# POWER CONTROL, MULTIUSER DETECTION AND INTERFERENCE AVOIDANCE IN CDMA SYSTEMS

**BY ŞENNUR ULUKUŞ**

**A dissertation submitted to the**

**Graduate School—New Brunswick**

**Rutgers, The State University of New Jersey**

**in partial fulfillment of the requirements**

**for the degree of**

**Doctor of Philosophy**

**Graduate Program in Electrical and Computer Engineering**

**Written under the direction of**

**Professor Roy D. Yates**

**and approved by**

____________________________________

____________________________________

____________________________________

____________________________________

____________________________________

**New Brunswick, New Jersey**

**October, 1998**

## ABSTRACT OF THE DISSERTATION

# Power Control, Multiuser Detection and Interference Avoidance in CDMA Systems

**by Şennur Ulukuş**

**Dissertation Director: Professor Roy D. Yates**

In CDMA systems, users create interference to each other to the extent of the correlation between their codes and the ratio of their received powers. Consequently, near-far effects in which strong users with higher received powers can overwhelm the communication quality of the weak users with lower received powers must be mitigated. This thesis presents techniques that lessen the effects of near-far problem through power control, multiuser detection, and interference avoidance.

For general signature sequence sets, the optimum multiuser detector has exponential computational complexity in number of users. This thesis identifies a class of signature sequence sets for which the optimum multiuser detector has a polynomial complexity. Because of the exponential complexity of optimum multiuser detector, many suboptimum multiuser detectors have been developed. Among them, the decorrelating detector is a linear suboptimum detector which suppresses the multiaccess interference totally. This thesis presents a blind adaptive implementation for the decorrelating detector.

Power control is a resource allocation technique that balances the received powers of the users so that no one user creates excessive interference. Deterministic power control, which assumes the availability of perfect measurements of parameters such as signal to interference ratio, received interference power, or bit error rate, has been well studied;

however, almost no effort has been made to analyze practical power control algorithms that must extract measurements from the received signal. This thesis analyzes the convergence properties of a class of stochastic power control algorithms that use readily available random observations.

Even when perfect power control is applied, the interference can be further suppressed by multiuser detection. Similarly, even when optimal linear multiuser receivers are used to maximize the signal to interference ratio, power control can be used to further diminish the near-far effect. This thesis introduces a combined power control and multiuser detection algorithm in which communication between the users and the base stations are optimized by controlling both the linear receiver filters and the power levels of the users.

This thesis introduces an iterative and distributed interference avoidance algorithm that operates by updating the signature sequences of the users. When users perform signature updates sequentially, the signature sequence set as a whole converges to an optimum set which maximizes the capacity of a single cell CDMA system.

## Acknowledgements

First, and foremost, I would like to express my gratitude to my advisor Professor Roy D. Yates for his unlimited guidance, support and encouragement throughout my Ph.D. studies at Rutgers. Working with him has been a great learning experience for me. I have greatly learned from his strong knowledge and exceptional insight on the technical subject area, as well as from his style of conducting research. I am deeply grateful to Professor Yates for his constant and generous advising, help and encouragement.

I would like to thank Professors David J. Goodman, Narayan B. Mandayam and Christopher Rose for reading my thesis, participating in my dissertation committee, and for their valuable comments and suggestions. I would like to thank Professor Sergio Verdú of Princeton University for taking time from his busy schedule to read my thesis and participate in my dissertation committee as an outside member, and for his useful comments and suggestions. I would like to thank Professor Jack M. Holtzman for his suggestions and encouragement in early stages of this thesis. I would like to thank Professor Goodman for giving me the opportunity to be a part of WINLAB.

I would like to thank all of my friends at WINLAB for making my stay at WINLAB a very enjoyable one and for many useful technical discussions. I would like to specifically acknowledge Aylin Yener and Mohammad Saquib in that respect. Aylin has contributed to my life at Rutgers in many ways. She has been a close friend who has supported and encouraged me in all aspects of my life including my Ph.D. studies, a valuable colleague who contributed to this thesis through many technical discussions and suggestions, and an excellent roommate.

I would like to thank my mother and sisters Ümran, Aynur and Gülnur for their endless support in all phases of my life including the time I have spent producing this thesis.

# Dedication

I dedicate this thesis to my mother Sevim Ulukuş.

# Table of Contents


**Abstract** . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . ii

**Acknowledgements** . . . . . . . . . . . . . . . . . . . . . . . . . . . . . iv

**Dedication** . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . v

**List of Figures** . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . ix

**1. Introduction** . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 1

1.1. Organization of the Thesis . . . . . . . . . . . . . . . . . . . . . . . . 9

**2. Tractable Cases of Optimum Multiuser Detection** . . . . . . . . . . . 11

2.1. System Model . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 11

2.2. The Optimum Multiuser Detection Problem . . . . . . . . . . . . . . . . 12

2.3. Network Preliminaries . . . . . . . . . . . . . . . . . . . . . . . . . . . 12

2.4. Optimum Multiuser Detection Problem as a Minimum Cut . . . . . . . 13

2.4.1. A Simple Example: The Two-User Case . . . . . . . . . . . . . . 14

2.5. Optimum Multiuser Detectors with Polynomial Complexity . . . . . . . 16

2.6. Chapter Summary and Conclusions . . . . . . . . . . . . . . . . . . . . 18

**3. A Blind Adaptive Decorrelating Detector for CDMA Systems** . . . 19

3.1. System Model . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 21

3.2. The Decorrelating Detector . . . . . . . . . . . . . . . . . . . . . . . . 22

3.3. A Blind Adaptive Decorrelating Detector (BADD) . . . . . . . . . . . . 24

3.4. BADD Convergence . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 25

3.5. The BADD in an Asynchronous CDMA System . . . . . . . . . . . . . . 29

3.6. Effects of Noise Variance Estimation . . . . . . . . . . . . . . . . . . . 31

3.6.1. Estimating $\sigma^2$ Before Filter Updates . . . . . . . . . . . . . . . . 32

3.6.2. Estimating $\sigma^2$ at Every Filter Update . . . . . . . . . . . . . . . . 34

3.7. Simulation Results . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 35

3.8. Chapter Summary and Conclusions . . . . . . . . . . . . . . . . . . . . 39

3.A. Appendix: Additional Proofs . . . . . . . . . . . . . . . . . . . . . . . 40

**4. Stochastic Power Control for Cellular Radio Systems** . . . . . . . . . 45

4.1. System Model . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 48

4.2. Deterministic Power Control . . . . . . . . . . . . . . . . . . . . . . . 49

4.3. Stochastic Power Control . . . . . . . . . . . . . . . . . . . . . . . . . 52

4.4. Stochastic Convergence Results . . . . . . . . . . . . . . . . . . . . . 56

4.4.1. Convergence Results for Fixed $a_n$ . . . . . . . . . . . . . . . . . . 59

4.4.2. Convergence Results for Varying $a_n$ . . . . . . . . . . . . . . . . . 60

4.5. Discussion . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 61

4.6. Simulation Results . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 64

4.7. Chapter Summary and Conclusions . . . . . . . . . . . . . . . . . . . . 73

4.A. Appendix: Matrix Properties . . . . . . . . . . . . . . . . . . . . . . . 74

4.B. Appendix: Additional Proofs . . . . . . . . . . . . . . . . . . . . . . . 75

**5. Combined Power Control and Interference Suppression** . . . . . . . . 81

5.1. System Model . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 82

5.2. Problem Definition . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 83

5.3. Derivation for the Filter Coefficients . . . . . . . . . . . . . . . . . . . 84

5.4. Power Control Algorithm . . . . . . . . . . . . . . . . . . . . . . . . . 85

5.5. Implementation of the Power Control Algorithm . . . . . . . . . . . . . 87

5.6. Simulation Results . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 90

5.7. Chapter Summary and Conclusions . . . . . . . . . . . . . . . . . . . . 95

**6. Iterative Construction of Optimum Sequence Sets to Maximize the Capacity of CDMA Systems** . . . . . . . . . . . . . . . . . . . . . . . . . 97

6.1. System Model . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 98

6.2. Capacity of a Single Cell CDMA System . . . . . . . . . . . . . . . . . . 99
6.3. WBE Sequences and a Simple Observation . . . . . . . . . . . . . . . . 102
6.4. An Iterative Algorithm . . . . . . . . . . . . . . . . . . . . . . . . . . 105
6.5. Convergence of the Proposed Algorithm . . . . . . . . . . . . . . . . . . 106
6.5.1. Properties of the Fixed Point: $N \leq G$ . . . . . . . . . . . . . . . . 108
6.5.2. Properties of the Fixed point: $N > G$ . . . . . . . . . . . . . . . . 108
6.6. Optimum Orthogonal Partitioning . . . . . . . . . . . . . . . . . . . . . 110
6.7. Simulation Results . . . . . . . . . . . . . . . . . . . . . . . . . . . . 112
6.8. Chapter Summary and Conclusions . . . . . . . . . . . . . . . . . . . . . 116
6.A. Appendix: Additional Proofs . . . . . . . . . . . . . . . . . . . . . . . 117

**7. Conclusion** . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 122
7.1. Suggestions for Further Study . . . . . . . . . . . . . . . . . . . . . . 124

**References** . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 125

**Vita** . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 131

# List of Figures


2.1. Graph associated with the optimum MUD problem for $N = 2$. . . . . . 15

3.1. Asynchronous CDMA model. . . . . . . . . . . . . . . . . . . . . . . . . 31

3.2. Averaged normalized squared error (NSE) of the desired user. . . . . . . 36

3.3. Averaged SIR of the desired user. . . . . . . . . . . . . . . . . . . . . . . 37

3.4. Averaged SIR of the desired user (algorithm of [69]). . . . . . . . . . . . 37

3.5. Averaged SIR of the desired user with mismatch in noise variance. Noise variance estimate is calculated only once before the filter updates. Noise variance used in the updates: $\tilde{\sigma}^2 = \beta\sigma^2$. . . . . . . . . . . . . . . . . . 38

4.1. Asynchronous CDMA model. . . . . . . . . . . . . . . . . . . . . . . . . 49

4.2. Simulation environment for $N = 500$. Symbols $\circ$ and $\times$ denote the base stations and the users, respectively. . . . . . . . . . . . . . . . . . . . . . 64

4.3. NSE as a function of $n$ for stochastic power control algorithms with $a_n = 1/n$ and $a_n = \epsilon$ for $\epsilon = 10^{-2}, 10^{-3}, 10^{-4}$. No averaging is used, $L = 1$. 65

4.4. 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. . . . . . . . . . . . . . . . . . . . . . . . . . . . 66

4.5. 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. . . . . . . . . . . . . . . . . . . . . . . . . . . . 66

4.6. Average SIR 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, $L = 1$. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 68

4.7. Average deviation of the SIRs from the target SIR 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, $L = 1$. . . . . . . . . . . . . 68
4.8. 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 $N = 200$ and $\gamma^* = 4$ ( $\approx 6$ dB). Converging point SIR $\tilde{\gamma}^* = 4.73$ ($\approx 6.75$ dB). Horizontal dotted lines show the levels $\gamma^*$ and $\tilde{\gamma}^*$ in dB. . . . . . . . . . 69
4.9. Average deviation of SIRs from the target SIR as a function of $n$ for stochastic power control algorithms with $a_n = 1/n$ for no quantization (infinite precision) and for quantization with 2, 3 and 4 bits. Averaging over $L = 100$ is implemented. . . . . . . . . . . . . . . . . . . . . . . . . 70
4.10. Fraction of users fail to converge to within 10% (0.4 dB) of the target SIR versus number of iterations divided by 1000, for the stochastic power control algorithm with $a_n = 1/n$. No averaging is used, $L = 1$. $N =$ number of users. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 71
4.11. Fraction of users fail to converge to within 10% ( 0.4 dB ) of the SIR target versus time (n) / 1000, for the stochastic power control algorithm with $a_n = 1/n$. Number of users in the system is $N = 200$. Different curves correspond to $L = 1, 10, 100$ where $L$ is the number of bits over which the observations are averaged. . . . . . . . . . . . . . . . . . . . . 72
4.12. Fraction of users fail to converge to within 10% ( 0.4 dB ) of the SIR target versus time (n) / 1000, for the stochastic power control algorithm with $a_n = 1/n$. Number of users in the system is $N = 400$. Different curves correspond to $L = 1, 10, 100$ where $L$ is the number of bits over which the observations are averaged. . . . . . . . . . . . . . . . . . . . . 72
5.1. Simulation environment for $N = 1000$. Symbols o and x denote the base stations and the users, respectively. . . . . . . . . . . . . . . . . . . . . . 90

5.2. Total transmitter power for the conventional power control algorithm (Conv.-PC) and the MMSE power control algorithm (MMSE-PC) for $N = 250, 500$ and 1000. . . . . . . . . . . . . . . . . . . . . . . . . . . . 91
5.3. Signal to Interference Ratio (SIR) of all the users as a function of $n$ for the conventional power control algorithm. SIR target value $\gamma_i^* = 4$ ($\approx 6$ dB) for $i = 1, \ldots, N$, number of users $N = 250$. . . . . . . . . . . . . . . . 92
5.4. Signal to Interference Ratio (SIR) of all the users as a function of $n$ for the MMSE power control algorithm. SIR target value $\gamma_i^* = 4$ ($\approx 6$ dB) for $i = 1, \ldots, N$, number of users $N = 250$. . . . . . . . . . . . . . . . . 92
5.5. Signal to Interference Ratio (SIR) of all the users as a function of $n$ for the conventional power control algorithm. SIR target value $\gamma_i^* = 4$ ($\approx 6$ dB) for $i = 1, \ldots, N$, number of users $N = 500$. . . . . . . . . . . . . . . . 93
5.6. Signal to Interference Ratio (SIR) of all the users as a function of $n$ for the MMSE power control algorithm. SIR target value $\gamma_i^* = 4$ ($\approx 6$ dB) for $i = 1, \ldots, N$, number of users $N = 500$. . . . . . . . . . . . . . . . . 93
5.7. Signal to Interference Ratio (SIR) of all the users as a function of $n$ for the conventional power control algorithm. SIR target value $\gamma_i^* = 4$ ($\approx 6$ dB) for $i = 1, \ldots, N$, number of users $N = 1000$. . . . . . . . . . . . . . 94
5.8. Signal to Interference Ratio (SIR) of all the users as a function of $n$ for the MMSE power control algorithm. SIR target value $\gamma_i^* = 4$ ($\approx 6$ dB) for $i = 1, \ldots, N$, number of users $N = 1000$. . . . . . . . . . . . . . . . . 94
6.1. Minimum and maximum eigenvalues of matrix $\boldsymbol{S}^\top(n)\boldsymbol{S}(n)$. . . . . . . . 113
6.2. Total squared correlation, TSC($n$). . . . . . . . . . . . . . . . . . . . . . . 113
6.3. Minimum and maximum eigenvalues of matrix $\boldsymbol{S}(n)\boldsymbol{S}^\top(n)$. . . . . . . . 114
6.4. Total squared correlation, TSC($n$). . . . . . . . . . . . . . . . . . . . . . . 114
6.5. Minimum and maximum eigenvalues of matrix $\boldsymbol{S}(n)\boldsymbol{S}^\top(n)$. . . . . . . . 115
6.6. Total squared correlation, TSC($n$). . . . . . . . . . . . . . . . . . . . . . . 115

# Chapter 1

# Introduction

Code Division Multiple Access (CDMA) is a promising access scheme for future wireless systems because of its many advantages [9,22,38,45,66]. In contrast with Time Division (TDMA) and Frequency Division Multiple Access (FDMA) where time or frequency is partitioned among the users, in CDMA, all users transmit at the same time in the same frequency band. In CDMA systems, the performance of the individual an user degrades gracefully as the number of users increases. Since there is no need for precise time or frequency coordination between the users, users can transmit asynchronously without the need for a prior scheduling of the channel. For the same reason, the entrance of a new user to the system is simpler, making CDMA preferable for cellular systems where handovers occur. CDMA is also superior to the fixed time or frequency allocation systems in terms of intentional jamming, privacy and security and provides greater immunity to multipath. By efficient processing of multipath, voice activity, antenna sectorization, and soft handovers [15, 67, 68], the capacity of a CDMA system can be larger than that of systems using TDMA or FDMA.

In CDMA systems, users spread their signals to a larger bandwidth by modulating their information bits with high frequency codes (signature sequences). It is not easy to implement codes that are orthogonal in an asynchronous CDMA system. Since the relative delays of the users can take arbitrary values, maintaining the orthogonality of the users requires spreading codes which are orthogonal for each pair of users for any arbitrary relative delay. Thus, even with careful design of the spreading codes, users can still have a nonzero cross correlation depending on their relative delays at the receivers (base stations). This non-zero correlation between the codes of the users can cause a phenomenon called the *near-far effect* where nearby users with higher received powers degrade the performance of the faraway users with lower received powers.

