to $b'_{n+1} = \alpha_0 b'_n + \epsilon^2 c_0$ and $b''_{n+1} = \alpha_1 b''_n + \epsilon^2 d_0$. Those two sequences converge to finite numbers iff ϵ is chosen such that $\mid \alpha_0 \mid < 1$ and $\mid \alpha_1 \mid < 1$.

We note that α_0 and α_1 are equal to 1 at $\epsilon = 0$. We also note that both α_0 and α_1 are locally decreasing as ϵ increases, since

$$\left.\frac{d\alpha_0}{d\epsilon}\right|_{\epsilon=0} = -k_0 < 0 \qquad \text{and} \qquad \left.\frac{d\alpha_1}{d\epsilon}\right|_{\epsilon=0} = -k'_0 < 0 \tag{48}$$

This means that we can always choose a small nonzero ϵ so that $|\alpha_0| < 1$ and $|\alpha_1| < 1$ in which case the sequences b'_n and b''_n converge and the limiting G-norm MSE, i.e., $\lim_{n\to\infty} b_n$, has finite lower and upper bounds. From the sandwich theorem, we have

$$\frac{\epsilon^2 d_0}{1-\alpha_1} \leq \lim_{n\to\infty} b_n \leq \frac{\epsilon^2 c_0}{1-\alpha_0} \tag{49}$$

We can evaluate the values of the lower and upper bounds in the extreme case when $\epsilon \to 0$ as,

$$\begin{aligned}
\lim_{\epsilon\to 0} \frac{\epsilon^2 d_0}{1-\alpha_1} &= \lim_{\epsilon\to 0} \frac{\epsilon d_0}{k'_0 - k_1\epsilon} = 0 \\
\lim_{\epsilon\to 0} \frac{\epsilon^2 c_0}{1-\alpha_0} &= \lim_{\epsilon\to 0} \frac{\epsilon c_0}{k_0 - (k'_1 + c_1)\epsilon} = 0
\end{aligned}$$

Therefore, for arbitrarily small ϵ, both lower and upper bounds for b_n approach zero, implying that the limiting G-norm MSE goes to zero as well. In this case, Lemma 2 implies that the limiting MSE goes to zero and the stochastic power control algorithm (25) converges to the unique optimal power vector $\bar{\boldsymbol{p}}$. However, as ϵ approaches zero, α_0 and α_1 approach 1 which slows the convergence rate. Thus it is undesirable to choose ϵ too small. Furthermore, it is also undesirable to choose ϵ too large. In particular, we observe that if $\epsilon = 1$ then $\alpha_1 = 1 - k'_0 + k_1$. Note that $|\alpha_1|$ may not be less than unity even if the deterministic power control problem is feasible. In this case, the lower bound derived above does not converge, and the limiting G-norm MSE and therefore limiting MSE diverge. This unfortunate situation reflects the fact that in practical systems, power control with unreliable measurements can be unstable even if the SIR targets are feasible.

In order to solve for the value of ϵ that gives rise to acceptable values of α_0 and α_1 in terms of the convergence rate and lower and upper bounds on limiting MSE, one needs to know the constants k_0, k'_0, k_1, k'_1 and c_1 which depend on the global system parameters such as the eigenvalues of matrix $\boldsymbol{\Upsilon H}^{-1}\boldsymbol{A}$. Finding these numbers is difficult because it requires the knowledge of relative delays, corresponding cross correlations and channel gains of all users, and is very computationally expensive as the number of users N increases. In order to overcome this difficulty, we propose to use a coefficient sequence a_n that is a function of n.

5.2 Convergence Results for Varying a_n

In this section, we will examine the variable step size iteration (27). We will show that if the coefficient sequence a_n satisfies the following two conditions

$$\sum_{n=1}^{\infty} a_n = \infty \qquad \text{and} \qquad \sum_{n=1}^{\infty} a_n^2 < \infty \tag{50}$$

then the power control algorithm (27) converges to the optimal power vector $\bar{\boldsymbol{p}}$ in the mean square sense. Note that $a_n = a/n$ satisfies conditions given in (50). It was shown in [19] that by that selection of a_n the convergence rate is proportional to $1/n$. In particular we will choose $a = 1$.

We will follow Sakrison's approach [19, pages 60–61] in the following derivation. We will need only the upper bound given in (44). Since a_n is a monotonically decreasing sequence, there exists n_0 and δ, $0 < \delta < 1$, such that for $n \geq n_0$,

$$\left(1 - k_0 a_n + (k_1' + c_1) a_n^2\right) \leq (1 - (1-\delta) k_0 a_n) \tag{51}$$

Furthermore, we can choose n_0 such that for $n \geq n_0$, we have $(1 - (1-\delta)k_0 a_n) > 0$. For $n \geq n_0$, the inequality (44) can be further bounded as,

$$b_{n+1} \leq \{1 - (1-\delta) k_0 a_n\} b_n + a_n^2 c_0 \tag{52}$$

Starting at $n = n_0$ and executing the recursion repeatedly yields

$$b_n \leq b_{n_0} \beta_{n_0, n-1} + c_0 \sum_{k=n_0}^{n-1} a_k^2 \beta_{k+1, n-1} \tag{53}$$

where

$$\beta_{k,n} = \begin{cases} \prod_{j=k}^{n} [1 - a_j (1-\delta) k_0] & 0 \leq k \leq n \\ 1 & k > n \end{cases} \tag{54}$$

For $0 < x < 1$, we can use the inequality $\ln(1-x) \leq -x$ to show

$$\beta_{k,n} = \exp\left(\sum_{j=k}^{n} \ln [1 - a_j (1-\delta) k_0]\right) \leq \exp\left\{-(1-\delta) k_0 \sum_{j=k}^{n} a_j\right\} \tag{55}$$

By the first condition in (50) and the fact that $(1-\delta)k_0 > 0$, the exponent in the above equation diverges to negative infinity and we have, $\lim_{n\to\infty}\beta_{k,n-1}=0$. This implies that $b_{n_0}\beta_{n_0,n-1}$ on the right hand side of (53) goes to zero as n goes to infinity. Now we will investigate the second term. Let $u(x)$ be the unit step function whose value is 1 for non-negative x and 0 otherwise. Then,

$$\lim_{n\to\infty} c_0 \sum_{k=n_0}^{n-1} a_k^2 \beta_{k+1,n-1} = c_0 \lim_{n\to\infty} \sum_{k=n_0}^{\infty} a_k^2 \beta_{k+1,n-1} u(n-1-k) \tag{56}$$

$$= c_0 \sum_{k=n_0}^{\infty} a_k^2 \lim_{n\to\infty} \beta_{k+1,n-1} u(n-1-k) = 0 \tag{57}$$

since $\lim_{n\to\infty}\beta_{k,n-1}=0$. We could exchange the limit and summation to obtain (57) from (56), because the sum on the right side of (56) is absolutely convergent. Finally, combining the result in (57) and $\lim_{n\to\infty}\beta_{k,n-1}=0$, and the fact that b_n is a non-negative sequence, we obtain $\lim_{n\to\infty} b_n = 0$. Using this result and Lemma 2, we conclude that the algorithm converges to the unique global optimal power vector in the mean squared sense, i.e., $\lim_{n\to\infty} E\left[\|\boldsymbol{p}(n)-\bar{\boldsymbol{p}}\|^2\right]=0$.

6 Discussion

Throughout this paper it is assumed that the channel gains are fixed. In general the channel gains change randomly in time as a result of log-normal or fast fading. In order to cope with the random nature of the channel gains, it was suggested in [23] to use larger SIR target values than needed allowing for a fade margin. In this paper, we only deal with difficulty of estimating the interference arising from the randomness of transmitted bits and ambient channel noise. The fact that the channel gains are changing randomly is not particular to stochastic power control but it is a problem of deterministic power control as well. In the following we will discuss the fixed channel gain case assuming that the SIR targets (γ_i^*, $i=1,\cdots,N$) are chosen properly to compensate for the fading.

The proposed algorithms need only a subset of the channel gains. Each user has M channel gains one corresponding to each one of M base stations, and only one of them, namely the channel gain of the user to its assigned base station, is needed to be used in the power update equations. The convergence results are developed with the assumption that the required channel gains, h_{ii} for $i=1,\cdots,N$, are known or estimated perfectly by the users. In this section, we will show that if the users use unbiased estimates of the random channel gains, then the proposed algorithms converge to effective target SIRs which are different than the intended ones.