Two methods have been used to overcome the near-far problem: multiuser detection and power control. In very general terms, multiuser detection is concerned with designing receiver filters to suppress the interference of users with fixed received powers, and power control is concerned with balancing the users' received powers when the users' receiver filters are simple conventional detectors.

For a single cell system with $N$ active users, the received signal at the base station is given as

$$r(t) = \sum_{i=1}^{N} \sqrt{p_i h_i} a_i s_i(t) + n(t) \tag{1.1}$$

where $p_i$, $h_i$, $a_i$ and $s_i(t)$ are the transmitter power, channel gain, information bit and signature waveform of user $i$ and $n(t)$ is the additive white Gaussian noise (AWGN) process with power spectral density $\sigma^2$. The detection problem is to determine what is transmitted by each user ($a_i, i = 1, \cdots, N$) given the received signal $r(t)$. It was shown [62] that the matched filter outputs

$$y_i = \int_0^{T_b} r(t) s_i(t) dt \qquad i = 1, \cdots, N \tag{1.2}$$

are sufficient statistics for the detection problem. Perhaps the simplest detection scheme is the single user (conventional) detection where the bit decisions are made according to

$$\hat{a}_i = \mathrm{sgn}(y_i) \qquad i = 1, \cdots, N \tag{1.3}$$

Conventional detection is optimal only if the noise and interference are white and Gaussian. The matched filter output for user $i$ has the signal and interference (total noise) components

$$y_i = \sqrt{p_i h_i} a_i + \underbrace{\sum_{j \neq i} \sqrt{p_j h_j} \Gamma_{ij} a_j}_{\text{MAI}} + \underbrace{n_i}_{\text{AWGN}} \tag{1.4}$$

where

$$\Gamma_{ij} = \int_0^{T_b} s_i(t) s_j(t) dt \tag{1.5}$$

is the cross correlation between the signatures of user $i$ and $j$ and $n_i = \int_0^{T_b} s_i(t) n(t) dt$ is a Gaussian random variable with zero mean and variance $\sigma^2$. The total interference term is composed of two components: Multi Access Interference (MAI) and AWGN. For the total interference to be white and Gaussian the MAI should be white and Gaussian as well. It was shown in [64] that this is not the case. Thus, the conventional detection is suboptimal.

In [62], optimum multiuser detection problem was studied. The optimum detection rule is based on the maximum likelihood estimation of the transmitted bits. The optimum multiuser detection problem turns out to be a vector 0–1 quadratic programming problem where the variable of the problem is the vector of transmitted bits. However, 0–1 quadratic programming problem is known to be NP-hard [14], meaning that its computational complexity increases exponentially with the dimension of the problem, which in the optimum multiuser detection case corresponds to the number of users [63]. Thus for a reasonable number of users, solving the optimum multiuser detection problem is not tractable. In Chapter 2, we investigate the tractable cases of optimum multiuser detection problem. Trivially, if the signatures of the users are orthogonal, the optimum multiuser detection problem reduces to $N$ single user detection problems and has a linear computational complexity in the number of users.

In Chapter 2, we show the equivalence of the optimum multiuser detection problem to the minimum capacity cut problem in a related network. Not surprisingly, the general minimum cut problem is NP-hard too. Under certain conditions, the minimum cut problem is solvable in polynomial time (specifically, in $O(N^3)$ time). These conditions translate to the cross correlation between the signatures of each pair of users being negative in the optimum multiuser detection problem. Thus, the conclusion of Chapter 2 is that if the signature set is designed to have all negative cross correlations, the bits of the users can be detected optimally in polynomial time. As an example, shifted versions of a basic $m$-sequence [39] can be used by users since $m$-sequences have negative autocorrelation for all shifts except zero shift.

Since the optimum multiuser detector has an exponential computational complexity in the number of active users, several suboptimum detectors have been proposed to achieve a performance comparable to that of the optimum detector while keeping the complexity low. The suboptimum detectors that have been developed to date can be classified as linear or nonlinear suboptimum multiuser detectors. In linear suboptimum detection, the matched filter outputs are passed through a linear transformation and then the sign operation is applied to obtain the final bit decisions. Examples of linear suboptimum multiuser detectors include the decorrelating detector [28, 29] and the minimum mean squared error (MMSE) detector [30, 71]. Examples of nonlinear suboptimum multiuser detection include the decision feedback detector [11], and the multistage detectors [59, 60].

The decorrelating detector yields the maximum likelihood estimate of the transmitted bits when the received powers of the users are assumed unknown. The principle of the decorrelating detector is to suppress the MAI totally at the cost of enhancing the ambient noise. The MMSE detector suppresses the combined effect of the MAI and the ambient noise; the MMSE detector is the linear filter which minimizes the mean squared error between the transmitted information bit and the output of the receiver filter. MMSE is also referred to as the optimum linear receiver in the sense of maximizing the receiver output Signal to Interference Ratio (SIR) [65].

In their original forms, the decorrelating detector [28] and the MMSE detector [71] were centralized, non-iterative and deterministic in the sense that all the signature sequences were required to be known and the detectors were obtained by one matrix inversion. In addition to cross correlations between the signature sequences, the MMSE receiver of [71] required the received amplitudes of the users as well. References [30, 34, 40] and [7] present adaptive implementations of the MMSE and the decorrelating detectors, respectively, but these implementations still require the transmission of training sequences.

Blind adaptive multiuser detectors are desirable to overcome the need for the knowledge about the parameters of the interfering users and training sequences while constructing the receiver filter of a user. Blind adaptive multiuser detectors require the same information as the near-far limited single user detector namely the user's signature sequence and perform better than the single user detector since they are near-far resistant. A blind adaptive implementation of the MMSE multiuser detector which is based on the stochastic minimization of the output energy is presented in [21]. Subspace tracking based blind MMSE and decorrelating multiuser detectors are given in [69].

In Chapter 3, we present a stochastic gradient based blind adaptive decorrelating detector. The receiver coefficients of a user are updated recursively based on the observation of the previous outputs of the filter. Since the outputs of the filter are random due to randomness of the transmitted bits and ambient noise, the filter coefficients evolve stochastically. The convergence of the filter coefficients to the decorrelating detector is proved in the Mean Squared Error (MSE) sense. The proposed algorithm is shown to converge to the one-shot decorrelating detector [29, 62] in an asynchronous CDMA channel. The proposed algorithm requires the knowledge of only two parameters for the construction of the decorrelating detector for a desired user: the desired user's signature sequence and the variance of the AWGN. The variance of the AWGN is a fixed quantity (not time varying) and can be estimated easily. In Chapter 3 we also investigate the effects of noise variance estimation errors on the convergence of the algorithm.

In equation (1.4), we observe that a user $j$ affects the quality of service of another user $i$, through the product $\sqrt{p_j h_j}\Gamma_{ij}$. The reason $\Gamma_{ij}$ appears is that we have assumed that the receiver filter of user $i$ is the matched filter, $s_i(t)$. If we allow user $i$ to have an arbitrary filter $c_i(t)$, then the contribution of user $j$ to the received signal of user $i$ becomes $\sqrt{p_j h_j}\int_0^T s_j(t)c_i(t)dt$ where the integral can be viewed as the correlation between the receiver filter of user $i$ and signature sequence of user $j$. Thus a user $j$ affects the link quality of user $i$ to the extent of the correlation of the receiver filter of user $i$ with the signature sequence of user $j$, and the ratio of the received power of user $j$ to that of user $i$. The higher the received power of user $j$ and/or the higher the cross correlation between the signature sequence of user $j$ and the receiver filter of user $i$, the more user $j$ degrades the link quality of user $i$. Linear multiuser detectors aim at decreasing the effect of interfering users on any user by choosing the receiver filter of each user in a way to suppress the interference created by the others. Power control assumes the simple single user receiver structure and balances the received powers of the users so that no one user dominates in terms of received power destroying the quality of communication of other users. Thus the rationale of power control is to assign users power levels so as to minimize the interference they create to each other while meeting quality of service objectives defined in terms of the SIR.

Earlier work on power control [1, 16, 36, 78] identified the 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. These algorithms were non-iterative, synchronous and centralized in the sense that all the power vector components were found by a single matrix inversion and the matrix contained the global system parameters such as the channel gains of all users. Due to the computational complexity and the difficulty of knowing these global parameters, distributed power control algorithms have been developed which need only local measurements [8, 12, 33, 35, 73, 77, 79].

The power control algorithms that have been analyzed in the literature to date are *deterministic* in the sense that they require the *exact* knowledge or *perfect* estimates of some deterministic quantities such as SIR, received interference power or bit error rate (BER). 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 Chapter 4, we consider power control algorithms which need only available observables. We present and study the convergence of a class of stochastic power control algorithms which are based on the observation of the matched filter outputs that are readily available in CDMA systems using conventional receivers. The algorithms are obtained by applying the method of stochastic approximations [41] to simple extensions of the deterministic power control algorithms given in [12, 73].

The power control algorithms proposed in Chapter 4 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: the output of its own receiver filter (matched filter) at its assigned base station and its own channel gain to its assigned base station. The matched filter outputs are readily available at the base stations that transmit this feedback information back to the mobiles before each power update.

However, the uplink channel gain of each user should still be estimated. This estimation can be done by assuming the equality of the uplink and downlink channel gains and by measuring the pilot tone of the base stations at the mobiles. In reality the channel gains are also time varying in addition to being unknown. In Chapter 4 we also discuss the effect of estimation errors of channel gains on the proposed stochastic power control algorithms.

In Chapter 5, we combine the power control and multiuser detection approaches to overcome the near-far effect more effectively. In particular, we propose an algorithm which controls both the transmitter powers and the receiver filters of the users. The proposed algorithm is iterative and distributed. At each iteration, the receiver filter coefficients of the users are updated to suppress the interference optimally and then the transmitter powers of the users are assigned so that each user creates the least possible interference to others while satisfying its quality of service requirement. We show that the resulting power control algorithm converges to a fixed point power vector where all the users satisfy their SIR based quality of service requirements and that the linear receiver converges to the MMSE multiuser detector. The fixed point power vector $\bar{\boldsymbol{p}}$ satisfies $\bar{\boldsymbol{p}} \leq \boldsymbol{p}'$ for any power vector $\boldsymbol{p}'$ for which there are filter coefficients that yield acceptable SIR for all users.

The algorithm presented in Chapter 5 is implicitly a two stage algorithm. Each power control iteration includes an optimization of the linear receiver filter coefficients to maximally suppress the interference and then update of the power to overcome the residual interference. The algorithm is distributed in the sense that the parameters needed to update the receiver filter and power level of a user is available at the input and output of the receiver filter of that user.

We observe from Chapter 5 that when the transmitter powers and the receiver filters of the users are chosen optimally for a given set of signature sequences, the SIR of each user depends only on the optimal power vector and the signature sequences being used. Thus the performance of any user and hence the capacity of the CDMA system can be further improved by appropriate selection of the signature sequence set. Recently, [65] identified the sets of signature sequences that maximize the capacity of a single cell synchronous CDMA system. The capacity in [65] is defined as the maximum number of admissible users given the processing gain ($G$) and the common SIR target for all users. A set of users are called admissible if there are transmitter powers and signature sequences such that all users have SIRs that are at least as large as the fixed SIR target. It was shown in [65] that if the number of users is less than the processing gain ($N \leq G$) then a set of $N$ orthogonal signature sequences, and if the number of users is larger than the processing gain ($N > G$) then a set of $N$ sequences that satisfy Welch's bound [70] on the sum of the squares of the correlations in a set of equal energy vectors with equality (named WBE sequences) should be used in order to maximize the number of admissible users. In an earlier work [42], orthogonal signature sequences for $N \leq G$ and WBE sequences for $N > G$ were found to maximize the information theoretic capacity of a single cell CDMA system with equal received powers.

Whether defined from the point of number of admissible users or from the point of information theory, orthogonal signature sequences for $N \leq G$ and WBE sequences for $N > G$ maximize the capacity of a single cell CDMA system. Usage of these sequences facilitates interference avoidance at the transmitter as opposed to interference rejection techniques at the receiver. Although these sequences are known to maximize the capacity, no simple algorithmic scheme is known for the construction of these sequences for general $N$ and $G$. For $N \leq G$, a simple Gram-Schmidt algorithm can be used to construct $N$ orthogonal signature sequences (vectors in $G$ dimensional vector space) for the users, starting with $N$ linearly independent vectors. But for $N > G$, no simple algorithmic scheme is known for constructing WBE sequences for general $N$ and $G$.

In Chapter 6, we propose an iterative and distributed algorithm for updating a set of signature sequences. We prove that the algorithm converges to a set of orthogonal vectors if $N \leq G$, and to a WBE set if $N > G$, under very mild conditions on the initial set of signature sequences with which the algorithm is started. The algorithm updates the signature sequence of one user at a time. At each update the signature sequence of one user is replaced with a generalized normalized MMSE filter for that user. We call this normalized MMSE filter a *generalized* one because the parameters such as the powers of the interfering users and the variance of the AWGN necessary for the calculation of the normalized MMSE filter for a user can be chosen arbitrarily (but the same for all users). Since the MMSE receiver can be implemented distributedly, the proposed algorithm is amenable to distributed implementation. Thus with the proposed algorithm, each user can update its own signature sequence autonomously with some local interference measurements and the set of signature sequences converges to a global optimal set.

Most of the results in this thesis have been presented previously. Chapter 2 has been published in *IEEE Communications Letters* [56]. Chapter 3 was presented in part at IEEE Global Telecommunications Conference [55] and is accepted for publication in *IEEE Journal on Selected Areas in Communications*, Special Issue on Signal Processing for Wireless Communications [51]. Chapter 4 was presented in part at the 34th Annual Allerton Conference on Communications, Control and Computing [52] and has been published in *IEEE Transactions on Communications* [57]. Chapter 5 was presented in part at Conference on Information Sciences and Systems [54] and at IEEE International Conference on Communications [53] and is accepted for publication in *Baltzer/ACM Journal of Wireless Networks*, Special Issue on Multiuser Detection in Wireless Communications [50].

## 1.1 Organization of the Thesis

Chapter 2 starts with the definition of the optimum multiuser detection problem and presents some of its tractable cases. Chapter 3 starts with an introduction to the suboptimal decorrelating detector and presents a blind adaptive implementation. Chapter 4 reviews the literature on deterministic power control and presents a stochastic power control algorithm. Chapter 5 combines the power control and multiuser detection approaches and presents an adaptive power control and interference suppression algorithm. Chapter 6 presents an iterative algorithm for the construction of optimum signature sequence sets which maximize the capacity of a single cell CDMA system. Chapter 7 presents conclusions and suggestions for further study.

Traditionally, the multiuser detection literature has assumed single cell CDMA systems in contrast to the power control literature where multicell CDMA systems have been assumed. The reason for this is that the multiuser detection works on the receiver side and is concerned with designing the receiver filters and power control works on the transmitter side and is concerned with assigning transmitter powers to the users. Chapters 2 and 3 lie purely in the area of multiuser detection and the results are given for single cell systems. Chapter 4 lies purely in the area of power control and the results are given for general multicell CDMA systems. Chapter 5 lies at the intersection of both. In Chapter 5, we present results for general multicell systems. The iterative signature sequence set construction algorithm of Chapter 6 must be generalized to a multicell system but currently the results are for a single cell system. At the beginning of each chapter, we will describe the system model assumed in that chapter. This will allow us to concentrate only on the relevant issues in each chapter and will make each chapter readable, independent of the others.

# Chapter 2

# Tractable Cases of Optimum Multiuser Detection

Since CDMA uses non-orthogonal signaling, the information bits of the users should be detected jointly. The optimum multiuser detection problem, which is based on the maximum likelihood estimate of the transmitted bits, is studied in [62] and was shown to have an exponential computational complexity in the number of users [63]. In this chapter, we will investigate tractable cases of the optimum multiuser detection problem.