Let the estimate of the channel gain used by user i in the power updates be $\hat{h}_{ii}$. Then, from equation (28) the modified power update equation for user i can be written as

$$p_i(n+1) = [1 - a_n(1+\gamma_i^*)]\, p_i(n) + a_n \frac{\gamma_i^*}{\hat{h}_{ii}} \left(\frac{1}{L} \sum_{l=1}^{L} y_i^2(n,l) \right) \tag{58}$$

Let $\hat{\boldsymbol{H}}$ denote a diagonal matrix with $\hat{h}_{ii}$ as its ith diagonal element. The modified version of (30) reflecting the estimation error in the channel gains is given as

$$\boldsymbol{p}(n+1) = \boldsymbol{p}(n) - a_n \left[\boldsymbol{p}(n) + \boldsymbol{\Upsilon} \boldsymbol{p}(n) - \boldsymbol{\Upsilon} \hat{\boldsymbol{H}}^{-1} \boldsymbol{v}(n) \right] \tag{59}$$

Note that the iteration (59) converges to a point where expected value of the term in the square brackets on the right hand side of (59) is equal to zero, since it is the fixed point of the iteration [14,19]. Calculation of this expected value requires the calculation of $E[\hat{\boldsymbol{H}}^{-1}]$. Since $f(x) = 1/x$ is a convex function for $x \geq 0$, Jensen's inequality [24] yields

$$E\left[\frac{1}{\hat{h}_{ii}} \right] \geq \frac{1}{E\left[\hat{h}_{ii} \right]} = \alpha_i \frac{1}{h_{ii}} \tag{60}$$

for some $\alpha_i \geq 1$. The equality in (60) follows from the assumption that $\hat{h}_{ii}$ is an unbiased estimate, i.e., $E[\hat{h}_{ii}] = h_{ii}$. Defining a diagonal matrix $\boldsymbol{\Lambda}$ with $\Lambda_{ii} = \alpha_i$ the expected value of $\hat{\boldsymbol{H}}^{-1}$ can be written as $E[\hat{\boldsymbol{H}}^{-1}] = \boldsymbol{\Lambda} \boldsymbol{H}^{-1}$. Therefore, from (21), the expected value of the term in the squared brackets on the right hand side of (59) is equal to

$$\boldsymbol{p} + \boldsymbol{\Upsilon} \boldsymbol{p} - \boldsymbol{\Upsilon} \boldsymbol{\Lambda} \boldsymbol{H}^{-1} \boldsymbol{A} \boldsymbol{p} - \boldsymbol{\Upsilon} \boldsymbol{\Lambda} \boldsymbol{p} - \sigma^2 \boldsymbol{\Upsilon} \boldsymbol{\Lambda} \boldsymbol{H}^{-1} \boldsymbol{u} \tag{61}$$

Equating (61) to zero yields the fixed point solution

$$\left((\boldsymbol{I} + \boldsymbol{\Upsilon} (\boldsymbol{I} - \boldsymbol{\Lambda})) - \boldsymbol{\Upsilon} \boldsymbol{\Lambda} \boldsymbol{H}^{-1} \boldsymbol{A} \right) \boldsymbol{p} = \sigma^2 \boldsymbol{\Upsilon} \boldsymbol{\Lambda} \boldsymbol{H}^{-1} \boldsymbol{u} \tag{62}$$

Defining

$$\tilde{\boldsymbol{\Upsilon}} = (\boldsymbol{I} + \boldsymbol{\Upsilon} (\boldsymbol{I} - \boldsymbol{\Lambda}))^{-1} \boldsymbol{\Upsilon} \boldsymbol{\Lambda} \tag{63}$$

equation (62) can be written equivalently as

$$\left(\boldsymbol{I} - \tilde{\boldsymbol{\Upsilon}} \boldsymbol{H}^{-1} \boldsymbol{A} \right) \boldsymbol{p} = \sigma^2 \tilde{\boldsymbol{\Upsilon}} \boldsymbol{H}^{-1} \boldsymbol{u} \tag{64}$$

Comparing (10) and (64) we observe that $\tilde{\boldsymbol{\Upsilon}}$ is the matrix of modified SIR targets. With algorithm (58), the SIR of user i converges to the ith diagonal element of $\tilde{\mathbf{\Upsilon}}$, $\tilde{\gamma_i}^*$

$$\tilde{\gamma_i}^* = \frac{\alpha_i \gamma_i^*}{1 + \gamma_i^*(1 - \alpha_i)} \tag{65}$$

instead of the originally intended SIR target γ_i^*.

As a simple example, consider the case where the channel gain estimate of user i, $\hat{h}_{ii}$, is uniformly distributed in the interval between $(1 - \xi_i)h_{ii}$ and $(1 + \xi_i)h_{ii}$. Note that $E[\hat{h}_{ii}] = h_{ii}$ and $\hat{h}_{ii}$ is an unbiased estimate, and

$$E\left[\frac{1}{\hat{h}_{ii}}\right] = \frac{1}{2\xi_i} \ln\left(\frac{1 + \xi_i}{1 - \xi_i}\right) \frac{1}{h_{ii}} \tag{66}$$

and therefore $\alpha_i = 1/(2\xi_i) \ln\left((1 + \xi_i)/(1 - \xi_i)\right)$. Note that α_i monotonically increases with ξ_i, the percentage error. Also note from (65) that $\tilde{\gamma_i}^* \geq \gamma_i^*$ for all $\alpha_i \geq 1$ and $\tilde{\gamma_i}^*$ increases with α_i. Therefore, user i aims for an effective SIR which is more than its original objective. Clearly, users get these new SIR targets if they are feasible, otherwise the powers of the users increase without bound as a sign of the infeasibility of the power control problem. Therefore, a large value of uncertainty (estimation error variance) may transform a feasible power control problem into an infeasible one.

7 Simulation Results

In our simulations, we consider a single cell of a wireless cellular system with synchronous CDMA access scheme. We assume that the mobiles are uniformly distributed on a disk of radius R_0 (chosen as $R_0 = 1000$ meters) around the base station. Channel gains (h_i's) are then obtained by $h_i = \frac{1}{r_i^\alpha}$ where r_i is the distance of the mobile i from the base station and α is the propagation constant which is taken to be 4 in the simulations. Although the convergence theorems permit individual SIR targets γ_i^* for each user i, for the simulations we use a common SIR target $\gamma_i^* = \gamma^*$ for all i. Power control iterations are always started with the zero power vector. The simulations are over $10,000$ bits. For L bit measurement averaging, the number of power control iterations is $10,000/L$.

First we investigate the performance of the stochastic power control algorithms for a situation where $N = 20$ mobiles aim for an SIR target of $\gamma^* = 4$ (≈ 6 dB). We assume a processing gain of $G = 150$ and AWGN with standard deviation $\sigma = 0.1$. Each user is assigned a random signature sequence. The normalized squared error which we define as $\text{NSE}(n) = \|\boldsymbol{p}(n) - \bar{\boldsymbol{p}}\|^2 / \|\bar{\boldsymbol{p}}\|^2$ is plotted as a function of iteration index in Figure 2. The curves of Figure 2 show the performance of the stochastic power control for $a_n = \epsilon$ (for

$\epsilon = 10^{-2}, 10^{-3}, 10^{-4}$) and for $a_n = 1/n$. Figures 3 and 4 show the same performance criteria when averaging is implemented with $L = 10$ and $L = 100$, respectively.

We observe the trade-off between the convergence rate and the value of the limiting NSE: when ϵ is large, α_0 and α_1 are smaller, and the convergence rate is fast, but the limiting NSE is larger. Therefore, we observe an initial fast decrease in the NSE, but then oscillations around the limiting NSE; see the $\epsilon = 10^{-2}$ curve in Figure 2. On the other hand, if ϵ is close to zero, then the limiting value of the NSE is smaller, but since α_0 and α_1 are close to 1, the convergence rate is very slow. In this case, we observe a slowly but steadily decreasing NSE with little oscillation; see $\epsilon = 10^{-4}$ curve in Figure 2. Also we observe from Figures 2, 3 and 4 that the performance of the stochastic power control algorithm with $a_n = \epsilon$ is almost the same as the performance of the stochastic power control algorithm which uses averaging over L bits with $a_n = L\epsilon$.

To show the convergence of the users' SIRs to the target SIR, we ran the SPCA with $a_n = 1/n$ and with $a_n = \epsilon$ for $\epsilon = 10^{-2}, 10^{-3}$ and 10^{-4} and plotted the SIRs of all users as a function of the iteration index n in Figure 5. In this experiment we chose $N = 20$, $G = 150$, $\sigma = 0.1$ and $\gamma^* = 4$ (≈ 6 dB). We observed that with $a_n = 1/n$, SIRs converge to the target SIR as expected. For fixed $a_n = \epsilon$, we observed the trade-off between the convergence rate and oscillations around the convergence point: as ϵ increases the convergence rate increases, but the SIR sequences oscillate around the target value with increasing amplitude.

Figure 6 shows the effect of estimation error in the channel gains. The channel gain estimates are uniformly distributed around the correct counterparts with $\xi_i = 0.3$ for all i. This relatively large value is chosen to create a distinguishable gap between γ^* and $\tilde{\gamma}^*$. From (66) $\alpha_i = 1.03$ for all i. With $\gamma^* = 4$, the common convergence point SIR for all users is calculated to be $\hat{\gamma}^* = 4.73$ from (65). Other parameters of this experiment are the same as the previous one, i.e., $N = 20$, $\sigma = 0.1$, $G = 150$. In Figure 6 the average SIR of all the users in the system is plotted for SPCA with $a_n = 1/n$, i.e., the graphs show $(1/N)\sum_{i=1}^{N} \mathrm{SIR}_i(n)$.