## 2.1 System Model

In CDMA systems, users are assigned unique signature waveforms which they use to modulate their information bits. Let the signature sequence of the $i$th user be $s_i(t)$ for $t \in [0, T_b]$ where $T_b$ is the bit duration. The received signal for a synchronous CDMA system with BPSK modulation is given by

$$r(t) = \sum_{i=1}^{N} A_i a_i s_i(t) + n(t) \tag{2.1}$$

where $A_i$ and $a_i$ are received amplitude and the transmitted bit ($\pm 1$ equiprobably) of the $i$th user and $n(t)$ is the additive white Gaussian noise (AWGN) process with power spectral density $\sigma^2$. The received signal vector at the output of the conventional receivers is given by

$$\boldsymbol{y} = \boldsymbol{\Gamma}\boldsymbol{\Lambda}\boldsymbol{a} + \boldsymbol{n} \tag{2.2}$$

The vector $\boldsymbol{y}$ is a sufficient statistics for the multiuser detection problem. In (2.2), $\boldsymbol{\Gamma}$ is a nonnegative definite matrix where $\Gamma_{ij}$ is $\int_0^{T_b} s_i(t)s_j(t)dt$, $\boldsymbol{\Lambda}$ is a diagonal matrix containing the received amplitudes of the users with $\Lambda_{ii} = A_i$, $\boldsymbol{a}$ is the vector containing the information bits of the users and $\boldsymbol{n}$ is a Gaussian random vector with autocovariance matrix $E[\boldsymbol{n}\boldsymbol{n}^\top] = \sigma^2\boldsymbol{\Gamma}$.

## 2.2 The Optimum Multiuser Detection Problem

The aim of multiuser detection is to recover the information bits transmitted by the users in this multiaccess environment. Optimum multiuser detection [62] is based on the maximum likelihood criteria. The optimum multiuser detector chooses $\boldsymbol{a}^*$ as the transmitted bit vector if for $\boldsymbol{a} = \boldsymbol{a}^*$ the conditional probability density of $\boldsymbol{y}$ given $\boldsymbol{a}$ is maximized. Denoting the probability density function of $\boldsymbol{n}$ by $f_{\boldsymbol{n}}(\cdot)$, the optimum detection problem is given as

$$\begin{aligned} \boldsymbol{a}^* &= \arg \max_{\boldsymbol{a} \in \{-1,1\}^N} f_{\boldsymbol{n}}(\boldsymbol{y} - \boldsymbol{\Gamma \Lambda a}) \\ &= \arg \min_{\boldsymbol{a} \in \{-1,1\}^N} \boldsymbol{a}^\top \boldsymbol{R a} - 2\boldsymbol{a}^\top \boldsymbol{\Lambda y} \end{aligned} \tag{2.3}$$

where $\boldsymbol{R} = \boldsymbol{\Lambda \Gamma \Lambda}$ with $R_{ij} = A_i A_j \Gamma_{ij}$. By introducing a vector $\boldsymbol{b}$ where $\boldsymbol{b} = (\boldsymbol{a} + \boldsymbol{u})/2$ and $\boldsymbol{u}$ is an $N$ dimensional vector of all ones, $\boldsymbol{u} = [1\ 1\ 1 \ldots 1]^\top$, we can convert (2.3) to the 0–1 programming problem

$$\boldsymbol{b}^* = \arg \min_{\boldsymbol{b} \in \{0,1\}^N} \boldsymbol{b}^\top \boldsymbol{R b} - \boldsymbol{b}^\top \tilde{\boldsymbol{y}} \tag{2.4}$$

where $\tilde{\boldsymbol{y}} = \boldsymbol{R u} + \boldsymbol{\Lambda y}$. Note that the solutions of (2.3) and (2.4) are related by the one-to-one relationship $\boldsymbol{a}^* = 2\boldsymbol{b}^* - \boldsymbol{u}$.

The optimum multiuser detection problem which is written in the form of a 0–1 programming program in (2.4) is NP-hard [14] i.e., its computational complexity is believed to increase exponentially with the number of users [63].

## 2.3 Network Preliminaries

Consider a network $G = [V, A]$ with vertices $V = \{0, 1, \ldots, N+1\}$ and arcs $A$. For any two vertices $i$ and $j$ in $G$, $c_{ij}$ denotes the capacity of the arc connecting $(i,j)$. Let the nodes 0 and $N+1$ represent the source and the sink respectively. A cut separating 0 and $N+1$ is a partition of the nodes $(\mathcal{S}, \bar{\mathcal{S}})$ where $0 \in \mathcal{S}$, $N+1 \in \bar{\mathcal{S}}$, $\mathcal{S} \cup \bar{\mathcal{S}} = V$ and $\mathcal{S} \cap \bar{\mathcal{S}} = \emptyset$. The capacity of the cut $(\mathcal{S}, \bar{\mathcal{S}})$ is given by [2]

$$C(\mathcal{S}, \bar{\mathcal{S}}) = \sum_{i \in \mathcal{S}} \sum_{j \in \bar{\mathcal{S}}} c_{ij} \tag{2.5}$$

The minimum cut separating nodes 0 and $N+1$ is defined to be the cut separating nodes 0 and $N+1$ having the minimum capacity.

In [23], it was shown that any cut separating nodes 0 and $N+1$ can be represented by a vector $(1, b_1, b_2, \ldots, b_N, 0)$ where $b_i \in \{0,1\}$ for $i = 1, \ldots, N$ is an indicator for membership in $\mathcal{S}$. That is $\mathcal{S} = \{i \mid b_i = 1\}$ and $\bar{\mathcal{S}} = \{i \mid b_i = 0\}$. It was also shown in [23] that the capacity of the cut $(\mathcal{S}, \bar{\mathcal{S}})$ is given by

$$C(\boldsymbol{b}) = \sum_{i=0}^{N+1} \sum_{j=0}^{N+1} c_{ij} b_i (1 - b_j) \tag{2.6}$$

where $b_0 = 1$ and $b_{N+1} = 0$. Substituting the values of $b_0$ and $b_{N+1}$ in (2.6) and using the fact that $b_i^2 = b_i$ yields

$$C(\boldsymbol{b}) = \sum_{j=1}^{N+1} c_{oj} - \sum_{i=1}^{N} \left( c_{0i} - \sum_{j=1}^{N+1} c_{ij} \right) b_i - \sum_{i=1}^{N} \sum_{j=1}^{N} c_{ij} b_i b_j \tag{2.7}$$

Similar to [37], we show the equivalence of the optimum multiuser detection problem to a minimum cut problem in a related network. We then identify conditions under which the optimum multiuser detection problem can be solved in polynomial time.

## 2.4 Optimum Multiuser Detection Problem as a Minimum Cut

Let us define the objective function of the 0–1 programming version the optimum multiuser detection problem given in (2.4) to be

$$F(\boldsymbol{b}) = -\sum_{i=1}^{N} \tilde{y}_i b_i + \sum_{i=1}^{N} \sum_{j=1}^{N} R_{ij} b_i b_j \tag{2.8}$$

Comparing (2.7) and (2.8) we observe that if the capacities $c_{ij}$ are chosen properly, the cost function of (2.8) in $N$ variables can be represented as the capacity of a cut in an associated network with $(N+2)$ nodes. That is a network $G$ with arc capacities $c_{ij}$ satisfying:

$$\begin{array}{lll} \text{(a)} \quad c_{ij} = -R_{ij} & \text{for} \quad i,j = 1, \ldots, N \\ \text{(b)} \quad c_{0i} - \sum_{j=1}^{N+1} c_{ij} = \tilde{y}_i & \text{for} \quad i = 1, \ldots, N \\ \text{(c)} \quad \sum_{j=0}^{N+1} c_{0j} = 0 & \end{array} \tag{2.9}$$

has $C(\boldsymbol{b}) = F(\boldsymbol{b})$ for all $\boldsymbol{b}$ such that $b_i \in \{0,1\}$ for $i = 1, \ldots, N$. Given the parameters of the quadratic 0-1 programming problem $\tilde{\boldsymbol{y}}$ and $\boldsymbol{R}$, the construction of the corresponding network satisfying above three conditions is not unique. Condition (2.9)-(a) fixes the capacities of all arcs not incident to the source and the sink. Since there are more arcs than the conditions to be satisfied given in (2.9)-(b) and (2.9)-(c), some of the arc capacities can be assigned arbitrary numbers while conserving the connectivity of the overall network. One such assignment equates $c_{N+1,j}$ to zero for $j = 0, \ldots, N$ and $c_{0j}$ to arbitrary positive numbers for $j = 1, \ldots, N$. The rest of the $(N+1)$ arc capacities can be determined uniquely from conditions (2.9)-(b) and (c). From (2.9)-(b)

$$c_{i,N+1} = -\tilde{y}_i - \sum_{j=1}^{N} c_{ij} + c_{0i} = -A_i y_i + c_{0i} \tag{2.10}$$

where we used $\tilde{y}_i = \sum_{j=1}^{N} R_{ij} + A_i y_i$ from the definition of $\tilde{\boldsymbol{y}}$ and (2.9)-(a) which implies $\sum_{j=1}^{N} R_{ij} = -\sum_{j=1}^{N} c_{ij}$. The only remaining capacity, $c_{0,N+1}$, can be found from (2.9)-(c). The following rules construct the related network for the optimum multiuser detection problem:

(i) Set $c_{ij} = -R_{ij}$ for $i, j = 1, \ldots, N$

(ii) Set $c_{i0} = 0$ and assign arbitrary positive numbers to $c_{0i}$ for $i = 1, \ldots, N$

(iii) Set $c_{0,N+1} = -\sum_{j=1}^{N} c_{0j}$

(iv) Set $c_{i,N+1} = -A_i y_i + c_{0i}$ for $i = 1, \ldots, N$

In the following subsection we demonstrate the equivalence of the optimum multiuser detection problem and the minimum cut problem in the related network obtained by the rules given above for the case of two users.

### 2.4.1 A Simple Example: The Two-User Case

Consider a two user system with the cross correlation between the signature sequences of the users $\Gamma_{12} = \Gamma_{21} = -\rho$, for some $0 \leq \rho \leq 1$. Then $R_{12} = R_{21} = -A_1 A_2 \rho$. Since the optimum multiuser detection problem (2.3) is insensitive to the diagonal elements of $\boldsymbol{R}$ we take $R_{11} = R_{22} = 0$. For two users, the optimum multiuser detection problem of (2.4) is to minimize

$$F(b_1, b_2) = -2A_1 A_2 \rho b_1 b_2 + (A_1 A_2 \rho - A_1 y_1) b_1 + (A_1 A_2 \rho - A_2 y_2) b_2 \tag{2.11}$$

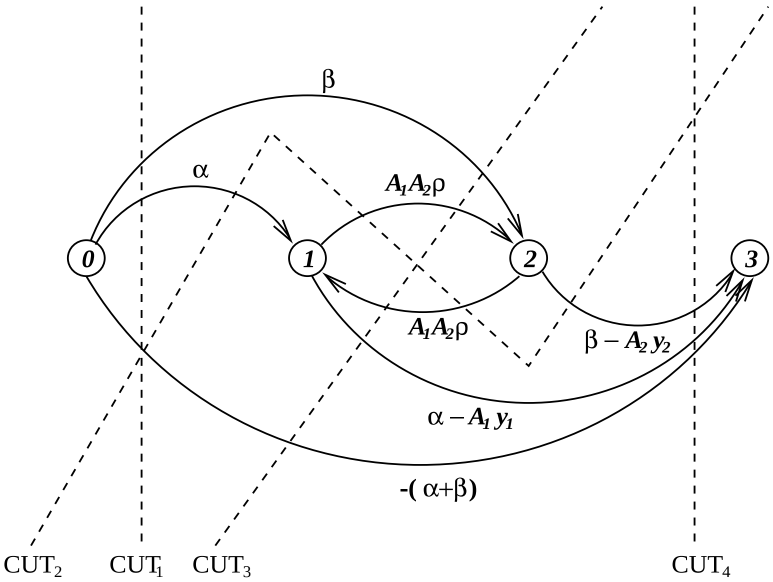


Figure 2.1: Graph associated with the optimum MUD problem for $N = 2$.

For all possible transmitted bits, the cost function $F(b_1, b_2)$ has the following values:

$$\begin{aligned} F(0,0) &= 0 \\ F(0,1) &= A_1A_2\rho - A_2y_2 \\ F(1,0) &= A_1A_2\rho - A_1y_1 \\ F(1,1) &= -A_1y_1 - A_2y_2 \end{aligned} \tag{2.12}$$

The optimum detection rule chooses that realization of $(b_1, b_2)$ for which the cost function $F(b_1, b_2)$ is minimum. Clearly, the minimum value depends on cross correlation value $(\rho)$, received amplitudes of the users $(A_1,\ A_2)$ and the matched filter outputs $(y_1,\ y_2)$. The graph corresponding to this multiuser detection problem is shown in Figure 2.1. The rules given above are used to construct the graph:

(i) $c_{12} = c_{21} = A_1A_2\rho$,

(ii) $c_{10} = c_{20} = c_{30} = 0$, $c_{01} = \alpha$, $c_{02} = \beta$,

(iii) $c_{03} = -(\alpha + \beta)$,

(iv) $c_{13} = \alpha - A_1y_1$, $c_{23} = \beta - A_2y_2$.

The capacities of the cuts corresponding to 4 possible cases for $(b_1, b_2)$, $C(b_1, b_2)$, are calculated from Figure 2.1 to be:

$$\begin{aligned}
\mathrm{CUT}_1: \quad & C(0,0) = 0 \\
\mathrm{CUT}_2: \quad & C(0,1) = A_1 A_2 \rho - A_2 y_2 \\
\mathrm{CUT}_3: \quad & C(1,0) = A_1 A_2 \rho - A_1 y_1 \\
\mathrm{CUT}_4: \quad & C(1,1) = -A_1 y_1 - A_2 y_2
\end{aligned} \tag{2.13}$$

Note the equivalence of the values of the cost function in (2.12) and the capacities of the cuts calculated above.

## 2.5 Optimum Multiuser Detectors with Polynomial Complexity

The arguments in Section 2.4 show that the optimum multiuser detection problem is equivalent to finding a minimum capacity cut between the source and the sink nodes of a related network. For the general optimum multiuser detection problem the capacities of the arcs in the related network can be both positive and negative depending on the cross correlations between the signature sequences. Unfortunately, in this general case, the minimum cut problem is an NP-hard problem as well [37]. However, in the special case when the capacities of the arcs not incident to both source and sink are nonnegative, the minimum cut problem is solvable in polynomial time. The minimum cut problem is the dual of the maximum flow problem for which there are many polynomial time algorithms [2]. For example the computational complexity of the FIFO preflow-push algorithm [2] is $O(N^3)$ where $N$ is the number of nodes of the network. This implies that the computational complexity of the optimum multiuser detection problem is $O(N^3)$ where $N$ is the number of active users, because if the dimensionality of the optimum multiuser detection problem is $N$ then the number of nodes in the associated network is $N + 2$.

From rule (i), we observe that the capacities of the arcs not incident to source and sink are guaranteed to be nonnegative only if $R_{ij} \leq 0$ for all $i \neq j$. Note that $R_{ij} \leq 0$ does not apply to the diagonal elements of $\boldsymbol{R}$, since the optimum multiuser detection problem of (2.3) is insensitive to the values of the diagonal elements of $\boldsymbol{R}$, because

$a_i^2 = 1$ for $a_i \in \{-1, 1\}$. Moreover, since $R_{ij} = A_i A_j \Gamma_{ij}$ the condition that $R_{ij} \leq 0$ for $i \neq j$ directly translates to the condition that $\Gamma_{ij} \leq 0$ for $i \neq j$. Thus the optimum multiuser detection problem has a polynomial computational complexity only if the signature sequences of the users are chosen such that all of the cross correlations are negative. This fact suggests the desirability of appropriate design of the users' signature sequence set. One such set of signature sequences is obtained when shifted versions of an *$m$-sequence* are used (see [75] for example). The autocorrelation of an $m$-sequence is equal to $-1/G$ for all nonzero shifts, where $G$ is the length of the sequence. Therefore, if shifted versions of a length $G$ m-sequence are used as the signature sequences, the cross correlation between any two users becomes $\Gamma_{ij} = -1/G$.

For an $N$ user system with all negative cross correlations, the steps of the polynomial time optimum multiuser detection algorithm are:

1. Construct the related network with $(N + 2)$ nodes,
2. Solve the minimum cut problem in this network using a polynomial time algorithm such as FIFO preflow-push [2],
3. Using the membership function, declare $\boldsymbol{b}^*$.

Note that while constructing the related network only a linear number of arc capacities (the ones that depend on $\boldsymbol{y}$) needs to be changed from bit to bit, because others depend only on $\boldsymbol{R}$ which is fixed once the signature sequences and the received amplitudes of the users are fixed.

Reference [58] gives a decision feedback detector which achieves the asymptotic efficiency of the optimum multiuser detector under certain conditions on the cross correlation matrix and users' received powers. The contribution of [58] differs from that of this work in two ways. First, the decision mechanism of [58] achieves the *asymptotic* efficiency of the optimum detector. In other words, it achieves the probability of error of the optimum detector in the low background noise region ($\sigma \to 0$). Second, the conditions of [58] are more restrictive than those given here in the sense that the conditions of [58] are defined in terms of both cross correlations and received powers of the users. The detection algorithm of this work, on the other hand, achieves the probability of error of the optimum detector irrespective of the users' powers or the level of background noise as long as the cross correlations between the signature sequences of the users are all negative.

## 2.6 Chapter Summary and Conclusions

The optimum multiuser detection rule in CDMA is based on the maximum likelihood estimate of the vector of transmitted bits. The maximum likelihood estimate is equivalent to a 0–1 quadratic programming problem which is NP-hard in general. This chapter suggests that if the signatures of the users are selected carefully, the optimum multiuser detection problem can be solved in polynomial time. In this chapter, the equivalence of the optimum multiuser detection and minimum capacity cut problems in a related network is shown. Although the minimum capacity cut problem is NP-hard too in general, it can be solved in polynomial time under certain conditions on the capacities of the links in the network. These conditions on the capacities directly translate to the optimum multiuser detection problem as a condition which requires that the cross correlation between the signatures of any two users to be negative. Thus, if the signature sequence set is chosen such that all the cross correlations between the signature sequences are negative, then the optimum multiuser detection problem can be solved in $O(N^3)$ time, where $N$ is the number of users. $M$-sequences are identified as a set of sequences which satisfy this condition.

# Chapter 3

# A Blind Adaptive Decorrelating Detector for CDMA Systems

The difficulty of bit detection in CDMA arises because of the non-orthogonality of the signature sequences of the users. Since the simple single user detector is not near-far resistant and the optimum multiuser detector is exponentially complex, many suboptimum detectors have been proposed. Among these is the decorrelating detector [28] which suppresses the multiaccess interference totally at the cost of enhancing the ambient noise. The decorrelating detector of [28, 29] is centralized and non-iterative. The construction of the decorrelating detector filter for a certain user requires the knowledge of the signature sequences of all the interfering users as well as the signature sequence of the user of interest. In addition, for an $N$ user system, the construction requires inversion of the $N \times N$ cross correlation matrix. Blind adaptive algorithms are desirable to overcome the need for the knowledge about the parameters of the interfering users and iterative algorithms are needed to avoid the matrix inversion which may have a large computational complexity. A blind adaptive algorithm based on the minimization of the output energy was given in [21]. This algorithm was shown to converge to the MMSE multiuser detector [30, 71].

In [7], an adaptive multiuser detector which converges to the decorrelating detector is proposed. This detector still needs the signature sequences of all users and the transmission of training sequences. In [69], blind algorithms based on signal subspace tracking are investigated and two algorithms which converge to the decorrelating and MMSE multiuser detectors are proposed. The blind adaptive decorrelating detector proposed in [69] needs information about the variance of the additive white Gaussian channel noise and the number of users, both of which can be estimated again using subspace tracking techniques. The computational complexity of the algorithm of [69] is $O(GN)$ per iteration, where $G$ is the processing gain and $N$ is the number of active users.

In this chapter, we present a blind adaptive multiuser detector which uses observables that are readily available at the receiver and which converges to a decorrelating detector. The detector is constructed via a distributed iterative algorithm which updates the receiver filter coefficients of a desired user by using the previous output of the filter under construction. Since the filter output is random due to the randomness of the multiaccess interference and existence of ambient Gaussian channel noise, the algorithm evolves stochastically. We prove the convergence of the filter coefficients to a decorrelating detector in the mean squared error (MSE) sense. We develop lower and upper bounds for the MSE of the filter coefficients at iteration $n+1$ in terms of the MSE of the filter coefficients at iteration $n$, and investigate the conditions under which the MSE sequence converges to zero. Two kinds of algorithms, one using a fixed step size sequence and the other using a time dependent step size sequence, are proposed. For the algorithm using a fixed step size, we obtain the lower and upper bounds for the limiting MSE as $n$ approaches infinity in terms of the step size. We show that the limiting MSE can be made arbitrarily small by choosing a small enough step size. On the other hand we directly prove that the MSE converges to zero when a particular type of time-varying step size sequence is used.

The computational complexity of the proposed algorithm is $O(G)$ per iteration. The proposed algorithm requires the knowledge of only two parameters for the construction of the filter coefficients for a desired user: the desired user's signature sequence and the variance of the additive white Gaussian noise (AWGN). The variance of the AWGN is a time invariant deterministic quantity that can be estimated easily, perhaps before the information transfer starts. In this case, samples of AWGN can be observed without any interference at the output of an arbitrary (non-zero) receiver filter and a reliable estimate of the variance of AWGN can be obtained by time-averaging the squared filter outputs. An alternative implementation is to allocate a unit energy filter that is orthogonal to all the signatures being used. In this case again the square of the output of the filter will be an unbiased estimate for the noise variance. In the latter case, since the filter is orthogonal to all the signatures, one can continue to have estimates for the noise variance while the users are transmitting and the decorrelating receiver filter coefficients are being updated. In Sections 3.1-3.5, we will assume that the noise variance is known perfectly. In Section 3.6, we investigate the effects of using estimated values of noise variance on the convergence of the proposed algorithm.

## 3.1 System Model

In this chapter, we consider the uplink of a single cell CDMA system with $N$ users. The modulation scheme is assumed to be BPSK to simplify the analysis. Until Section 3.5, where the convergence of the proposed algorithm for an asynchronous CDMA channel will be investigated, we will assume a synchronous CDMA system.

The signature waveform of user $i$, denoted by $s_i(t)$, is non-zero only in the bit interval $[0, T_b]$ and is normalized to unit energy, i.e., $\int_0^{T_b} s_i^2(t)dt = 1$. At the front end of the base station receiver filters, the baseband received signal, $r(t)$, in one bit interval is given by

$$r(t) = \sum_{i=1}^{N} \sqrt{p_i} a_i s_i(t) + n(t) \tag{3.1}$$

where $p_i$ and $a_i$ are the received power and the information bit of user $i$ ($+1$ or $-1$ with equal probability) and $n(t)$ is an additive white Gaussian noise (AWGN) process.

We define the chip waveform to be $\psi(t)$, $t \in [0, T_c]$ and 0 elsewhere, where $T_c$ is the chip duration. Thus $\{\psi(t - kT_c), k = 0, \ldots, G-1\}$, where $G = T_b/T_c$ is the processing gain, is a basis for the signal space. This allows us to represent both the signature sequences and the receiver filters of the users with $G$ dimensional vectors. We will use vector $\boldsymbol{s}_i$ to denote the pre-assigned unique signature sequence of user $i$. In terms of signal vectors, the received signal at the base station can be written as

$$\boldsymbol{r} = \sum_{i=1}^{N} \sqrt{p_i} a_i \boldsymbol{s}_i + \boldsymbol{n} \tag{3.2}$$

where $\boldsymbol{n}$ is a Gaussian random vector with $E\left[\boldsymbol{n}\boldsymbol{n}^\top\right] = \sigma^2 \boldsymbol{I}$.

Let us define a $G \times N$ dimensional matrix $\boldsymbol{S}$ with its $(i,j)$th element being $(\boldsymbol{s}_j)_i$, the $i$th component of $\boldsymbol{s}_j$. Therefore, the columns of $\boldsymbol{S}$ are the signature sequences of the users. For future use, we define a subspace $\mathcal{L}$ in $G$ dimensional vector space to be the subspace spanned by the signature sequences of the users, i.e.,

$$\mathcal{L} = \text{span}\{\boldsymbol{s}_1, \ldots, \boldsymbol{s}_N\} = \text{column space of } \boldsymbol{S} \tag{3.3}$$

In the following section we will briefly summarize the well known decorrelating detector and develop the necessary background for the presentation of the blind adaptive decorrelating detector.

## 3.2 The Decorrelating Detector

We consider the decorrelating detector for the first user without loss of generality and represent it by a $G$ dimensional vector $\boldsymbol{c}$. Then $\boldsymbol{c}$ should satisfy the following condition,

$$\boldsymbol{S}^\top \boldsymbol{c} = \alpha \boldsymbol{e}_1 \tag{3.4}$$

where $\boldsymbol{e}_1$ is the first unit vector in $N$ dimensional space, i.e., $\boldsymbol{e}_1 = [1\ 0\ 0 \ldots\ 0]^\top$, and $\alpha$ is a non-negative real number. Equation (3.4) states that the receiver filter for the desired (first) user should be orthogonal to the signature sequences of all the interfering users, and it should have a nonzero correlation with the desired user. Here this nonzero correlation is equal to the scaling factor $\alpha$. It is not difficult to show that the bit error rate (BER) performance of the decorrelating detector is insensitive to scaling of the filter as long as the filter eliminates the multiaccess interference totally. The reason for this is that the scalar factor multiplies both the received power level of the desired user and the AWGN. Thus, for any nonnegative value of $\alpha$ we have a decorrelating detector with the same BER performance. The importance of using $\alpha \neq 1$ will be evident in the next section where we will avoid the need for the knowledge about the received power of the desired user by a particular selection of $\alpha$.

We first note that (3.4) has more than one solution, because it has $N$ equalities and $G$ unknowns, and typically $G \gg N$. The unique decorrelating detector for the first user, $\boldsymbol{d}$, is given in [28] as

$$\boldsymbol{d} = \boldsymbol{S}(\boldsymbol{S}^\top \boldsymbol{S})^{-1} \boldsymbol{e}_1 \tag{3.5}$$

Let us denote any solution of (3.4) as $\bar{\boldsymbol{c}}$, then inserting (3.4) into (3.5) we obtain (assuming $\alpha = 1$ for the time being)

$$\begin{aligned} \boldsymbol{d} &= \boldsymbol{S}(\boldsymbol{S}^\top \boldsymbol{S})^{-1}\boldsymbol{S}^\top \bar{\boldsymbol{c}} && (3.6) \\ &= \boldsymbol{P}\bar{\boldsymbol{c}} && (3.7) \end{aligned}$$

Note that $\boldsymbol{P} = \boldsymbol{S}(\boldsymbol{S}^\top \boldsymbol{S})^{-1}\boldsymbol{S}^\top$ is the projection matrix which projects any vector onto the column space of $\boldsymbol{S}$. Also note that the column space of $\boldsymbol{S}$ is the subspace spanned by $\boldsymbol{s}_1, \ldots, \boldsymbol{s}_N$, i.e., the subspace which was previously denoted as $\mathcal{L}$. Therefore, although (3.4) has more than one solution, all of the solutions have the same projection onto $\mathcal{L}$; and this projection is equal to the unique decorrelating detector solution of [28].

Let $\boldsymbol{\Lambda}$ be an $N \times N$ diagonal matrix with received power of user $i$, $p_i$, being its $i$th diagonal element. Multiplying both sides of (3.4) first with $\boldsymbol{\Lambda}$, then with $\boldsymbol{S}$, we obtain

$$\boldsymbol{S}\boldsymbol{\Lambda}\boldsymbol{S}^\top \boldsymbol{c} = \alpha p_1 \boldsymbol{s}_1 \tag{3.8}$$

We observe that although (3.8) has $G$ equations in $G$ unknowns it does not have a unique solution for $\boldsymbol{c}$ since $\boldsymbol{S}\boldsymbol{\Lambda}\boldsymbol{S}^\top$ has rank $N$ and thus $G - N$ zero eigenvalues. However, the following Lemma relates the solution spaces of (3.4) and (3.8).

**Lemma 1** *If $\boldsymbol{s}_1, \ldots, \boldsymbol{s}_N$ are linearly independent, then all solutions of (3.4) and (3.8) coincide.*

A proof for Lemma 1 is given in Appendix 3.A.

Let us define $\boldsymbol{A} = \boldsymbol{S}\boldsymbol{\Lambda}\boldsymbol{S}^\top$. At this point, we choose $\alpha = 1/p_1$ and devise the following gradient descent algorithm:

$$\boldsymbol{c}(n+1) = \boldsymbol{c}(n) - \mu\left(\boldsymbol{A}\boldsymbol{c}(n) - \boldsymbol{s}_1\right) \tag{3.9}$$

Note that the iterative algorithm given in (3.9) converges to the solution of (3.8) and equivalently, by Lemma 1, to the solution of (3.4). The fact that the solution of (3.4) is not unique is noted earlier. Note that for any $\boldsymbol{c}$

$$\boldsymbol{P}\boldsymbol{A}\boldsymbol{c} = \boldsymbol{S}(\boldsymbol{S}^\top \boldsymbol{S})^{-1}\boldsymbol{S}^\top \boldsymbol{S}\boldsymbol{\Lambda}\boldsymbol{S}^\top \boldsymbol{c} = \boldsymbol{S}\boldsymbol{\Lambda}\boldsymbol{S}^\top \boldsymbol{c} = \boldsymbol{A}\boldsymbol{c} \tag{3.10}$$

This means that for any $\boldsymbol{c}(n)$, $\boldsymbol{A}\boldsymbol{c}(n)$ lies entirely in $\mathcal{L}$. Since $\boldsymbol{s}_1 \in \mathcal{L}$ by definition, the correction term added to $\boldsymbol{c}(n)$ at every iteration in (3.9) is always in $\mathcal{L}$. Starting from $n = 0$, by induction we obtain

$$\boldsymbol{c}(n) = \boldsymbol{P}\boldsymbol{c}(n) + \boldsymbol{c}(0) - \boldsymbol{P}\boldsymbol{c}(0) \tag{3.11}$$

Therefore, if the iterative algorithm (3.9) is started in the subspace $\mathcal{L}$, implying $\boldsymbol{c}(0) = \boldsymbol{P}\boldsymbol{c}(0)$, then from (3.11) we will have: $\boldsymbol{c}(n) = \boldsymbol{P}\boldsymbol{c}(n)$ for all $n$. In this case $\boldsymbol{c}(n)$ will always stay in $\mathcal{L}$ and it will converge to the scaled version of the unique decorrelating detector solution $\boldsymbol{c}^* = \frac{1}{p_1}\boldsymbol{d}$ as $n$ goes to infinity. Note that the algorithm converges to the scaled version of $\boldsymbol{d}$, the decorrelating detector of [28, 29], instead of converging to $\boldsymbol{d}$, because $\alpha$ was chosen to be $1/p_1$ instead of 1. Note also that $\boldsymbol{d}$ and $\boldsymbol{c}^* = \frac{1}{p_1}\boldsymbol{d}$ have the same BER performance as discussed earlier.

The restriction that the iterations should be started in $\mathcal{L}$ for algorithm (3.9) to converge to a decorrelating detector is fairly mild. The vectors $\boldsymbol{c}(0) = \boldsymbol{0}$, or $\boldsymbol{c}(0) = \boldsymbol{s}_1$, or any linear combination of the signature sequences all belong to $\mathcal{L}$. The signature sequences of all users must be known for the algorithm given in (3.9). Also, the algorithm of (3.9) is an off-line algorithm in the sense that it does not utilize any real time measurements or observations. The algorithm of (3.9) can be run before the users start transmitting information bits. After running the algorithm for some time, the filter coefficients would converge to the decorrelating detector filter and then communication can be started. In this chapter, our aim is to develop a blind adaptive algorithm which would converge to the decorrelating detector solution in a stochastic sense by using real random measurements while the users are active and transmitting bits. To this end, we will propose an algorithm which can be viewed as the stochastic version of the deterministic algorithm given in (3.9).