The rest of the simulations examine the SPCA with $a_n = 1/n$. The performance measure used was the fraction of users who fail to converge to within 10%, or equivalently 0.4 dB, of their SIR target over 500 realizations of the random signature sequences. For each trial, we chose a processing gain of $G = 150$, $\sigma = 0.1$, $\gamma^* = 2$ (≈ 3 dB). In Figure 7, we varied the number of users between 10 and 70. We observed that the components of the power vector with similar values converge relatively quickly, and the algorithm is slowed typically by a small number of users with dissimilar required powers. It is also noted that when the cross correlations between the signature waveforms are small, SPCA converges faster.

We implemented the SPCA with $L = 1$, 10 and 100 bits of measurement averaging for $N = 20$ and $N = 50$ users. All other parameters of the system are taken to be the same as in the previous experiment: $\gamma^* = 2$, $\sigma = 0.1$, $G = 150$. The fraction of users who fail to converge within 10% of their SIR target is plotted in Figure 8 for $N = 20$ and in Figure 9 for $N = 50$. In general, we observed that when the system is lightly loaded, M/L iterations of SPCA with averaging over L bits performs as well as M iterations of SPCA without averaging. However, when the system is highly loaded, averaging over a large number of bits L slows the convergence.

8 Conclusion

We have proposed two classes of stochastic power control algorithms, using a fixed coefficient sequence $a_n = \epsilon$ in the first class and $a_n = 1/n$ in the second class. We investigated the conditions under which we can have lower and upper bounds on the limiting value of mean squared error (MSE) for the first class of stochastic power control algorithms. We also investigated the effect of averaging on the MSE. For the second class algorithms with $a_n = 1/n$ coefficient sequence, we showed that the limiting value of MSE goes to zero, given that the deterministic power control problem is feasible.

The proposed algorithms are distributed in the sense that they require each user know only its own channel gain to its assigned base station and its own matched filter output at its assigned base station.

We observe that the convergence of the algorithm may seem slow compared to existing deterministic power control algorithms [25,26]. But one should note that deterministic algorithms require perfect measurements of parameters such as SIR or the interference experienced by each user. If these quantities are not readily available (which is the case in a typical application), one needs to estimate them possibly via an iterative algorithm. Therefore, when we account for the time needed for this estimation, seemingly faster algorithms may become drastically slower.

A Matrix Properties

The definition of a *stable* matrix given below can be found in [27, page 403].

Definition 1 *A real matrix $\boldsymbol{S}$ is stable iff all of its eigenvalues have negative real parts.*

We also state the classical result of Lyapunov; see [27, page 405] or [28, page 224].

Theorem 3 *A matrix $-\boldsymbol{S}$ is stable iff for any positive definite symmetric matrix $\boldsymbol{C}$ there exists a positive definite symmetric matrix $\boldsymbol{G}$ such that,*

$$\boldsymbol{S}^\top \boldsymbol{G} + \boldsymbol{G}\boldsymbol{S} = \boldsymbol{C} \tag{67}$$

The following simple theorem on vector norms can be found in [29, pages 352–354].

Theorem 4 *If $\boldsymbol{G}$ is an $N \times N$ symmetric, positive definite matrix then $\|\boldsymbol{x}\|_G = (\boldsymbol{x}^\top \boldsymbol{G} \boldsymbol{x})^{1/2}$ is a norm on $\boldsymbol{R}^N$.*

For any matrix $\boldsymbol{M}$, we will use $\lambda_{\min}^M$ and $\lambda_{\max}^M$ to denote the smallest and largest eigenvalues of $\boldsymbol{M}$. The relationship between the norm $\|\cdot\|_G$ and the Euclidean norm is clarified by use of the following result known as the Rayleigh quotient [30, page 349].

Theorem 5 *For any vector $\boldsymbol{x}$ and symmetric matrix $\boldsymbol{M}$,*

$$\lambda_{\min}^M \|\boldsymbol{x}\|^2 \leq \boldsymbol{x}^\top \boldsymbol{M} \boldsymbol{x} \leq \lambda_{\max}^M \|\boldsymbol{x}\|^2 \tag{68}$$

In terms of the eigenvalues of the symmetric, positive definite matrix $\boldsymbol{G}$, the Rayleigh quotient states that

$$\lambda_{\min}^G \|\boldsymbol{x}\|^2 \leq \|\boldsymbol{x}\|_G^2 \leq \lambda_{\max}^G \|\boldsymbol{x}\|^2 \qquad \text{and} \qquad \frac{1}{\lambda_{\max}^G} \|\boldsymbol{x}\|_G^2 \leq \|\boldsymbol{x}\|^2 \leq \frac{1}{\lambda_{\min}^G} \|\boldsymbol{x}\|_G^2 \tag{69}$$

We combine (69) and the Rayleigh quotient in the following useful results.

Lemma 6 *For any symmetric positive definite matrix $\boldsymbol{G}$ and symmetric matrix $\boldsymbol{M}$*

$$\frac{\lambda_{\min}^M}{\lambda_{\max}^G} \|\boldsymbol{x}\|_G^2 \leq \boldsymbol{x}^\top \boldsymbol{M} \boldsymbol{x} \leq \frac{\lambda_{\max}^M}{\lambda_{\min}^G} \|\boldsymbol{x}\|_G^2 \tag{70}$$

Corollary 1 *For any matrix $\boldsymbol{S}$ and vector $\boldsymbol{x}$,*

$$\lambda_{\min}^{S^\top S} \|\boldsymbol{x}\|^2 \leq \|\boldsymbol{S}\boldsymbol{x}\|^2 \leq \lambda_{\max}^{S^\top S} \|\boldsymbol{x}\| \tag{71}$$

B Additional Proofs

Proof: Lemma 2 From (69), for any realization of the random variables, we have

$$\frac{1}{\lambda_{\max}^G} \|\boldsymbol{p}(n) - \bar{\boldsymbol{p}}\|_G^2 \leq \|\boldsymbol{p}(n) - \bar{\boldsymbol{p}}\|^2 \leq \frac{1}{\lambda_{\min}^G} \|\boldsymbol{p}(n) - \bar{\boldsymbol{p}}\|_G^2 \tag{72}$$

Applying the expectation operator and taking the limit as $n \to \infty$ yields

$$\frac{1}{\lambda_{\max}^G} \lim_{n\to\infty} E\left[\|\boldsymbol{p}(n) - \bar{\boldsymbol{p}}\|_G^2\right] \leq \lim_{n\to\infty} E\left[\|\boldsymbol{p}(n) - \bar{\boldsymbol{p}}\|^2\right] \leq \frac{1}{\lambda_{\min}^G} \lim_{n\to\infty} E\left[\|\boldsymbol{p}(n) - \bar{\boldsymbol{p}}\|_G^2\right] \tag{73}$$

By the hypothesis of the lemma, $\lim_{n\to\infty} E\Big[\|\boldsymbol{p}(n)-\bar{\boldsymbol{p}}\|_G^2\Big] = 0$ and our desired result follows. $\square$

Proof: Lemma 3 Using λ_i and μ_i to denote the eigenvalues of $\boldsymbol{B}$ and $\boldsymbol{\Upsilon}\boldsymbol{H}^{-1}\boldsymbol{A}$, respectively, we note that $\lambda_i = 1-\mu_i$. If the deterministic power control problem is feasible, Theorem 1 and Lemma 1 imply $|\mu_i| < 1$. Hence, $|Re\{\mu_i\}| < 1$ and the result follows. $\square$

Proof: Lemma 5 We observe that $\boldsymbol{x}(n) = \boldsymbol{x}$ iff $\boldsymbol{p}(n) = \boldsymbol{x} + \bar{\boldsymbol{p}} = \boldsymbol{p}$. As a convenience, we express our conditioning as $\boldsymbol{p}(n) = \boldsymbol{p}$. First, we prove the upper bound in Lemma 5. Since $\boldsymbol{G}$ is symmetric positive definite, $\boldsymbol{\eta}(n)\boldsymbol{G}\boldsymbol{\eta}(n) \le \lambda_{\max}^G \boldsymbol{\eta}(n)^\top \boldsymbol{\eta}(n)$ for all realizations of the random vector $\boldsymbol{\eta}(n)$. Therefore, we have

$$E\Big[\|\boldsymbol{\eta}(n)\|_G^2 \mid \boldsymbol{p}(n) = \boldsymbol{p}\Big] \le \lambda_{\max}^G E\Big[\|\boldsymbol{\eta}(n)\|^2 \mid \boldsymbol{p}(n) = \boldsymbol{p}\Big] \tag{74}$$