## 3.3 A Blind Adaptive Decorrelating Detector (BADD)

From (3.2) we have,

$$E\left[\boldsymbol{r}\boldsymbol{r}^\top\right] = \sum_{i=1}^{N} p_i \boldsymbol{s}_i \boldsymbol{s}_i^\top + \sigma^2 \boldsymbol{I} = \boldsymbol{A} + \sigma^2 \boldsymbol{I} \tag{3.12}$$

Therefore, using (3.12), the deterministic iteration of (3.9) can be written in an exact form as

$$\boldsymbol{c}(n+1) = \boldsymbol{c}(n) - \mu \left[ \left( E\left[ \boldsymbol{r}\boldsymbol{r}^\top \right] - \sigma^2 \boldsymbol{I} \right) \boldsymbol{c}(n) - \boldsymbol{s}_1 \right] \tag{3.13}$$

Note that $\boldsymbol{r}\boldsymbol{r}^\top - \sigma^2\boldsymbol{I}$ is an *unbiased* estimate for $\boldsymbol{A}$ matrix. In order to obtain a practical algorithm, we replace $E\left[\boldsymbol{r}\boldsymbol{r}^\top\right]$ in (3.13) with the estimate $\boldsymbol{r}(n)\boldsymbol{r}^\top(n)$ where $\boldsymbol{r}(n)$ is the received signal vector at the $n$th iteration.

$$\boldsymbol{c}(n+1) = \boldsymbol{c}(n) - \mu \left[ \left( \boldsymbol{r}(n)\boldsymbol{r}^\top(n) - \sigma^2 \boldsymbol{I} \right) \boldsymbol{c}(n) - \boldsymbol{s}_1 \right] \tag{3.14}$$

Before analyzing the convergence of (3.14), we state it in terms of available observations, and list the parameters needed at each iteration. We note that the output of the receiver filter of the desired user at time $n$ is $y(n) = \boldsymbol{r}^\top(n)\boldsymbol{c}(n)$. Thus an implementation of the algorithm (3.14) is

$$\boldsymbol{c}(n+1) = \left(1 + \sigma^2\mu\right) \boldsymbol{c}(n) - \mu \left( \boldsymbol{r}(n)y(n) - \boldsymbol{s}_1 \right) \tag{3.15}$$

Since $\boldsymbol{r}(n)$ and $y(n)$ are readily available at the input and output of the receiver filter that is under construction, only two system parameters are needed to be known in order to run the algorithm: the signature sequence of the desired user $\boldsymbol{s}_1$, and the variance of the AWGN, $\sigma^2$.

## 3.4 BADD Convergence

In this Section we will investigate the convergence of the blind adaptive decorrelating detector (BADD) proposed in the previous section. Let us define the zero mean random vector $\boldsymbol{\eta}(n)$ as

$$\boldsymbol{\eta}(n) = \left( \boldsymbol{r}\boldsymbol{r}^\top - \boldsymbol{A} - \sigma^2 \boldsymbol{I} \right) \boldsymbol{c}(n) \tag{3.16}$$

Noting that $\boldsymbol{A}\boldsymbol{c}^* = \boldsymbol{s}_1$, we can write the stochastic iterations (3.14) and (3.15) as

$$\boldsymbol{c}(n+1) = \boldsymbol{c}(n) - \mu \left[ \boldsymbol{A}\left( \boldsymbol{c}(n) - \boldsymbol{c}^* \right) + \boldsymbol{\eta}(n) \right] \tag{3.17}$$

Subtracting $\boldsymbol{c}^*$ from both sides of (3.17) and defining $\boldsymbol{\epsilon}(n) = \boldsymbol{c}(n) - \boldsymbol{c}^*$ we obtain

$$\boldsymbol{\epsilon}(n+1) = \boldsymbol{\epsilon}(n) - \mu \left( \boldsymbol{A}\boldsymbol{\epsilon}(n) + \boldsymbol{\eta}(n) \right) \tag{3.18}$$

Note that $\|\boldsymbol{\epsilon}(n)\|^2$, the norm of $\boldsymbol{\epsilon}(n)$, is a measure of the distance of the receiver filter at iteration $n$ from the convergence point. In order to relate the distance measures at iterations $n$ and $n+1$, we square both sides of (3.18) and obtain

$$\begin{aligned} \|\boldsymbol{\epsilon}(n+1)\|^2 &= \|\boldsymbol{\epsilon}(n)\|^2 - 2\mu\boldsymbol{\epsilon}(n)^\top \boldsymbol{A}\boldsymbol{\epsilon}(n) + 2\mu^2\boldsymbol{\epsilon}(n)^\top \boldsymbol{A}\boldsymbol{\eta}(n) - 2\mu\boldsymbol{\epsilon}(n)^\top \boldsymbol{\eta}(n) \\ &\quad + \mu^2\boldsymbol{\epsilon}(n)^\top \boldsymbol{A}^2\boldsymbol{\epsilon}(n) + \mu^2 \|\boldsymbol{\eta}(n)\|^2 \end{aligned} \tag{3.19}$$

Taking the conditional expectation of both sides of (3.19), conditioned on $\boldsymbol{\epsilon}(n) = \boldsymbol{\epsilon}$, and observing that $E[\boldsymbol{\eta}(n) \mid \boldsymbol{\epsilon}(n) = \boldsymbol{\epsilon}] = \mathbf{0}$ yields

$$E\left[\|\boldsymbol{\epsilon}(n+1)\|^2 \mid \boldsymbol{\epsilon}(n) = \boldsymbol{\epsilon}\right] = \|\boldsymbol{\epsilon}\|^2 - 2\mu\boldsymbol{\epsilon}^\top \boldsymbol{A}\boldsymbol{\epsilon} + \mu^2\boldsymbol{\epsilon}^\top \boldsymbol{A}^2\boldsymbol{\epsilon} + \mu^2 E\left[\|\boldsymbol{\eta}(n)\|^2 \mid \boldsymbol{\epsilon}(n) = \boldsymbol{\epsilon}\right] \tag{3.20}$$

We will be using the results of the following Lemmas to develop bounds for the right hand side of (3.20).

**Lemma 2** *If* $\boldsymbol{P}\boldsymbol{\epsilon} = \mathbf{0}$ *then* $\boldsymbol{\epsilon}^\top \boldsymbol{A}\boldsymbol{\epsilon} = 0$ *and* $\boldsymbol{\epsilon}^\top \boldsymbol{A}^2\boldsymbol{\epsilon} = 0$. *If* $\boldsymbol{P}\boldsymbol{\epsilon} \neq \mathbf{0}$ *then there exist* $0 < k_0 \leq k_1 < \infty$, *such that*

$$k_0 \|\boldsymbol{\epsilon}\|^2 \leq \boldsymbol{\epsilon}^\top \boldsymbol{A}\boldsymbol{\epsilon} \leq k_1 \|\boldsymbol{\epsilon}\|^2 \qquad \textit{and} \qquad k_0^2 \|\boldsymbol{\epsilon}\|^2 \leq \boldsymbol{\epsilon}^\top \boldsymbol{A}^2\boldsymbol{\epsilon} \leq k_1^2 \|\boldsymbol{\epsilon}\|^2 \tag{3.21}$$

**Lemma 3** *There exist* $0 < c_0 \leq c_1 < \infty$, *such that*

$$0 \leq E\left[\|\boldsymbol{\eta}(n)\|^2 \mid \boldsymbol{\epsilon}(n) = \boldsymbol{\epsilon}\right] \leq c_0 + c_1 \|\boldsymbol{\epsilon}\|^2 \tag{3.22}$$

The proofs of Lemmas 2 and 3 can be found in Appendix 3.A.

If $\boldsymbol{P}\boldsymbol{\epsilon} = \mathbf{0}$, we can use Lemmas 2 and 3 to develop the following bounds for the right side of (3.20)

$$E\left[\|\boldsymbol{\epsilon}(n+1)\|^2 \mid \boldsymbol{\epsilon}(n) = \boldsymbol{\epsilon}, \boldsymbol{P}\boldsymbol{\epsilon} = \mathbf{0}\right] \leq \left(1 + \mu^2 c_1\right) \|\boldsymbol{\epsilon}\|^2 + c_0\mu^2 \tag{3.23}$$

$$E\left[\|\boldsymbol{\epsilon}(n+1)\|^2 \mid \boldsymbol{\epsilon}(n) = \boldsymbol{\epsilon}, \boldsymbol{P}\boldsymbol{\epsilon} = \mathbf{0}\right] \geq \|\boldsymbol{\epsilon}\|^2 \tag{3.24}$$

If $\boldsymbol{P}\boldsymbol{\epsilon} \neq \mathbf{0}$, again we can use Lemmas 2 and 3 to obtain the following bounds for the right side of (3.20)

$$E\left[\|\boldsymbol{\epsilon}(n+1)\|^2 \mid \boldsymbol{\epsilon}(n) = \boldsymbol{\epsilon}, \boldsymbol{P}\boldsymbol{\epsilon} \neq \mathbf{0}\right] \leq \left(1 - 2\mu k_0 + (k_1^2 + c_1)\mu^2\right) \|\boldsymbol{\epsilon}\|^2 + c_0\mu^2 \tag{3.25}$$

$$E\left[\|\boldsymbol{\epsilon}(n+1)\|^2 \mid \boldsymbol{\epsilon}(n) = \boldsymbol{\epsilon}, \boldsymbol{P}\boldsymbol{\epsilon} \neq \mathbf{0}\right] \geq \left(1 - 2\mu k_1 + k_0^2\mu^2\right) \|\boldsymbol{\epsilon}\|^2 \tag{3.26}$$

An important observation toward the convergence proof is that the projection of $\boldsymbol{c}(n)$ (equivalently the projection of $\boldsymbol{\epsilon}(n)$) on $\mathcal{L}$ is non-zero *almost surely (a.s.)* [48]. This means that for any $n$ the probability of the event $[\boldsymbol{P}\boldsymbol{\epsilon}(n) = \mathbf{0}]$ is zero. Similar to the deterministic result in (3.11), it can be shown using induction on (3.18) that (see also [18, Eqn. (16)]),

$$\boldsymbol{\epsilon}(n) = \boldsymbol{P}\boldsymbol{\epsilon}(n) + \boldsymbol{\epsilon}(0) - \boldsymbol{P}\boldsymbol{\epsilon}(0) \qquad \text{a. s.} \tag{3.27}$$

If $\boldsymbol{\epsilon}(0) \in \mathcal{L}$ then $\boldsymbol{\epsilon}(0) = \boldsymbol{P}\boldsymbol{\epsilon}(0)$ and (3.27) can be written as

$$\boldsymbol{\epsilon}(n) = \boldsymbol{P}\boldsymbol{\epsilon}(n) \qquad \text{a. s.} \tag{3.28}$$

Note that $\boldsymbol{\epsilon}(0) = \boldsymbol{c}(0) - \boldsymbol{c}^*$. Since $\boldsymbol{c}^* \in \mathcal{L}$, $\boldsymbol{\epsilon}(0) \in \mathcal{L}$ is equivalent to $\boldsymbol{c}(0) \in \mathcal{L}$. Thus, (3.28) is valid if the iteration (3.14) is started in $\mathcal{L}$. The upper bound on $E\left[\|\boldsymbol{\epsilon}(n+1)\|^2 \mid \boldsymbol{\epsilon}(n) = \boldsymbol{\epsilon}\right]$ can be developed as follows.

$$\begin{aligned}
E\left[\|\boldsymbol{\epsilon}(n+1)\|^2 \mid \boldsymbol{\epsilon}(n) = \boldsymbol{\epsilon}\right] &= E\left[\|\boldsymbol{\epsilon}(n+1)\|^2 \mid \boldsymbol{\epsilon}(n) = \boldsymbol{\epsilon}, \boldsymbol{P}\boldsymbol{\epsilon} = \mathbf{0}\right] \mathrm{Prob}\left[\boldsymbol{P}\boldsymbol{\epsilon} = \mathbf{0}\right] \\
&\quad + E\left[\|\boldsymbol{\epsilon}(n+1)\|^2 \mid \boldsymbol{\epsilon}(n) = \boldsymbol{\epsilon}, \boldsymbol{P}\boldsymbol{\epsilon} \neq \mathbf{0}\right] \mathrm{Prob}\left[\boldsymbol{P}\boldsymbol{\epsilon} \neq \mathbf{0}\right] \\
&= E\left[\|\boldsymbol{\epsilon}(n+1)\|^2 \mid \boldsymbol{\epsilon}(n) = \boldsymbol{\epsilon}, \boldsymbol{P}\boldsymbol{\epsilon} \neq \mathbf{0}\right] \\
&\leq \left(1 - 2\mu k_0 + (k_1^2 + c_1)\mu^2\right) \|\boldsymbol{\epsilon}\|^2 + c_0\mu^2
\end{aligned} \tag{3.29}$$

Note that we have used the facts $\mathrm{Prob}\left[\boldsymbol{P}\boldsymbol{\epsilon} = \mathbf{0}\right] = 0$ and $\mathrm{Prob}\left[\boldsymbol{P}\boldsymbol{\epsilon} \neq \mathbf{0}\right] = 1$ and the result in (3.25). Following similar steps to those in (3.29), and using the result given in (3.26), we can obtain the following lower bound for the same term

$$E\left[\|\boldsymbol{\epsilon}(n+1)\|^2 \mid \boldsymbol{\epsilon}(n) = \boldsymbol{\epsilon}\right] \geq \left(1 - 2\mu k_1 + k_0^2\mu^2\right) \|\boldsymbol{\epsilon}\|^2 \tag{3.30}$$

Taking the expectation of both sides of the inequalities in (3.29) and (3.30) with respect to $\boldsymbol{\epsilon}(n)$ and letting $b_n = E\left[\|\boldsymbol{\epsilon}(n)\|^2\right]$, we obtain the following lower and upper bounds for $b_{n+1}$, the mean squared error (MSE) of the filter coefficients at iteration $n+1$ from $\boldsymbol{c}^*$, in terms of $b_n$

$$\left(1 - 2\mu k_1 + k_0^2\mu^2\right) b_n \leq b_{n+1} \leq \left(1 - 2\mu k_0 + (k_1^2 + c_1)\mu^2\right) b_n + c_0\mu^2 \tag{3.31}$$

By defining

$$\alpha_0 = 1 - 2\mu k_0 + (k_1^2 + c_1)\mu^2 \qquad \alpha_1 = 1 - 2\mu k_1 + k_0^2\mu^2 \tag{3.32}$$

we can rewrite (3.31) as

$$\alpha_1 b_n \leq b_{n+1} \leq \alpha_0 b_n + c_0 \mu^2 \tag{3.33}$$

We observe from (3.33) that the nonnegative sequence $b_n$ is sandwiched between the two sequences generated according to $b'_{n+1} = \alpha_0 b'_n + c_0 \mu^2$ and $b''_{n+1} = \alpha_1 b''_n$. These two sequences converge to finite values if and only if $\mu$ is chosen such that $\mid \alpha_0 \mid < 1$ and $\mid \alpha_1 \mid < 1$. Note that both $\alpha_0$ and $\alpha_1$ are equal to 1 at $\mu = 0$. We also note that both $\alpha_0$ and $\alpha_1$ are locally decreasing as $\mu$ increases, since

$$\left.\frac{d\alpha_0}{d\mu}\right|_{\mu=0} = -2k_0 < 0 \qquad \text{and} \qquad \left.\frac{d\alpha_1}{d\mu}\right|_{\mu=0} = -2k_1 < 0 \tag{3.34}$$

This implies that we can always choose $\mu$ small enough so that

$$|\alpha_0| < 1 \qquad \text{and} \qquad |\alpha_1| < 1 \tag{3.35}$$

and in this case the sequences $b'_n$ and $b''_n$ converge and the limiting MSE, i.e., $\lim_{n\to\infty} b_n$, has finite lower and upper bounds. From the sandwich theorem, we have

$$0 \leq \lim_{n\to\infty} b_n \leq \frac{c_0 \mu^2}{1 - \alpha_0} \tag{3.36}$$

We can evaluate the value of the upper bound in the extreme case when $\mu \to 0$ as,

$$\lim_{\mu\to 0} \frac{c_0 \mu^2}{1-\alpha_0} = \lim_{\mu \to 0} \frac{c_0 \mu}{2k_0 - (k_1^2 + c_1)\mu} = 0 \tag{3.37}$$

Therefore, if the step size ($\mu$) is chosen extremely small, then the MSE of the filter coefficients goes to zero as the number of iterations grows to infinity. But note that as $\mu \to 0$, the numbers $\alpha_0$ and $\alpha_1$ approach 1, and the convergence rate goes to zero. Thus we observe the tradeoff between the limiting value of the MSE and the convergence rate: if the step size $\mu$ is large, then the convergence rate is faster but the limiting MSE is larger; and if the step size $\mu$ is small, the limiting MSE is small but the convergence rate is slow. Hence, a time dependent step size sequence which takes large values at the beginning and smaller values at the end may be preferable. A step size sequence that depends on the iteration index $n$ can be used to accomplish this. Replacing the fixed step size $\mu$ in (3.15) with the time varying step size sequence $a_n$, we obtain the new algorithm

$$\boldsymbol{c}(n+1) = \left(1 + \sigma^2 a_n\right) \boldsymbol{c}(n) - a_n \left(\boldsymbol{r}(n) y(n) - \boldsymbol{s}_1\right) \tag{3.38}$$

If the step size sequence satisfies the Robbins-Monro conditions [41]:

$$\sum_{n}^{\infty} a_n = \infty \qquad \text{and} \qquad \sum_{n}^{\infty} a_n^2 < \infty \tag{3.39}$$

then the filter coefficients converge to the decorrelating detector solution in the MSE sense. The convergence is stated in the following Lemma which is proven in Appendix 3.A.

**Lemma 4** *Stochastic iteration given in (3.38) converges to* $\mathbf{c}^*$ *in the MSE sense, i.e.,*

$$\lim_{n\to 0} E\left[\|\boldsymbol{c}(n) - \mathbf{c}^*\|^2\right] = 0$$

*if* $a_n$ *satisfies the conditions of (3.39).*

Note that $a_n = a/(n+n_0)$, for any $a > 0$ and $n_0 > 0$ is a sequence which satisfies (3.39).

## 3.5 The BADD in an Asynchronous CDMA System

In sections 3.1-3.4, we assumed that the system was synchronous. Now suppose $\tau_i$ denote the delay of the $i$th user. Without loss of generality, we will assume that $\tau_1$, the desired user's delay, is equal to zero and $\tau_i \geq 0$, for $i > 1$. As in [64], each asynchronous interfering user $i$, can be replaced with two synchronous users $2i-2$ and $2i-1$ with modified signature sequences; see Figure 3.1. The modified signature sequences, $\bar{s}_i(t)$, of these synchronous users are given as

$$\begin{aligned} \bar{s}_{2i-2}(t) &= \begin{cases} s_i(t) & 0 \leq t \leq \tau_i \\ 0 & \tau_i < t \leq T_b \end{cases} \\ \bar{s}_{2i-1}(t) &= \begin{cases} 0 & 0 \leq t \leq \tau_i \\ s_i(t) & \tau_i < t \leq T_b \end{cases} \end{aligned}$$

where $T_b$ is the bit duration. In this case the received signal before the receiver filter of the desired user at the $n$th bit interval is given as

$$r_n(t) = \sum_{i=1}^{N} \sqrt{p_i}\,(a_i(n-1)\bar{s}_{2i-2}(t) + a_i(n)\bar{s}_{2i-1}(t)) + n(t) \tag{3.40}$$

where $a_i(n)$ denotes the information bit transmitted by user $i$ in the $n$th bit interval. Denoting the modified signature sequence $\bar{s}_i(t)$ in a vector form as $\bar{\boldsymbol{s}}_i$, the chip sampled received signal before the receiver filter of the desired user at the $n$th bit interval, $\boldsymbol{r}(n)$, can be written as

$$\boldsymbol{r}(n) = \sum_{i=1}^{N} \sqrt{p_i}\left(a_i(n-1)\bar{\boldsymbol{s}}_{2i-2} + a_i(n)\bar{\boldsymbol{s}}_{2i-1}\right) + \boldsymbol{n} \tag{3.41}$$

In (3.40) and (3.41), for convenience, the desired (first) user is also written in terms of the signature of two equivalent users. Note that the signature sequence $\bar{\boldsymbol{s}}_0$ is identically zero and $\bar{\boldsymbol{s}}_1 = \boldsymbol{s}_1$.

From the mutual independence of the transmitted bits and the AWGN we have

$$\begin{aligned} E\left[\boldsymbol{r}(n)\boldsymbol{r}^\top(n)\right] &= \sum_{i=1}^{N} p_i\left(\bar{\boldsymbol{s}}_{2i-2}\bar{\boldsymbol{s}}_{2i-2}^\top + \bar{\boldsymbol{s}}_{2i-1}\bar{\boldsymbol{s}}_{2i-1}^\top\right) + \sigma^2\boldsymbol{I} \\ &= \sum_{i=1}^{2N-1} \bar{p}_i\bar{\boldsymbol{s}}_i\bar{\boldsymbol{s}}_i^\top + \sigma^2\boldsymbol{I} \\ &= \bar{\boldsymbol{A}} + \sigma^2\boldsymbol{I} \end{aligned} \tag{3.42}$$

where $\bar{p}_{2i-2}$ and $\bar{p}_{2i-1}$ are equal to $p_i$ since they denote the received power levels of the two synchronous users representing the asynchronous user $i$.

If the stochastic iteration (3.14) is used for the asynchronous system described in this section, the received filter of the desired user will converge to the point $\bar{\boldsymbol{c}}^*$ where

$$\bar{\boldsymbol{A}}\bar{\boldsymbol{c}}^* = \boldsymbol{s}_1 \tag{3.43}$$

We define a $G \times (2N-1)$ matrix $\bar{\boldsymbol{S}}$ with its $i$th column being $\bar{\boldsymbol{s}}_i$. Thus, $\bar{\boldsymbol{S}}$ contains the signature sequences of the equivalent synchronous system with $(2N-1)$ users. Equation (3.43) is equivalent to

$$\bar{\boldsymbol{S}}\bar{\boldsymbol{\Lambda}}\bar{\boldsymbol{S}}^\top\bar{\boldsymbol{c}}^* = \boldsymbol{s}_1 \tag{3.44}$$

where $\bar{\boldsymbol{\Lambda}}$ is a $(2N-1) \times (2N-1)$ dimensional diagonal matrix with $\bar{\boldsymbol{\Lambda}}_{ii} = \bar{p}_i$. By Lemma 1, the solution of (3.44) is equal to the solution of

$$\bar{\boldsymbol{S}}^\top\bar{\boldsymbol{c}}^* = \boldsymbol{e}_1 \tag{3.45}$$

if the columns of $\bar{\boldsymbol{S}}$, equivalently the modified signature sequences, are linearly independent. This condition is the *Linear Independence Assumption (LIA)* of [29] which

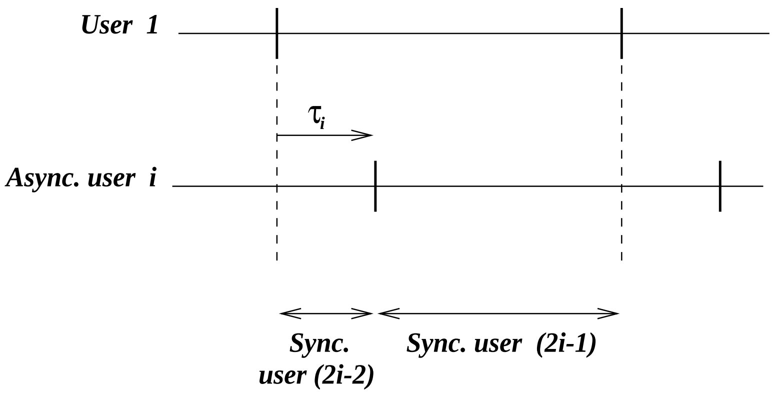


Figure 3.1: Asynchronous CDMA model.

requires $(2N-1) \leq G$. In this chapter, for the asynchronous channel, we will assume that LIA is in effect as was done in [29].

The convergence point $\bar{\boldsymbol{c}}^*$ is the *one-shot decorrelator* [29, 62] which is the decorrelating receiver filter when only one bit of the desired user is considered. The one-shot decorrelator has lower complexity than asynchronous decorrelating detectors which take into account the dependence of the received signal in different bit intervals. However, the one-shot decorrelating detector does suffer some loss in performance [29, 62].

## 3.6 Effects of Noise Variance Estimation

The proposed algorithm (3.15) requires that the variance of the AWGN is known and used in the update of the filter coefficients. In this section, we investigate the effect of missing the exact value of noise variance $\sigma^2$.

In wireless communication channels, the noise variance is a fixed parameter, unlike other parameters such as channel gains that are subject to fading. Although it is fixed, it is unknown and needs to be estimated. There can be two approaches:

1. The noise variance can be estimated before the information transmission hence the update of the filter coefficients starts and this estimate (a fixed deterministic number) is used in every update of the filter.

2. A real time estimator which produces a new (random) estimate for the noise variance can be built, and its estimates can be used in the filter updates. The

physical implementation of this estimator can be realized in many ways. One method could be to allocate a unit energy filter $\boldsymbol{v}$ which is orthogonal to all signature sequences being used by the users and use the output of this filter for estimating $\sigma^2$. Since the filter is orthogonal to all signature sequences, the square of its output is equal to

$$Z = (\boldsymbol{v}^\top \boldsymbol{n})^2 \tag{3.46}$$

where $\boldsymbol{n}$ is the same as that in (3.2). Note that since $\boldsymbol{v}$ is unit energy, $E[Z] = \sigma^2$ and $Z$ is an unbiased estimate of $\sigma^2$.

Whether an estimate of $\sigma^2$ is obtained before the bit transmission starts and this same fixed estimate is used in every update of the filter coefficients or a new unbiased estimate is used at each filter update is an implementation issue. Below, we will investigate the performance of the proposed algorithm in both cases.

### 3.6.1 Estimating $\sigma^2$ Before Filter Updates

In this section we assume that a deterministic number $\tilde{\sigma}^2$ is used in place of $\sigma^2$ in (3.15) or equivalently, in (3.14). From (3.14), the update equation becomes:

$$\boldsymbol{c}(n+1) = \boldsymbol{c}(n) - \mu \left[ \left( \boldsymbol{r}(n)\boldsymbol{r}^\top(n) - \tilde{\sigma}^2 \boldsymbol{I} \right) \boldsymbol{c}(n) - \boldsymbol{s}_1 \right] \tag{3.47}$$

The deterministic version of the algorithm becomes

$$\boldsymbol{c}(n+1) = \boldsymbol{c}(n) - \mu \left[ \tilde{\boldsymbol{A}} \boldsymbol{c}(n) - \boldsymbol{s}_1 \right] \tag{3.48}$$

where

$$\tilde{\boldsymbol{A}} = \boldsymbol{S}\boldsymbol{\Lambda}\boldsymbol{S}^\top + (\sigma^2 - \tilde{\sigma}^2)\boldsymbol{I} = \sum_{i=1}^{N} p_i \boldsymbol{s}_i \boldsymbol{s}_i^\top + (\sigma^2 - \tilde{\sigma}^2)\boldsymbol{I} \tag{3.49}$$

If $\tilde{\sigma}^2 \leq \sigma^2$, then $\tilde{\boldsymbol{A}}$ is a positive semidefinite matrix and the algorithm in (3.48) converges deterministically. Further, the stochastic algorithm of (3.47), converges in the MSE sense to $\tilde{\boldsymbol{c}}$ satisfying

$$\tilde{\boldsymbol{A}} \tilde{\boldsymbol{c}} = \boldsymbol{s}_1 \tag{3.50}$$

If $\tilde{\sigma}^2 < \sigma^2$, then $\tilde{\boldsymbol{A}}$ is invertible and

$$\tilde{\boldsymbol{c}} = \tilde{\boldsymbol{A}}^{-1} \boldsymbol{s}_1 \tag{3.51}$$

The convergence of the stochastic iteration (3.47) to $\tilde{\boldsymbol{c}}$ can be shown by writing (3.47) as

$$\boldsymbol{c}(n+1) = \boldsymbol{c}(n) - \mu \left[ \left( \boldsymbol{r}(n)\boldsymbol{r}^\top(n) - \tilde{\boldsymbol{A}} + \tilde{\boldsymbol{A}} - \tilde{\sigma}^2 \boldsymbol{I} \right) \boldsymbol{c}(n) - \boldsymbol{s}_1 \right] \tag{3.52}$$

so that (3.50) implies

$$\boldsymbol{c}(n+1) = \boldsymbol{c}(n) - \mu \left[ \tilde{\boldsymbol{A}} \left( \boldsymbol{c}(n) - \tilde{\boldsymbol{c}} \right) + \left( \boldsymbol{r}(n)\boldsymbol{r}^\top(n) - (\tilde{\boldsymbol{A}} + \tilde{\sigma}^2 \boldsymbol{I}) \right) \boldsymbol{c}(n) \right] \tag{3.53}$$

Note that the random term $(\boldsymbol{r}(n)\boldsymbol{r}^\top(n) - (\tilde{\boldsymbol{A}} + \tilde{\sigma}^2\boldsymbol{I}))\boldsymbol{c}(n)$ in (3.53) is identical to $\boldsymbol{\eta}(n)$ defined in (3.16), and thus (3.53) can be written as

$$\boldsymbol{c}(n+1) = \boldsymbol{c}(n) - \mu \left[ \tilde{\boldsymbol{A}} \left( \boldsymbol{c}(n) - \tilde{\boldsymbol{c}} \right) + \boldsymbol{\eta}(n) \right] \tag{3.54}$$

Subtracting $\tilde{\boldsymbol{c}}$ from both sides of (3.54) and defining $\tilde{\boldsymbol{\epsilon}}(n)$ as $\tilde{\boldsymbol{\epsilon}}(n) = \boldsymbol{c}(n) - \tilde{\boldsymbol{c}}$ yields

$$\tilde{\boldsymbol{\epsilon}}(n+1) = \tilde{\boldsymbol{\epsilon}}(n) - \mu \left( \tilde{\boldsymbol{A}} \tilde{\boldsymbol{\epsilon}}(n) + \boldsymbol{\eta}(n) \right) \tag{3.55}$$

Comparing (3.55) with (3.18), we observe that the Lemmas 2 and 3 are still valid, and the algorithm converges in the MSE sense. Note also that since there is a unique solution for $\tilde{\boldsymbol{c}}$ when $\tilde{\sigma}^2 \neq \sigma^2$, the arguments involving the projection onto signal space are not necessary. Also unnecessary is the condition that the algorithm should be started in the signal space.

Note that $\tilde{\boldsymbol{c}}$ in (3.51) is a scaled version of the MMSE multiuser detector where a reduced value for the noise variance is used, i.e., $(\sigma^2 - \tilde{\sigma}^2)$ is used instead of the true noise variance $\sigma^2$.

In two extreme cases, the detector $\tilde{\boldsymbol{c}}$ reduces to a well known multiuser detector. If $\tilde{\sigma}^2 = 0$, then $\tilde{\boldsymbol{c}}$ is the true MMSE multiuser detector and if $\tilde{\sigma}^2 = \sigma^2$ then $\tilde{\boldsymbol{c}}$ is the decorrelating multiuser detector. Defining $\tilde{\boldsymbol{A}}_1$ as $\tilde{\boldsymbol{A}}_1 = \sum_{i\neq 1}^{N} p_i \boldsymbol{s}_i \boldsymbol{s}_i^\top + (\sigma^2 - \tilde{\sigma}^2)\boldsymbol{I}$, the filter in (3.51) can be written as $\tilde{\boldsymbol{c}} = \beta \tilde{\boldsymbol{A}}_1^{-1} \boldsymbol{s}_1$, where $\beta$ is a positive factor. The SIR of the desired (first) user is given by

$$\begin{aligned} \mathrm{SIR}_1 &= \frac{p_1 (\tilde{\boldsymbol{c}}^\top \boldsymbol{s}_1)^2}{\sum_{j\neq 1} p_j (\tilde{\boldsymbol{c}}^\top \boldsymbol{s}_j)^2 + \sigma^2 (\tilde{\boldsymbol{c}}^\top \tilde{\boldsymbol{c}})} \\ &= \frac{p_1 \tilde{\boldsymbol{c}}^\top \boldsymbol{s}_1 \boldsymbol{s}_1^\top \tilde{\boldsymbol{c}}}{\tilde{\boldsymbol{c}}^\top \left( \sum_{j\neq 1} p_j \boldsymbol{s}_j \boldsymbol{s}_j^\top + \sigma^2 \boldsymbol{I} \right) \tilde{\boldsymbol{c}}} \\ &= \frac{p_1 \tilde{\boldsymbol{c}}^\top \boldsymbol{s}_1 \boldsymbol{s}_1^\top \tilde{\boldsymbol{c}}}{\tilde{\boldsymbol{c}}^\top \left( \tilde{\boldsymbol{A}}_1 + \tilde{\sigma}^2 \boldsymbol{I} \right) \tilde{\boldsymbol{c}}} \end{aligned} \tag{3.56}$$

Inserting $\tilde{\boldsymbol{c}} = \beta \tilde{\boldsymbol{A}}_1^{-1} \boldsymbol{s}_1$ yields

$$\mathrm{SIR}_1 = \frac{p_1 \left( \boldsymbol{s}_1^\top \tilde{\boldsymbol{A}}_1^{-1} \boldsymbol{s}_1 \right)^2}{\boldsymbol{s}_1^\top \tilde{\boldsymbol{A}}_1^{-1} \boldsymbol{s}_1 + \tilde{\sigma}^2 \boldsymbol{s}_1^\top \tilde{\boldsymbol{A}}_1^{-2} \boldsymbol{s}_1} \tag{3.57}$$

If we denote the value of SIR of the desired (first) user obtained when $\tilde{\sigma}^2$ is used as $\mathrm{SIR}_1(\tilde{\sigma}^2)$, then it is not difficult to show that $\mathrm{SIR}_1(\tilde{\sigma}^2)$ is a decreasing function of $\tilde{\sigma}^2$. When $\tilde{\sigma}^2 = 0$, the highest value of SIR is attained. This result is expected because at $\tilde{\sigma}^2 = 0$, the detector $\tilde{\boldsymbol{c}}$ is the MMSE multiuser detector and MMSE is known to maximize the SIR among all linear detectors [30]. When $\tilde{\sigma}^2$ is increased to $\sigma^2$, the decorrelating detector is reached.