From (33), $\|\boldsymbol{\eta}(n)\|^2 = \left\|\boldsymbol{\Upsilon}\boldsymbol{H}^{-1}\boldsymbol{w}(n)\right\|^2$ so that applying Corollary 1 and taking the conditional expectation yields

$$E\Big[\|\boldsymbol{\eta}(n)\|^2 \mid \boldsymbol{p}(n) = \boldsymbol{p}\Big] \le \frac{(\gamma_{\max}^*)^2}{h_{\min}^2} E\Big[\|\boldsymbol{w}(n)\|^2 \mid \boldsymbol{p}(n) = \boldsymbol{p}\Big] \tag{75}$$

In deriving (75), we note that the eigenvalues of diagonal matrices $\boldsymbol{\Upsilon}$ and $\boldsymbol{H}^{-1}$ are equal to their diagonal elements. Therefore, the largest eigenvalues of $\boldsymbol{\Upsilon}$ and $\boldsymbol{H}^{-1}$ are equal to $\gamma_{\max}^*$ and $1/h_{\min}$ where $\gamma_{\max}^*$ and $h_{\min}$ are defined as $\gamma_{\max}^* = \max_i \gamma_i^*$ and $h_{\min} = \min_i h_{ii}$. Combining the results of (74) and (75) yields,

$$E\Big[\|\boldsymbol{\eta}(n)\|_G^2 \mid \boldsymbol{p}(n) = \boldsymbol{p}\Big] \le \frac{(\gamma_{\max}^*)^2 \lambda_{\max}^G}{h_{\min}^2} \sum_{i=1}^N E\Big[w_i^2(n) \mid \boldsymbol{p}(n) = \boldsymbol{p}\Big] \tag{76}$$

Using (18), $E[w_i^2(n)]$ can be expressed as

$$E\Big[w_i^2(n) \mid \boldsymbol{p}(n){=}\boldsymbol{p}\Big] {=} \frac{1}{L^2}\left(\sum_{l=1}^L E\Big[w_i^2(n,l) \mid \boldsymbol{p}(n){=}\boldsymbol{p}\Big] + \sum_{l=1}^L \sum_{k\ne l} E[w_i(n,l)w_i(n,k) \mid \boldsymbol{p}(n){=}\boldsymbol{p}]\right) \tag{77}$$

Thus, we need to evaluate $E[w_i^2(n,l) \mid \boldsymbol{p}(n){=}\boldsymbol{p}]$ and $E[w_i(n,l)w_i(n,k) \mid \boldsymbol{p}(n){=}\boldsymbol{p}]$ for $k \ne l$. This requires the computation of cross correlations between $\hat{b}_{ij}(l)$ terms. Note from the definition of $\hat{b}_{ij}(l)$ given in (4) that

$$E\Big[\hat{b}_{ij}(l)\hat{b}_{ik}(m)\Big] = \begin{cases} 0 & j \neq k \\ 0 & j = k, |l-m| > 1 \\ \Gamma_{ij}(\bar{\Gamma}_{ij} + \tilde{\Gamma}_{ij}) & j = k, |l-m| = 1 \\ \bar{\Gamma}^2_{ij} + \Gamma^2_{ij} + \tilde{\Gamma}^2_{ij} & j = k, l = m \end{cases} \tag{78}$$

Note also that

$$E\Big[\hat{b}^4_{ij}(l)\Big] = \left(\bar{\Gamma}^2_{ij} + \Gamma^2_{ij} + \tilde{\Gamma}^2_{ij}\right)^2 + 4\Gamma^2_{ij}\left(\bar{\Gamma}^2_{ij} + \tilde{\Gamma}^2_{ij}\right) \tag{79}$$

and for $k \neq l$

$$E\Big[\hat{b}^2_{ij}(l)\hat{b}^2_{ij}(k)\Big] = \left(\bar{\Gamma}^2_{ij} + \Gamma^2_{ij} + \tilde{\Gamma}^2_{ij}\right)^2 \tag{80}$$

First we will evaluate $E[w^2_i(n,l) \mid \boldsymbol{p}(n) = \boldsymbol{p}]$. From (15),

$$E\Big[w^2_i(n,l) \mid \boldsymbol{p}(n) = \boldsymbol{p}\Big] = E\Big[v^2_i(n,l) \mid \boldsymbol{p}(n) = \boldsymbol{p}\Big] - \left(\sum_{j=1}^{N} p_j h_{ij}\left(\bar{\Gamma}^2_{ij} + \Gamma^2_{ij} + \tilde{\Gamma}^2_{ij}\right) + \sigma^2\right)^2 \tag{81}$$

To find $E[v^2_i(n,l) \mid \boldsymbol{p}(n) = \boldsymbol{p}]$, we note that

$$\begin{aligned} v^2_i(n,l) \;=\; & \underbrace{\sum_{j=1}^{N}\sum_{k=1}^{N}\sum_{s=1}^{N}\sum_{t=1}^{N} \sqrt{p_j}\sqrt{p_k}\sqrt{p_s}\sqrt{p_t}\sqrt{h_{ij}}\sqrt{h_{ik}}\sqrt{h_{is}}\sqrt{h_{it}}\hat{b}_{ij}(l)\hat{b}_{ik}(l)\hat{b}_{is}(l)\hat{b}_{it}(l)}_{T_1} \\ & \underbrace{+4\, n^2_i(l)\sum_{j=1}^{N}\sum_{k=1}^{N} \sqrt{p_j}\sqrt{p_k}\sqrt{h_{ij}}\sqrt{h_{ik}}\hat{b}_{ij}(l)\hat{b}_{ik}(l)}_{T_2} + \underbrace{n^4_i(l)}_{T_3} \\ & \underbrace{+4\, n_i(l)\sum_{j=1}^{N}\sum_{k=1}^{N}\sum_{s=1}^{N} \sqrt{p_j}\sqrt{p_k}\sqrt{p_s}\sqrt{h_{ij}}\sqrt{h_{ik}}\sqrt{h_{is}}\hat{b}_{ij}(l)\hat{b}_{ik}(l)\hat{b}_{is}(l)}_{T_4} \\ & \underbrace{+2\, n^2_i(l)\sum_{j=1}^{N}\sum_{k=1}^{N} \sqrt{p_j}\sqrt{p_k}\sqrt{h_{ij}}\sqrt{h_{ik}}\hat{b}_{ij}(l)\hat{b}_{ik}(l)}_{T_5} \underbrace{+2\, n^3_i(l)\sum_{j=1}^{N} \sqrt{p_j}\sqrt{h_{ij}}\hat{b}_{ij}(l)}_{T_6} \end{aligned} \tag{82}$$

To find the expected value of T_1, we note that the expected value of $\hat{b}_{ij}(l)\hat{b}_{ik}(l)\hat{b}_{is}(l)\hat{b}_{it}(l)$ is nonzero for $(j = k, s = t, j \neq s)$, $(j = s, k = t, j \neq k)$, $(j = t, k = s, j \neq k)$ and $(j = k = s = t)$. Therefore, using (78) and (79),

$$E[T_1 \mid \boldsymbol{p}(n) = \boldsymbol{p}] \;=\; 3\sum_{j=1}^{N}\sum_{k\neq j} p_j p_k h_{ij} h_{ik}(\bar{\Gamma}^2_{ij} + \Gamma^2_{ij} + \tilde{\Gamma}^2_{ij})(\bar{\Gamma}^2_{ik} + \Gamma^2_{ik} + \tilde{\Gamma}^2_{ik})$$

$$+\sum_{j=1}^{N} p_j^2 h_{ij}^2 \left((\bar{\Gamma}_{ij}^2 + \Gamma_{ij}^2 + \tilde{\Gamma}_{ij}^2)^2 + 4\Gamma_{ij}^2 (\bar{\Gamma}_{ij}^2 + \tilde{\Gamma}_{ij}^2) \right) \tag{83}$$

From (78) and the fact that $E[n_i^2(l)] = \sigma^2$,

$$E[T_2 \mid \boldsymbol{p}(n) = \boldsymbol{p}] = E[T_5 \mid \boldsymbol{p}(n) = \boldsymbol{p}] = \sigma^2 \sum_{j=1}^{N} p_j h_{ij} (\bar{\Gamma}_{ij}^2 + \Gamma_{ij}^2 + \tilde{\Gamma}_{ij}^2) \tag{84}$$

Since $n_i(l)$ and $\hat{b}_{ij}(l)$ are independent and have zero mean

$$E[T_4 \mid \boldsymbol{p}(n) = \boldsymbol{p}] = E[T_6 \mid \boldsymbol{p}(n) = \boldsymbol{p}] = 0 \tag{85}$$