### 3.6.2 Estimating $\sigma^2$ at Every Filter Update

In this section, we will investigate the convergence of the stochastic decorrelating detector algorithm where a new random unbiased estimate of $\sigma^2$ is used in every update of the filter coefficients. Let $\hat{\sigma}^2(n)$ denote the estimate of $\sigma^2$ at iteration $n$. Since we assume that the estimate is unbiased $E[\hat{\sigma}^2(n)] = \sigma^2$. As discussed earlier, one possible way of obtaining such an estimate is to allocate a filter $\boldsymbol{v}$ that is orthogonal to all signatures being used. Then the square of the output of the filter is an unbiased estimate for $\sigma^2$. In this section we will assume: $\hat{\sigma}^2 = (\boldsymbol{v}^\top \boldsymbol{n})^2$.

The filter update equation is given as:

$$\boldsymbol{c}(n+1) = \boldsymbol{c}(n) - \mu \left[ \left( \boldsymbol{r}(n)\boldsymbol{r}^\top(n) - \hat{\sigma}^2(n)\boldsymbol{I} \right) \boldsymbol{c}(n) - \boldsymbol{s}_1 \right] \tag{3.58}$$

The fixed point of this algorithm would be $\hat{\boldsymbol{c}}$ for which the expected value of the correction term on the right hand side equals zero. That is, $\boldsymbol{A}\hat{\boldsymbol{c}} = \boldsymbol{s}_1$, and $\hat{\boldsymbol{c}}$ is the decorrelating detector. From (3.58), we have

$$\begin{aligned} \boldsymbol{c}(n+1) &= \boldsymbol{c}(n) - \mu \left[ \left( \boldsymbol{r}(n)\boldsymbol{r}^\top(n) - \boldsymbol{A} + \boldsymbol{A} - \hat{\sigma}^2(n)\boldsymbol{I} \right) \boldsymbol{c}(n) - \boldsymbol{s}_1 \right] \\ &= \boldsymbol{c}(n) - \mu \left[ \boldsymbol{A} \left( \boldsymbol{c}(n) - \hat{\boldsymbol{c}} \right) + \left( \boldsymbol{r}(n)\boldsymbol{r}^\top(n) - \boldsymbol{A} - \hat{\sigma}^2 \boldsymbol{I} \right) \boldsymbol{c}(n) \right] \\ &= \boldsymbol{c}(n) - \mu \left[ \boldsymbol{A} \left( \boldsymbol{c}(n) - \hat{\boldsymbol{c}} \right) + \boldsymbol{\theta}(n) \right] \end{aligned} \tag{3.59}$$

where $\boldsymbol{\theta}(n) = (\boldsymbol{r}(n)\boldsymbol{r}^\top(n) - \boldsymbol{A} - \hat{\sigma}^2\boldsymbol{I})\boldsymbol{c}(n)$. Subtracting $\hat{\boldsymbol{c}}$ from both sides of the final equation in (3.59) and defining $\hat{\boldsymbol{\epsilon}}(n) = \boldsymbol{c}(n) - \hat{\boldsymbol{c}}$ yields

$$\hat{\boldsymbol{\epsilon}}(n+1) = \hat{\boldsymbol{\epsilon}}(n) - \mu \left[\boldsymbol{A}\hat{\boldsymbol{\epsilon}}(n) + \boldsymbol{\theta}(n)\right] \tag{3.60}$$

Comparing (3.60) with (3.18), we see that Lemma 2 applied to $\hat{\boldsymbol{\epsilon}}(n)$ is still valid. In order to complete the convergence proof we need the following lemma similar to Lemma 3.

**Lemma 5** *There exist* $0 < c_0 \leq c_1 < \infty$, *such that*

$$0 \leq E\left[\|\boldsymbol{\theta}(n)\|^2 \mid \hat{\boldsymbol{\epsilon}}(n) = \boldsymbol{\epsilon}\right] \leq c_0 + c_1 \|\boldsymbol{\epsilon}\|^2 \tag{3.61}$$

The proof of Lemma 5 can be found in Appendix 3.A.

Thus, if an unbiased estimate for the noise variance is available, the proposed algorithm that uses the random noise variance estimates in every filter update converges to the decorrelating detector in the MSE sense.

## 3.7 Simulation Results

In the simulations, we consider a synchronous CDMA system with $N = 6$ users using randomly generated signature sequences with a processing gain of $G = 31$. Powers of the users are assigned so that the final signal to interference ratio (SIR) of all users will be 20 dB. Thus the power of the $i$th user is found by $p_i = 100\sigma^2[\boldsymbol{\Gamma}^{-1}]_{ii}$ where $\boldsymbol{\Gamma} = \boldsymbol{S}^\top \boldsymbol{S}$ is the cross correlation matrix. The algorithm is run for 100 times and averaged results are plotted in the figures. In Figure 3.2, we plot the averaged normalized squared error (NSE) of the filter coefficients of the desired (first) user versus iteration index $n$, where NSE at iteration $n$ is defined as

$$\text{NSE}(n) = \frac{\|\boldsymbol{c}(n) - \boldsymbol{c}^*\|^2}{\|\boldsymbol{c}^*\|^2} \tag{3.62}$$

The curves in Figure 3.2 correspond to the blind adaptive decorrelating detector algorithms with various fixed step sizes and that with a time dependent step size sequence of the form of $a_n = 1/(n + n_0)$ with $n_0 = 5$. We observe that if the step size $\mu$ is too large, then the algorithm does not converge; see increasing NSE curve for $\mu = 0.1$ in Figure 3.2. We also observe in the curves corresponding to $\mu = 0.01$ and $\mu = 0.001$ in

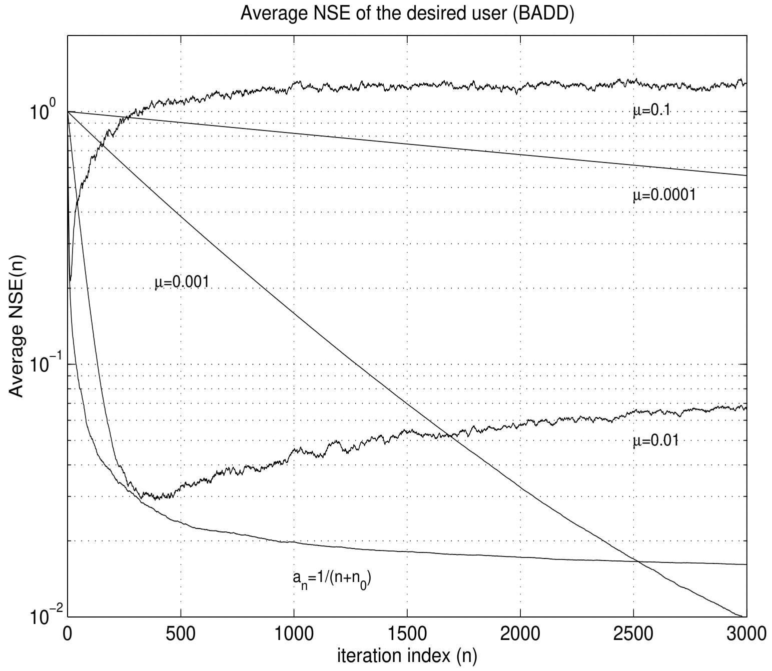


Figure 3.2: Averaged normalized squared error (NSE) of the desired user.

Figure 3.2, that for a large step size, the convergence rate is fast but the limiting NSE is large too.

Averaged (over 100 runs) SIR of the desired (first) user is plotted in Figure 3.3 for the same system. At iteration $n$, the SIR of the desired user is calculated as

$$\mathrm{SIR}(n) = \frac{p_1(\boldsymbol{s}_1^\top \boldsymbol{c}(n))^2}{\sum_{j\neq 1} p_j(\boldsymbol{s}_j^\top \boldsymbol{c}(n))^2 + \sigma^2(\boldsymbol{c}^\top(n)\boldsymbol{c}(n))} \tag{3.63}$$

We observe that the convergence of the SIRs to the *target* SIR (which is 20 dB in this experiment) is quantitatively similar to the convergence of the NSE to zero.

The blind adaptive decorrelating detector of [69] is also implemented for the same system and the SIR of the desired user is plotted as a function of iteration index in Figure 3.4 for different values of the forgetting factor, $\beta$ (see [72]). In comparing Figures 3.3 and 3.4, we see that in this instance the blind adaptive decorrelating detector in this work outperforms that of [69].

To verify the theoretical results of Section 3.6, we now run the proposed algorithm with a mismatch in the noise variance. First we use the approach detailed in Section 3.6.1 where we obtain an estimate for $\sigma^2$ once before the bit transmission starts and use this estimate in every filter update as in (3.47). The algorithm is run for three

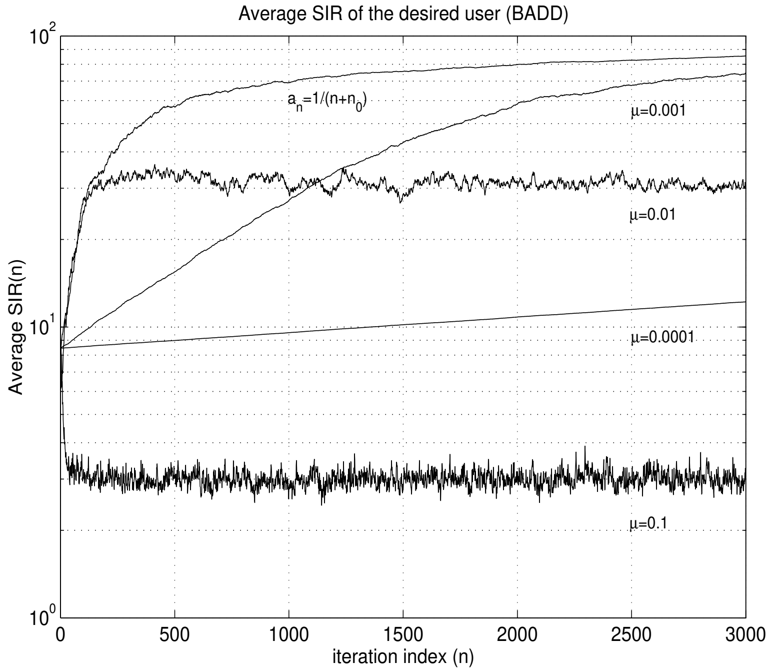


Figure 3.3: Averaged SIR of the desired user.

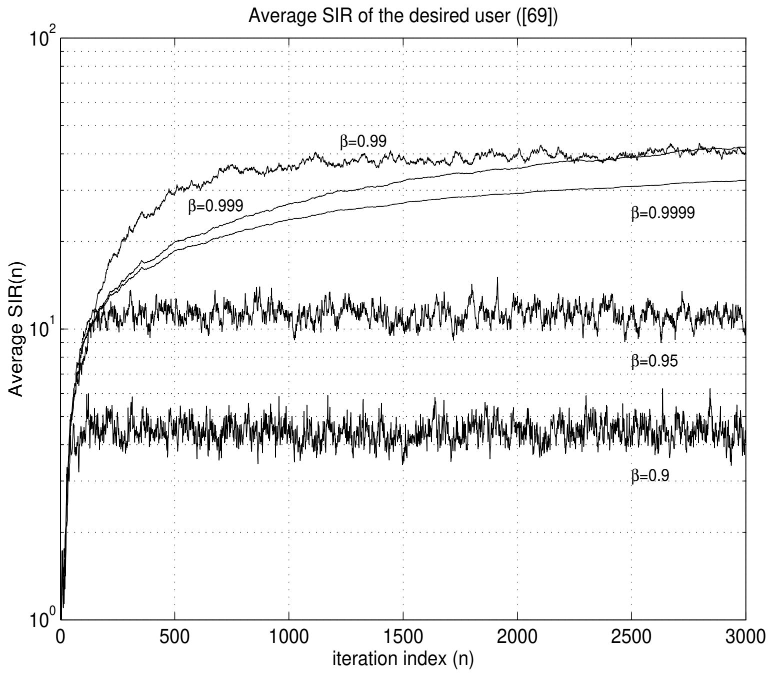


Figure 3.4: Averaged SIR of the desired user (algorithm of [69]).

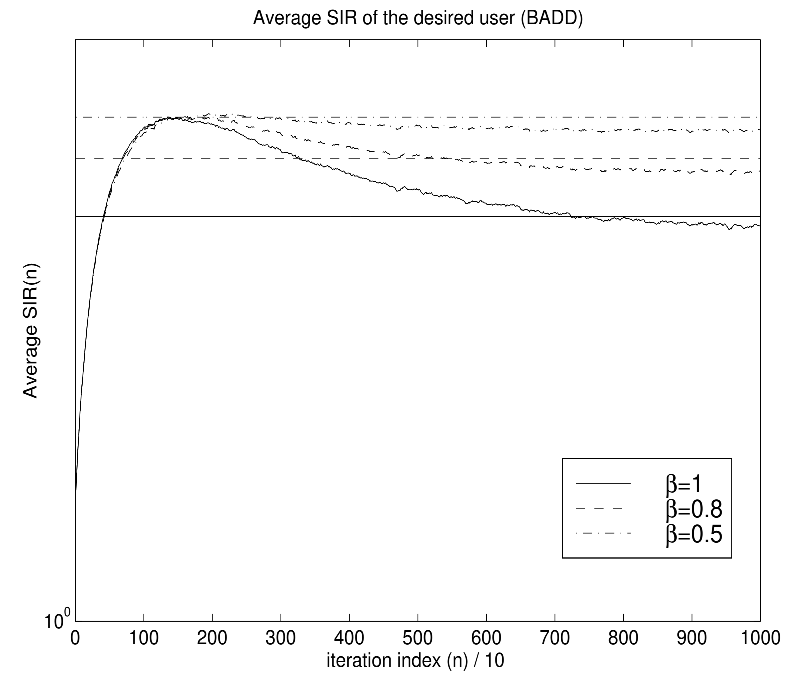


Figure 3.5: Averaged SIR of the desired user with mismatch in noise variance. Noise variance estimate is calculated only once before the filter updates. Noise variance used in the updates: $\tilde{\sigma}^2 = \beta\sigma^2$.

values of $\tilde{\sigma}^2$: $\tilde{\sigma}^2 = \beta\sigma^2$ for $\beta = 1$ (exact knowledge on $\sigma^2$), $\beta = 0.8$ (20% error in estimation) and $\beta = 0.5$ (50% error in estimation). Figure 3.5 shows the average SIR in these three cases along with the convergence point SIR calculated from (3.57) as horizontal lines. In order to make the difference where SIRs converge visible we assumed a lower target SIR of 5 dB ($\approx 3.16$) and a highly loaded system ($N = 20$) for the same processing gain ($G = 31$). We observe from Figure 3.5 that the average SIR of the desired user converges to the value predicted by the theoretical result in (3.57).

We also ran the algorithm given in (3.58) where a new estimate of the noise variance is obtained at every update of the filter coefficients. We observed that the average SIR converges to the same value when the algorithm is run with perfect knowledge of the noise variance (algorithm in (3.14)). The averaged SIRs in both cases follow almost exact trajectories. Since there is no visible difference between the SIRs in both cases, we do not include the figure here.

## 3.8 Chapter Summary and Conclusions

Multiuser detectors exploit the structure in the noise to suppress the multiaccess interference more effectively. Blind adaptive multiuser detectors are desirable in order to be able to construct the receiver filter of a desired user without knowing anything about the interfering users and without requiring training sequences which use up the already scarce bandwidth. The decorrelating detector is the linear multiuser detector which suppresses the multiaccess interference totally at the cost of enhancing the background noise.