And finally

$$E[T_3 \mid \boldsymbol{p}(n) = \boldsymbol{p}] = 3\sigma^4 \tag{86}$$

We obtain $E[v_i^2(n,l) \mid \boldsymbol{p}(n) = \boldsymbol{p}]$ by combining the results of (83), (84), (85) and (86) and insert the result in (81) to get

$$\begin{aligned} E\Big[w_i^2(n,l) \mid \boldsymbol{p}(n) = \boldsymbol{p}\Big] &= 2\sum_{j=1}^{N}\sum_{k\neq j} p_j p_k h_{ij} h_{ik} (\bar{\Gamma}_{ij}^2 + \Gamma_{ij}^2 + \tilde{\Gamma}_{ij}^2)(\bar{\Gamma}_{ik}^2 + \Gamma_{ik}^2 + \tilde{\Gamma}_{ik}^2) \\ &\quad + 4\sum_{j=1}^{N} p_j^2 h_{ij}^2 \Gamma_{ij}^2 (\bar{\Gamma}_{ij}^2 + \tilde{\Gamma}_{ij}^2) + 4\sigma^2 \sum_{j=1}^{N} p_j h_{ij} (\bar{\Gamma}_{ij}^2 + \Gamma_{ij}^2 + \tilde{\Gamma}_{ij}^2) + 2\sigma^4 \end{aligned} \tag{87}$$

By similar manipulations, it can be shown that

$$E[w_i(n,l) w_i(n,l-1) \mid \boldsymbol{p}(n) = \boldsymbol{p}] = 2\sum_{j=1}^{N}\sum_{k\neq j} p_j p_k h_{ij} h_{ik} \Gamma_{ij} \Gamma_{ik} (\bar{\Gamma}_{ij} + \tilde{\Gamma}_{ij})(\bar{\Gamma}_{ik} + \tilde{\Gamma}_{ik}) \tag{88}$$

and for $|l-k| > 1$

$$E[w_i(n,l) w_i(n,k) \mid \boldsymbol{p}(n) = \boldsymbol{p}] = 0 \tag{89}$$

By using the result (89) we observe that $E[w_i^2(n) \mid \boldsymbol{p}(n) = \boldsymbol{p}]$ given in (77) simplifies to

$$E\Big[w_i^2(n) \mid \boldsymbol{p}(n) = \boldsymbol{p}\Big] = \frac{1}{L^2}\Big\{ L E\Big[w_i^2(n,l) \mid \boldsymbol{p}(n) = \boldsymbol{p}\Big] + 2(L-1) E[w_i(n,l) w_i(n,l-1) \mid \boldsymbol{p}(n) = \boldsymbol{p}] \Big\} \tag{90}$$

Inserting (87) and (88) into (90) yields

$$E\Big[w_i^2(n) \mid \boldsymbol{p}(n) = \boldsymbol{p}\Big] = \frac{1}{L^2}\Bigg\{ 2L \sum_{j=1}^{N}\sum_{k\neq j} p_j p_k h_{ij} h_{ik} (\bar{\Gamma}_{ij}^2 + \Gamma_{ij}^2 + \tilde{\Gamma}_{ij}^2)(\bar{\Gamma}_{ik}^2 + \Gamma_{ik}^2 + \tilde{\Gamma}_{ik}^2)$$

$$+4L\sum_{j=1}^{N} p_j^2 h_{ij}^2 \Gamma_{ij}^2 (\bar{\Gamma}_{ij}^2 + \tilde{\Gamma}_{ij}^2) + 4L\sigma^2 \sum_{j=1}^{N} p_j h_{ij} (\bar{\Gamma}_{ij}^2 + \Gamma_{ij}^2 + \tilde{\Gamma}_{ij}^2) + 2L\sigma^4$$
$$+4(L-1)\sum_{j=1}^{N}\sum_{k\neq j} p_j p_k h_{ij} h_{ik} \Gamma_{ij}\Gamma_{ik}(\bar{\Gamma}_{ij}+\tilde{\Gamma}_{ij})(\bar{\Gamma}_{ik}+\tilde{\Gamma}_{ik})\Bigg\} \tag{91}$$

Combining the second and the last terms in (91) yields

$$\begin{aligned} E\Big[w_i^2(n) \mid \boldsymbol{p}(n) = \boldsymbol{p}\Big] &= \frac{1}{L^2}\Bigg\{2L\sum_{j=1}^{N}\sum_{k\neq j} p_j p_k h_{ij} h_{ik} (\bar{\Gamma}_{ij}^2 + \Gamma_{ij}^2 + \tilde{\Gamma}_{ij}^2)(\bar{\Gamma}_{ik}^2 + \Gamma_{ik}^2 + \tilde{\Gamma}_{ik}^2) \\ &\quad +4\sum_{j=1}^{N} p_j^2 h_{ij}^2 \Gamma_{ij}^2 (\bar{\Gamma}_{ij}^2 + \tilde{\Gamma}_{ij}^2) + 4L\sigma^2 \sum_{j=1}^{N} p_j h_{ij} (\bar{\Gamma}_{ij}^2 + \Gamma_{ij}^2 + \tilde{\Gamma}_{ij}^2) + 2L\sigma^4 \\ &\quad +4(L-1)\left(\sum_{j=1}^{N} p_j h_{ij} \Gamma_{ij} (\bar{\Gamma}_{ij} + \tilde{\Gamma}_{ij})\right)^2\Bigg\} \end{aligned} \tag{92}$$

Now we will derive an upper bound for (92). First we will include the terms corresponding to $j = k$ in the first double summation on the right hand side of (92). Since the included terms are nonnegative, this process yields an upper bound, and the double sum becomes equal to $\left(\sum_j p_j h_{ij}(\bar{\Gamma}_{ij}^2 + \Gamma_{ij}^2 + \tilde{\Gamma}_{ij}^2)\right)^2$. In order to derive an upper bound for the last term in (92), we note that $\Gamma_{ij}(\bar{\Gamma}_{ij}+\tilde{\Gamma}_{ij}) \leq \frac{1}{2}(\bar{\Gamma}_{ij}^2+\Gamma_{ij}^2+\tilde{\Gamma}_{ij}^2)$, for all i, j from the simple inequality $ab \leq (a^2+b^2)/2$ and the fact that $\bar{\Gamma}_{ij}\tilde{\Gamma}_{ij} = 0$. Upper bounding the coefficient $4(L-1)$ of this term with $4L$ and combining this term with the first term we get $3L\left(\sum_j p_j h_{ij}(\bar{\Gamma}_{ij}^2 + \Gamma_{ij}^2 + \tilde{\Gamma}_{ij}^2)\right)^2$. Since the third term (single sum) in (92) is non-negative we can upper bound it by changing its coefficient from $4L\sigma^2$ to $6L\sigma^2$. Finally we upper bound $2L\sigma^2$ with $3L\sigma^2$ to get

$$E\Big[w_i^2(n) \mid \boldsymbol{p}(n) = \boldsymbol{p}\Big] \leq \frac{3}{L}\left(\sum_{j=1}^{N} p_j h_{ij}(\bar{\Gamma}_{ij}^2 + \Gamma_{ij}^2 + \tilde{\Gamma}_{ij}^2) + \sigma^2\right)^2 + \frac{4}{L^2}\sum_{j=1}^{N} p_j^2 h_{ij}^2 \Gamma_{ij}^2 (\bar{\Gamma}_{ij}^2 + \tilde{\Gamma}_{ij}^2) \tag{93}$$

Note that the expression in the parenthesis on the right hand side of (93) is equal to the ith element of vector $((\boldsymbol{A} + \boldsymbol{H})\boldsymbol{p} + \sigma^2\boldsymbol{u})$. Inserting (93) into (76) yields

$$E\Big[\|\boldsymbol{\eta}(n)\|_G^2 \mid \boldsymbol{p}(n) = \boldsymbol{p}\Big] \leq \frac{(\gamma_{\max}^*)^2 \lambda_{\max}^G}{h_{\min}^2 L}\left\{3\left\|(\boldsymbol{A}+\boldsymbol{H})\boldsymbol{p} + \sigma^2\boldsymbol{u}\right\|^2 + \frac{4}{L}\sum_{i=1}^{N}\sum_{j=1}^{N} p_j^2 h_{ij}^2 \Gamma_{ij}^2 (\bar{\Gamma}_{ij}^2 + \tilde{\Gamma}_{ij}^2)\right\} \tag{94}$$

Let $\boldsymbol{F}$ be a diagonal matrix with its jth diagonal element $F_{jj} = \sum_{i=1}^{N} h_{ij}^2 \Gamma_{ij}^2 (\bar{\Gamma}_{ij}^2 + \tilde{\Gamma}_{ij}^2)$. Then the last summation in (94) becomes equal to $\boldsymbol{p}^\top \boldsymbol{F}\boldsymbol{p}$. Note that $\boldsymbol{F}$ is positive definite since all $F_{jj} > 0$. Then from Theorem 5 in Appendix A, we can conclude $\boldsymbol{p}^\top\boldsymbol{F}\boldsymbol{p} \leq \lambda_{\max}^F \|\boldsymbol{p}\|^2$, where

$\lambda_{\max}^F$ is the largest eigenvalue of $\boldsymbol{F}$. Then we can further bound (94) as

$$E\Big[\|\boldsymbol{\eta}(n)\|_G^2 \mid \boldsymbol{p}(n)=\boldsymbol{p}\Big] \leq \frac{(\gamma_{\max}^*)^2\lambda_{\max}^G}{h_{\min}^2 L}\left\{3\left\|(\boldsymbol{A}+\boldsymbol{H})\boldsymbol{p}+\sigma^2\boldsymbol{u}\right\|^2+\frac{4}{L}\lambda_{\max}^F\|\boldsymbol{p}\|^2\right\} \tag{95}$$

Note that for any two vectors $\boldsymbol{a}$ and $\boldsymbol{b}$, $\|\boldsymbol{a}+\boldsymbol{b}\|^2 \leq 2(\|\boldsymbol{a}\|^2+\|\boldsymbol{b}\|^2)$. Applying this to the first term in (95) and noting that $\|\boldsymbol{u}\|^2=N$ yields

$$E\Big[\|\boldsymbol{\eta}(n)\|_G^2 \mid \boldsymbol{p}(n)=\boldsymbol{p}\Big] \leq \frac{(\gamma_{\max}^*)^2\lambda_{\max}^G}{h_{\min}^2 L}\left\{6\left(\|(\boldsymbol{A}+\boldsymbol{H})\boldsymbol{p}\|^2+\sigma^2 N\right)+\frac{4}{L}\lambda_{\max}^F\|\boldsymbol{p}\|^2\right\} \tag{96}$$

Using Corollary 1 and denoting the largest eigenvalue of $(\boldsymbol{A}+\boldsymbol{H})^\top(\boldsymbol{A}+\boldsymbol{H})$ as μ we obtain

$$\begin{aligned} E\Big[\|\boldsymbol{\eta}(n)\|_G^2 \mid \boldsymbol{p}(n)=\boldsymbol{p}\Big] &\leq \frac{2(\gamma_{\max}^*)^2\lambda_{\max}^G}{h_{\min}^2 L}\left\{3(\mu\|\boldsymbol{p}\|^2+\sigma^2 N)+\frac{2}{L}\lambda_{\max}^F\|\boldsymbol{p}\|^2\right\} \\ &= \frac{2(\gamma_{\max}^*)^2\lambda_{\max}^G}{h_{\min}^2 L}\left\{\mu_1\|\boldsymbol{p}\|^2+3\sigma^2 N\right\} \end{aligned} \tag{97}$$

where $\mu_1 = 3\mu + 2\lambda_{\max}^F/L$. Note that $\|\boldsymbol{p}\|^2 = \|\boldsymbol{p}-\bar{\boldsymbol{p}}+\bar{\boldsymbol{p}}\|^2 \leq 2(\|\boldsymbol{p}-\bar{\boldsymbol{p}}\|^2+\|\bar{\boldsymbol{p}}\|^2)$ from the inequality $\|\boldsymbol{a}+\boldsymbol{b}\|^2 \leq 2(\|\boldsymbol{a}\|^2+\|\boldsymbol{b}\|^2)$. Applying this and then (69) to (97) yields the desired result,

$$E\Big[\|\boldsymbol{\eta}(n)\|_G^2 \mid \boldsymbol{p}(n)=\boldsymbol{p}\Big] \leq c_0 + c_1\|\boldsymbol{p}-\bar{\boldsymbol{p}}\|_G^2 \tag{98}$$

Now we prove the lower bound in Lemma 5. From Equations (33) and (69), we obtain,

$$\|\boldsymbol{\eta}(n)\|_G^2 \geq \lambda_{\min}^G\left\|\boldsymbol{\Upsilon}\boldsymbol{H}^{-1}\boldsymbol{w}(n)\right\|^2 \tag{99}$$

Applying Corollary 1 and noting that $h_{\max}=\max_i h_{ii}$, we have

$$\|\boldsymbol{\eta}(n)\|_G^2 \geq \frac{\lambda_{\min}^G(\gamma_{\min}^*)^2}{h_{\max}^2}\|\boldsymbol{w}(n)\|^2 \tag{100}$$

Taking the expectation conditioned on $\boldsymbol{x}(n)=\boldsymbol{x}$, or equivalently $\boldsymbol{p}(n)=\boldsymbol{p}$, yields

$$E\Big[\|\boldsymbol{\eta}(n)\|_G^2 \mid \boldsymbol{p}(n)=\boldsymbol{p}\Big] \geq \frac{\lambda_{\min}^G(\gamma_{\min}^*)^2}{h_{\max}^2}\sum_{i=1}^N E\Big[w_i^2(n) \mid \boldsymbol{p}(n)=\boldsymbol{p}\Big] \tag{101}$$

We note from (92) that

$$E\Big[w_i^2(n) \mid \boldsymbol{p}(n)=\boldsymbol{p}\Big] \geq \frac{2\sigma^4}{L} \tag{102}$$

since all of the remaining terms are non-negative. Applying this result to (101) and expressing our conditioning yields the desired result

$$E\Big[\|\boldsymbol{\eta}(n)\|_G^2 \mid \boldsymbol{x}(n) = \boldsymbol{x}\Big] \geq \frac{2\lambda_{\min}^G(\gamma_{\min}^*)^2 N\sigma^4}{h_{\max}^2 L} = d_0 \tag{103}$$

□

References

[1] J. M. Aein. Power balancing in system employing frequency reuse. *COMSAT Technical Review*, 3(2):277–300, Fall 1973.

[2] R. W. Nettleton and H. Alavi. Power control for a spread spectrum radio system. In *IEEE Vehicular Technology Conference VTC-83*, pages 242–246, 1983.

[3] S. A. Grandhi, R. Vijayan, D. J. Goodman, and J. Zander. Centralized power control in cellular radio systems. *IEEE Transactions on Vehicular Technology*, 42(4):466–468, November 1993.

[4] J. Zander. Performance of optimum transmitter power control in cellular radio systems. *IEEE Transactions on Vehicular Technology*, 41(1):57–62, February 1992.

[5] J. Zander. Transmitter power control for co-channel interference management in cellular radio systems. In *Fourth WINLAB Workshop on Third Generation Wireless Information Networks*, pages 241–247, 1993.

[6] S. C. Chen, N. Bambos, and G. J. Pottie. On distributed power control for radio networks. In *Proceedings of the International Conference on Communications ICC'94*, pages 1281–1285, May 1994.

[7] G. J. Foschini and Z. Miljanic. A simple distributed autonomous power control algorithm and its convergence. *IEEE Transactions on Vehicular Technology*, 42(4):641–646, November 1993.

[8] D. Mitra. An asynchronous distributed algorithm for power control in cellular radio systems. In *Fourth WINLAB Workshop on Third Generation Wireless Information Networks*, pages 249–259, 1993.

[9] R. D. Yates. A framework for uplink power control in cellular radio systems. *IEEE Journal on Selected Areas in Communications*, 13(7):1341–1347, September 1995.

[10] R. D. Yates and C. Y. Huang. Integrated power control and base station assignment. *IEEE Transactions on Vehicular Technology*, 44(3):638–644, August 1995.

[11] S. V. Hanly. An algorithm of combined cell-site selection and power control to maximize cellular spread spectrum capacity. *IEEE Journal on Selected Areas in Communications*, 13(7):1332–1340, September 1995.

[12] P. S. Kumar, R. D. Yates, and J. Holtzman. Power control based on bit error rate (BER) measurements. In *IEEE Military Communications Conference MILCOM'95*, pages 617–620, 1995.

[13] P. S. Kumar and J. Holtzman. Power control for a spread spectrum system with multiuser receivers. In *IEEE Personal Indoor and Mobile Radio Communications, PIMRC'95*, pages 955–959, 1995.

[14] H. Robbins and S. Monro. A stochastic approximation method. *Annals of Mathematical Statistics*, 22:400–407, 1951.

[15] J. R. Blum. Multidimentional stochastic approximation methods. *Annals of Mathematical Statistics*, 25:737–744, 1954.

[16] L. Gyorfi. Adaptive linear procedures under general conditions. *IEEE Transactions on Information Theory*, IT-30(2):262–267, March 1984.

[17] L. Gyorfi. Stochastic approximation from ergodic sample for linear regression. *Zeitschrift fur Wahrscheinlichkeitstheorie*, 54:47–55, 1980.

[18] L. Ljung. Strong convergence of a stochastic approximation algorithm. *Annals of Statistics*, 6(3):680–696, 1978.

[19] D. J. Sakrison. Stochastic approximation : A recursive method for solving regression problems. *Advances in Communication Systems 2*, pages 51–106, 1966. A. V. Balakrishnan, Ed.

[20] L. Schmetterer. Multidimensional stochastic approximation. *Multivariate Analysis 2, Proceedings of 2nd International Symp.*, pages 443–460, 1969. P.R. Krishnaiah, Ed.

[21] L. C. Yun and D. G. Messeschmitt. Power control for variable QOS on a CDMA channel. In *IEEE Military Communications Conference MILCOM-94*, pages 178–182, October 1994.

[22] E. Seneta. *Non-negative Matrices and Markov Chains*. Springer Verlag, 1981. 2nd edition.

[23] S. V. Hanly. Capacity and power control in spread spectrum macrodiversity radio networks. *IEEE Transactions on Communications*, 44(2):247–256, February 1996.

[24] T. M. Cover and J. A. Thomas. *Elements of Information Theory*. Wiley Interscience, 1991.

[25] J. Zander. Distributed cochannel interference control in cellular radio systems. *IEEE Transactions on Vehicular Technology*, 41(3):305–311, August 1992.

[26] C. Y. Huang and R. D. Yates. Rate of convergence for minimum power assignment algorithms in cellular radio systems. *ACM Wireless Networks*. to be published.

[27] D. D. Siljak. *Large-Scale Dynamic Systems : Stability and Structure*. North-Holland, 1978.

[28] F. R. Gantmacher. *Applications of the Theory of Matrices*. Interscience Publishers, 1959.

[29] P. Lancaster and M. Tismenetsky. *The Theory of Matrices*. Academic Press, 1985. Second Edition.

[30] G. Strang. *Linear Algebra and Its Applications*. Saunders College Publishing, 1988. 3rd edition.

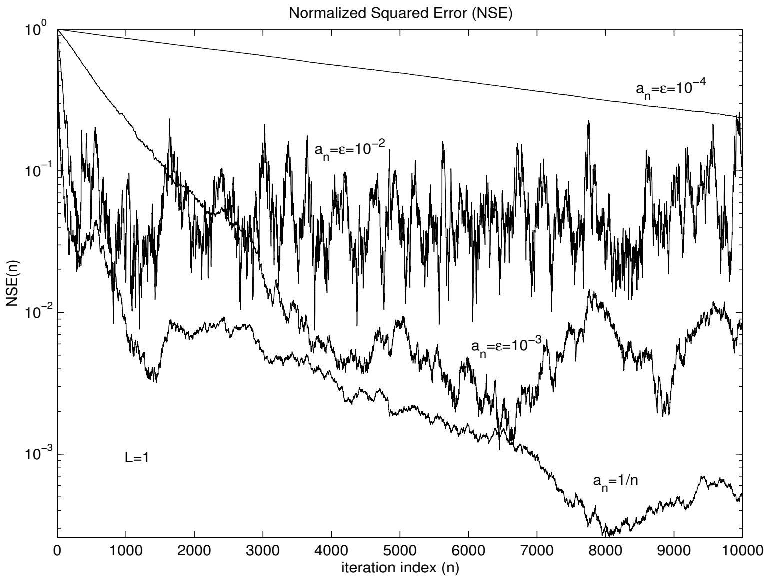

Figure 2: NSE as a function of n for stochastic power control algorithms with $a_n = 1/n$ and $a_n = \epsilon$ for $\epsilon = 10^{-2}, 10^{-3}, 10^{-4}$. No averaging is used. $(L = 1)$

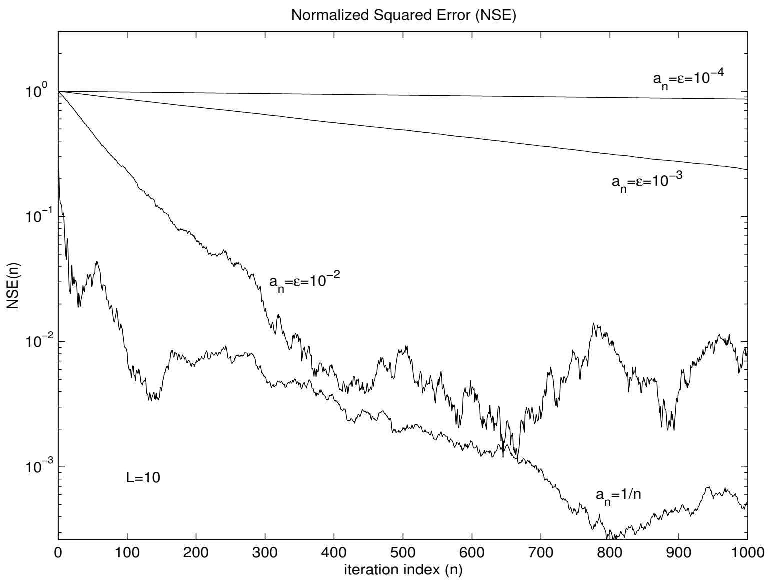

Figure 3: NSE as a function of n for stochastic power control algorithms with $a_n = 1/n$ and $a_n = \epsilon$ for $\epsilon = 10^2, 10^{-3}, 10^{-4}$. Averaging over $L = 10$ bits is implemented.

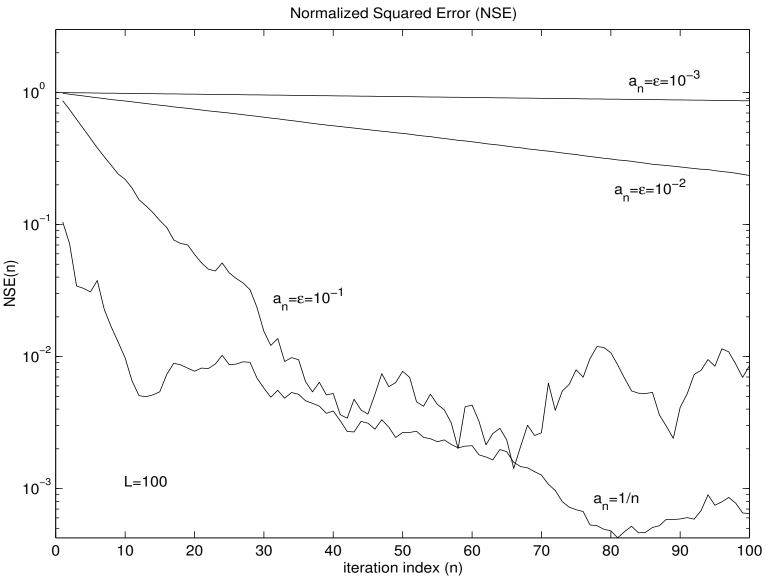

Figure 4: NSE as a function of n for stochastic power control algorithms with $a_n = 1/n$ and $a_n = \epsilon$ for $\epsilon = 10^{-1}, 10^{-2}, 10^{-3}$. Averaging over $L = 100$ bits is implemented.

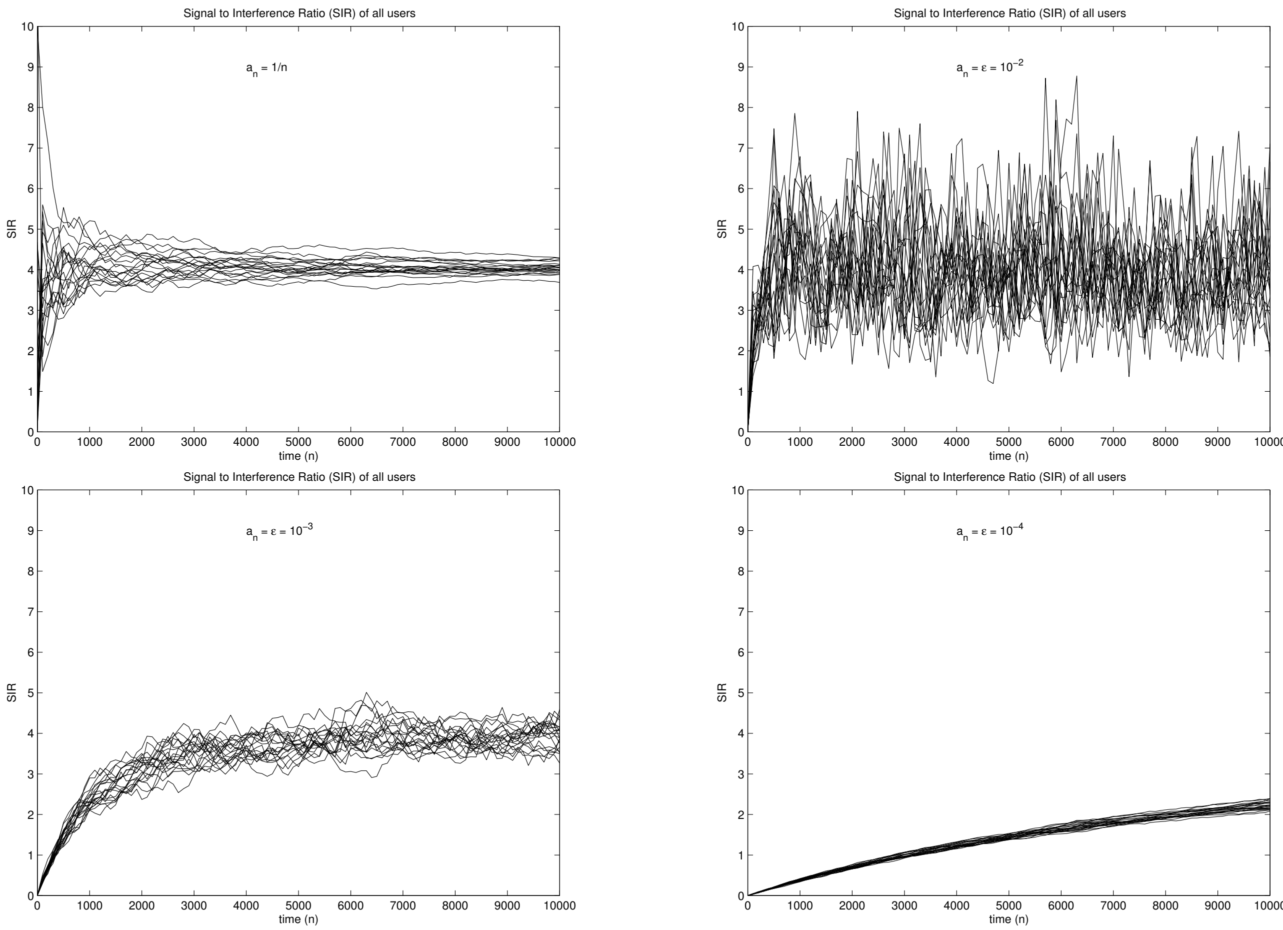

Figure 5: Signal to Interference Ratio (SIR) of all the users as a function of n for stochastic power control algorithms with $a_n = 1/n$ and $a_n = \epsilon$ with $\epsilon = 10^{-2}, 10^{-3}, 10^{-4}$. SIR target value $\gamma^* = 4$ (≈ 6 dB), number of users $N = 20$.

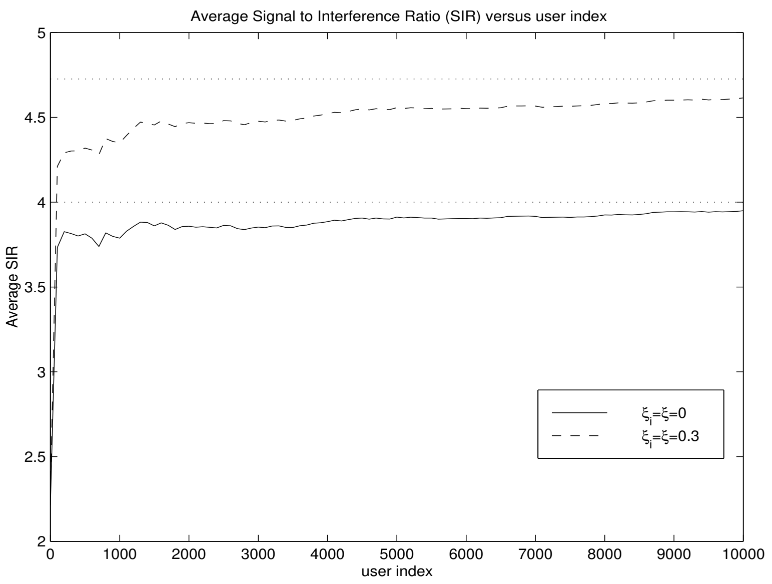

Figure 6: Average SIR of all users as a function of iteration index for the case with ($\xi_i = \xi = 0.3$) and without ($\xi_i = \xi = 0$) channel estimation error. Stochastic power control algorithm with $a_n = 1/n$ is used for $N = 20$ and $\gamma^* = 4$. Converging point SIR $\tilde{\gamma}^* = 4.73$. Horizontal dotted lines show the levels γ^* and $\tilde{\gamma}^*$.

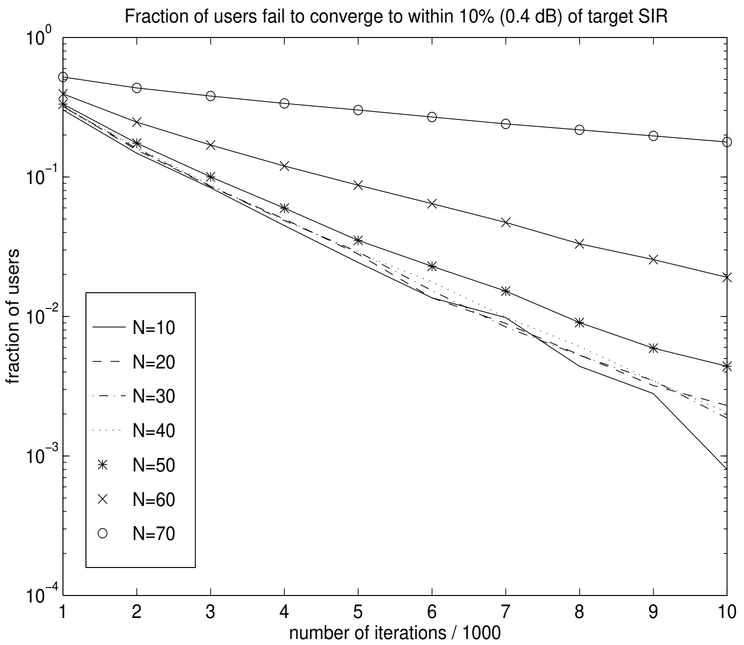

Figure 7: Fraction of users fail to converge to within 10% (0.4 dB) of the target SIR versus number of iterations divided by 1000. SPCA with $a_n = 1/n$, no averaging ($L = 1$). $N =$ number of users.

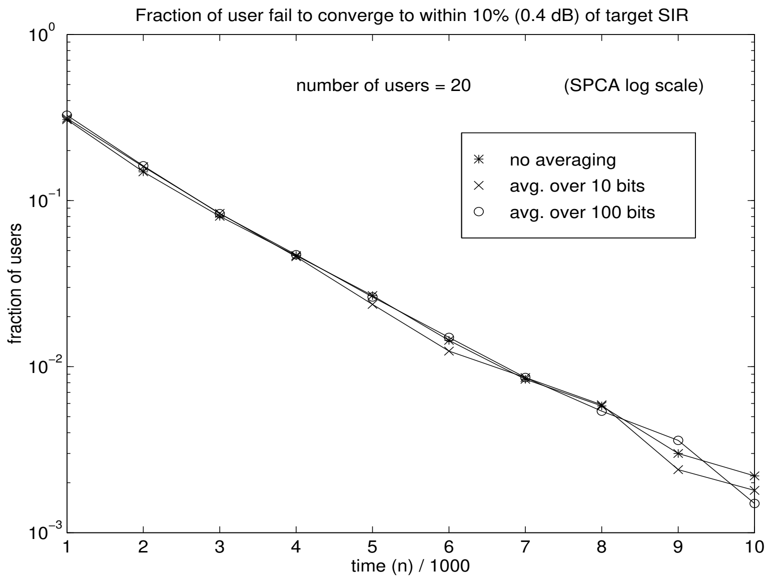

Figure 8: Fraction of users fail to converge to within 10% (0.4 dB)of the SIR target versus time (n) / 1000. Number of users in the system is 20. Different curves correspond to $L = 1, 10, 100$ where L is the number of bits over which the observations are averaged. (SPCA with $a_n = 1/n$)

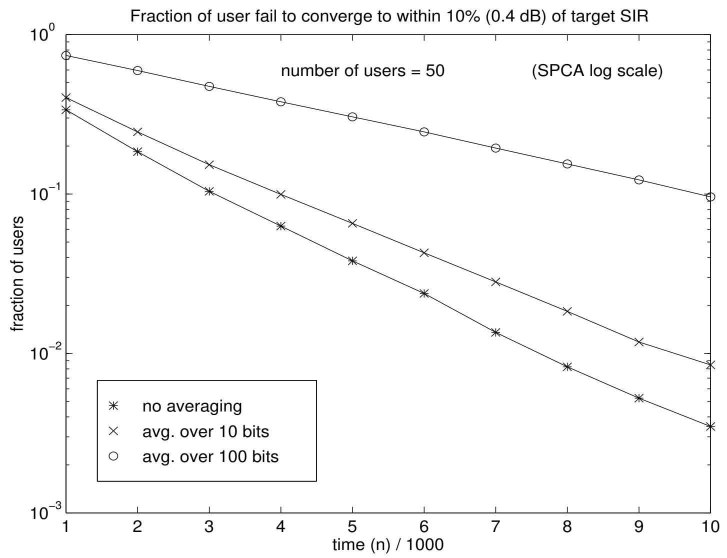

Figure 9: Fraction of users fail to converge to within 10% (0.4 dB)of the SIR target versus time (n) / 1000. Number of users in the system is 50. Different curves correspond to $L = 1, 10, 100$ where L is the number of bits over which the observations are averaged. (SPCA with $a_n = 1/n$)