In this chapter we presented an iterative and distributed blind adaptive decorrelating detector algorithm which is based on the observation of the available statistics, and studied its convergence. For the asynchronous CDMA channel, we showed that the filter coefficients converge to the one-shot asynchronous decorrelating detector. The update equation of the algorithm needs the observation of the chip sampled input signal before the receiver filter and the output of the filter with the current filter coefficients. For the implementation of the algorithm to construct the decorrelating receiver filter of a user only two parameters are needed: the user's signature sequence and the variance of the receiver noise. We investigated the effects of estimation errors of the noise variance on the performance of the algorithm. We suggested two possible implementations for the algorithm. In the first implementation, an estimate for the noise variance is obtained only once and this estimate is used in all filter updates, whereas in the second implementation a new estimate for the noise variance is obtained and used at each filter update. For the first case, we showed that if the estimate is smaller than the true value of the noise variance, the algorithm converges to a filter which performs between the MMSE and the decorrelating multiuser detectors in terms of the SIR. For the second case, we showed that the filter coefficients converge to the decorrelating detector as long as the estimates of the noise variance are unbiased.

In all cases, we studied the convergence of the proposed algorithm both for a fixed step size sequence and for a time varying step size sequence. For the first case we developed the conditions of having lower and upper bounds on the MSE and showed that as the step size goes to zero the algorithm converges in the MSE. For the second case, we directly proved the convergence in the MSE.

## 3.A Appendix: Additional Proofs

**Proof: Lemma 1** We need to show that if $\boldsymbol{s}_1, \ldots, \boldsymbol{s}_N$ are linearly independent, then (i) any solution of (3.4) is also a solution of (3.8), and (ii) any solution of (3.8) is also a solution of (3.4). Let $\hat{\boldsymbol{c}}$ be a solution of (3.4). Then $\boldsymbol{S}^\top \hat{\boldsymbol{c}} = \alpha \boldsymbol{e}_1$. Multiplying both sides by first $\boldsymbol{\Lambda}$ and then with $\boldsymbol{S}$ yields, $\boldsymbol{S}\boldsymbol{\Lambda}\boldsymbol{S}^\top \hat{\boldsymbol{c}} = \alpha p_1 \boldsymbol{s}_1$, and the proof of the first part is complete. Now let $\check{\boldsymbol{c}}$ be a solution of (3.8). Then $\boldsymbol{S}\boldsymbol{\Lambda}\boldsymbol{S}^\top \check{\boldsymbol{c}} = \alpha p_1 \boldsymbol{s}_1$ and equivalently

$$\boldsymbol{S}\left(\boldsymbol{\Lambda}\boldsymbol{S}^\top \check{\boldsymbol{c}} - \alpha p_1 \boldsymbol{e}_1\right) = \boldsymbol{0} \tag{3.64}$$

Note that any vector multiplying $\boldsymbol{S}$ yields a linear superposition of the columns of $\boldsymbol{S}$, i.e., $\boldsymbol{S}\boldsymbol{a} = \sum_{i=1}^{N} a_i \boldsymbol{s}_i$, for any vector $\boldsymbol{a}$. Combined with the independence of $\boldsymbol{s}_1, \ldots, \boldsymbol{s}_N$, (3.64) implies

$$\boldsymbol{\Lambda}\boldsymbol{S}^\top \check{\boldsymbol{c}} - \alpha p_1 \boldsymbol{e}_1 = \boldsymbol{0} \tag{3.65}$$

Multiplying both sides of (3.65) with $\boldsymbol{\Lambda}^{-1}$ yields $\boldsymbol{S}^\top \check{\boldsymbol{c}} = \alpha \boldsymbol{e}_1$, completing the second part of the proof. $\square$

**Proof: Lemma 2** First note that $\boldsymbol{A}$ is a positive semidefinite matrix and its eigenvalues are nonnegative. To be specific, since the rank of $\boldsymbol{A}$ is $N$ which is typically much less than $G$, $G - N$ eigenvalues of $\boldsymbol{A}$ are zero and the remaining $N$ eigenvalues are positive. Since $\boldsymbol{A}\boldsymbol{x} \in \mathcal{L}$ by (3.10), eigenvectors of $\boldsymbol{A}$ are either completely in $\mathcal{L}$ with positive eigenvectors, or completely out of $\mathcal{L}$ (meaning that their projections onto $\mathcal{L}$ are zero) with zero eigenvalues. In this case Lemma 2 is a simple result of the Rayleigh quotient [49] which states that for any symmetric matrix $\boldsymbol{M}$ and vector $\boldsymbol{x}$,

$$\lambda_{\min}^{M} \|\boldsymbol{x}\|^2 \leq \boldsymbol{x}^\top \boldsymbol{M}\boldsymbol{x} \leq \lambda_{\max}^{M} \|\boldsymbol{x}\|^2 \tag{3.66}$$

where $\lambda_{\min}^{M}$ and $\lambda_{\max}^{M}$ are the smallest and the largest eigenvalues of $\boldsymbol{M}$. Note that in (3.21) $k_1$ and $k_0$ are equal to the largest and the smallest non-zero eigenvalues of $\boldsymbol{A}$. $\square$

**Proof: Lemma 3** Using (3.16) and noting that $\boldsymbol{\epsilon}(n) = \boldsymbol{c}(n) - \boldsymbol{c}^*$, we observe that

$$E\left[\|\boldsymbol{\eta}(n)\|^2 \mid \boldsymbol{\epsilon}(n) = \boldsymbol{\epsilon}\right] = (\boldsymbol{\epsilon} + \boldsymbol{c}^*)^\top E\left[(\boldsymbol{r}\boldsymbol{r}^\top - (\boldsymbol{A} + \sigma^2 \boldsymbol{I}))^\top (\boldsymbol{r}\boldsymbol{r}^\top - (\boldsymbol{A} + \sigma^2 \boldsymbol{I}))\right] (\boldsymbol{\epsilon} + \boldsymbol{c}^*) \tag{3.67}$$

Note that $E\left[\boldsymbol{r}\boldsymbol{r}^\top\right] = \boldsymbol{A} + \sigma^2 \boldsymbol{I}$ from (3.12). Applying this result to (3.67) yields

$$E\left[\|\boldsymbol{\eta}(n)\|^2 \mid \boldsymbol{\epsilon}(n) = \boldsymbol{\epsilon}\right] = (\boldsymbol{\epsilon} + \boldsymbol{c}^*)^\top \left\{E\left[\boldsymbol{r}\boldsymbol{r}^\top \boldsymbol{r}\boldsymbol{r}^\top\right] - \left(\boldsymbol{A} + \sigma^2 \boldsymbol{I}\right)^2\right\} (\boldsymbol{\epsilon} + \boldsymbol{c}^*) \tag{3.68}$$

First we will evaluate $E\left[\boldsymbol{r}\boldsymbol{r}^\top \boldsymbol{r}\boldsymbol{r}^\top\right]$ below. Using (3.2), we note that

$$\boldsymbol{r}\boldsymbol{r}^\top = \sum_k \sum_l \sqrt{p_k}\sqrt{p_l} a_k a_l \boldsymbol{s}_k \boldsymbol{s}_l^\top + \left(\sum_k \sqrt{p_k} a_k \boldsymbol{s}_k\right) \boldsymbol{n}^\top + \boldsymbol{n} \left(\sum_k \sqrt{p_k} a_k \boldsymbol{s}_k^\top\right) + \boldsymbol{n}\boldsymbol{n}^\top \tag{3.69}$$

Therefore, $\boldsymbol{r}\boldsymbol{r}^\top \boldsymbol{r}\boldsymbol{r}^\top$ is given as

$$\begin{aligned}
&\sum_i \sum_j \sum_k \sum_l \sqrt{p_i}\sqrt{p_j}\sqrt{p_k}\sqrt{p_l} a_i a_j a_k a_l \boldsymbol{s}_i \boldsymbol{s}_j^\top \boldsymbol{s}_k \boldsymbol{s}_l^\top \\
&+ \sum_i \sum_j \sum_k \sqrt{p_i}\sqrt{p_j}\sqrt{p_k} a_i a_j a_k \boldsymbol{s}_i \boldsymbol{s}_j^\top \boldsymbol{s}_k \boldsymbol{n}^\top + \sum_i \sum_j \sum_k \sqrt{p_i}\sqrt{p_j}\sqrt{p_k} a_i a_j a_k \boldsymbol{s}_i \boldsymbol{s}_j^\top \boldsymbol{n} \boldsymbol{s}_k^\top \\
&+ \sum_i \sum_j \sqrt{p_i}\sqrt{p_j} a_i a_j \boldsymbol{s}_i \boldsymbol{s}_j^\top \boldsymbol{n}\boldsymbol{n}^\top + \sum_i \sum_j \sum_k \sqrt{p_i}\sqrt{p_j}\sqrt{p_k} a_i a_j a_k \boldsymbol{s}_i \boldsymbol{n}^\top \boldsymbol{s}_j \boldsymbol{s}_k^\top \\
&+ \sum_i \sum_j \sqrt{p_i}\sqrt{p_j} a_i a_j \boldsymbol{s}_i \boldsymbol{n}^\top \boldsymbol{s}_j \boldsymbol{n}^\top + \sum_i \sum_j \sqrt{p_i}\sqrt{p_j} a_i a_j \boldsymbol{s}_i \boldsymbol{n}^\top \boldsymbol{n} \boldsymbol{s}_j + \sum_i \sqrt{p_i} a_i \boldsymbol{s}_i \boldsymbol{n}^\top \boldsymbol{n}\boldsymbol{n}^\top \\
&+ \sum_i \sum_j \sum_k \sqrt{p_i}\sqrt{p_j}\sqrt{p_k} a_i a_j a_k \boldsymbol{n} \boldsymbol{s}_i^\top \boldsymbol{s}_j \boldsymbol{s}_k^\top + \sum_i \sum_j \sqrt{p_i}\sqrt{p_j} a_i a_j \boldsymbol{n} \boldsymbol{s}_i^\top \boldsymbol{s}_j \boldsymbol{n}^\top \\
&+ \sum_i \sum_j \sqrt{p_i}\sqrt{p_j} a_i a_j \boldsymbol{n} \boldsymbol{s}_i^\top \boldsymbol{n} \boldsymbol{s}_j^\top + \sum_i \sqrt{p_i} a_i \boldsymbol{n} \boldsymbol{s}_i^\top \boldsymbol{n}\boldsymbol{n}^\top + \boldsymbol{n}\boldsymbol{n}^\top \sum_i \sum_j \sqrt{p_i}\sqrt{p_j} a_i a_j \boldsymbol{s}_i \boldsymbol{s}_j^\top \\
&+ \boldsymbol{n}\boldsymbol{n}^\top \sum_i \sqrt{p_i} a_i \boldsymbol{s}_i \boldsymbol{n}^\top + \boldsymbol{n}\boldsymbol{n}^\top \boldsymbol{n} \sum_i \sqrt{p_i} a_i \boldsymbol{s}_i^\top + \boldsymbol{n}\boldsymbol{n}^\top \boldsymbol{n}\boldsymbol{n}^\top
\end{aligned} \tag{3.70}$$

Let $T_i$ denote the $i$th term on the right hand side of (3.70). Note that

$$\begin{aligned}
E[T_i] &= 0 \qquad i = 2, 3, 5, 8, 9, 12, 14, 15. \\
E[T_i] &= \sigma^2 \boldsymbol{A} \qquad i = 4, 6, 11, 13. \\
E[T_7] &= \sigma^2 G \boldsymbol{A} \\
E[T_{10}] &= \sigma^2 p_T \boldsymbol{I} \\
E[T_{16}] &= (G+1)\sigma^4 \boldsymbol{I} \\
E[T_1] &= \boldsymbol{A}^2 + \boldsymbol{C}
\end{aligned} \tag{3.71}$$

where $p_T > 0$ is the total receiver power, i.e., $p_T = \sum_i p_i$, and $\boldsymbol{C}$ is given as

$$\boldsymbol{C} = \sum_i \sum_{j \neq i} p_i p_j \boldsymbol{s}_i \boldsymbol{s}_i^\top + \sum_i \sum_{j \neq i} p_i p_j \boldsymbol{s}_i \boldsymbol{s}_i^\top \boldsymbol{s}_j \boldsymbol{s}_j^\top \tag{3.72}$$

Note that

$$\boldsymbol{C} = \frac{1}{2}\sum_{i}\sum_{j\neq i} p_i p_j \left(\boldsymbol{s}_i\boldsymbol{s}_i^\top + \boldsymbol{s}_j\boldsymbol{s}_j^\top\right)^2 \tag{3.73}$$

Thus $\boldsymbol{C}$ is positive semidefinite and symmetric. Combining the results in (3.71) yields

$$E\left[\boldsymbol{r}\boldsymbol{r}^\top\boldsymbol{r}\boldsymbol{r}^\top\right] = \boldsymbol{A}^2 + \boldsymbol{C} + (G+4)\sigma^2\boldsymbol{A} + \sigma^2 p_T\boldsymbol{I} + (G+1)\sigma^4\boldsymbol{I} \tag{3.74}$$

Inserting (3.74) into (3.68) and defining

$$\boldsymbol{B} = \boldsymbol{C} + (G+2)\sigma^2\boldsymbol{A} + \sigma^2 p_T\boldsymbol{I} + G\sigma^4\boldsymbol{I} \tag{3.75}$$

yields

$$E\left[\|\boldsymbol{\eta}(n)\|^2 \mid \boldsymbol{\epsilon}(n) = \boldsymbol{\epsilon}\right] = (\boldsymbol{\epsilon} + \boldsymbol{c}^*)^\top\boldsymbol{B}(\boldsymbol{\epsilon} + \boldsymbol{c}^*) \tag{3.76}$$

Note that $\boldsymbol{B}$ is a positive definite matrix since $\boldsymbol{A}$ and $\boldsymbol{C}$ are positive semidefinite and $\sigma > 0$. Using Rayleigh quotient [49], and denoting the smallest and largest eigenvalues of $\boldsymbol{B}$ by $\lambda_{\min}$ and $\lambda_{\max}$, we obtain

$$\lambda_{\min}\|\boldsymbol{\epsilon} + \boldsymbol{c}^*\|^2 \le E\left[\|\boldsymbol{\eta}(n)\|^2 \mid \boldsymbol{\epsilon}(n) = \boldsymbol{\epsilon}\right] \le \lambda_{\max}\|\boldsymbol{\epsilon} + \boldsymbol{c}^*\|^2 \tag{3.77}$$

Using the fact that $\|\boldsymbol{a}+\boldsymbol{b}\|^2 \le 2(\|\boldsymbol{a}\|^2 + \|\boldsymbol{b}\|^2)$ for any two vectors $\boldsymbol{a}$ and $\boldsymbol{b}$, we obtain the desired result

$$0 \le E\left[\|\boldsymbol{\eta}(n)\|^2 \mid \boldsymbol{\epsilon}(n) = \boldsymbol{\epsilon}\right] \le c_0 + c_1\|\boldsymbol{\epsilon}\|^2 \tag{3.78}$$

where $c_0 = 2\lambda_{\max}\|\boldsymbol{c}^*\|^2$ and $\boldsymbol{c}_1 = 2\lambda_{\max}$. $\square$

**Proof: Lemma 4** We need only the upper bound given in (3.31). Replacing $\mu$ with $a_n$ we obtain

$$b_{n+1} \le \left(1 - 2a_nk_0 + (k_1^2 + c_1)a_n^2\right)b_n + c_0a_n^2 \tag{3.79}$$

We will follow Sakrison's approach [43, pages 60–61] in the following derivation. Since $a_n$ is a monotonically decreasing sequence, there exists $n_0$ and $\delta$, $0 < \delta < 2$, such that for $n \ge n_0$,

$$\left(1 - 2k_0a_n + (k_1^2 + c_1)a_n^2\right) \le (1 - (2-\delta)k_0a_n) \tag{3.80}$$

Furthermore, we can choose $n_0$ such that for $n \ge n_0$, we have $(1 - (2-\delta)k_0a_n) > 0$. For $n \ge n_0$, the inequality (3.79) can be further bounded as,

$$b_{n+1} \le (1 - (2-\delta)k_0a_n)\,b_n + a_n^2c_0 \tag{3.81}$$

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

$$b_n \le b_{n_0}\beta_{n_0,n-1} + c_0 \sum_{k=n_0}^{n-1} a_k^2 \beta_{k+1,n-1} \tag{3.82}$$

where

$$\beta_{k,n} = \begin{cases} \prod_{j=k}^{n} [1 - a_j(2-\delta)k_0] & 0 \le k \le n \\ 1 & k > n \end{cases} \tag{3.83}$$

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

$$\beta_{k,n} = \exp\left(\sum_{j=k}^{n} \ln[1 - a_j(2-\delta)k_0]\right) \le \exp\left\{-(2-\delta)k_0 \sum_{j=k}^{n} a_j\right\} \tag{3.84}$$

By the first condition in (3.39) and the fact that $(2-\delta)k_0 > 0$, the exponent in the above equation diverges to negative infinity and we have,

$$\lim_{n\to\infty} \beta_{k,n-1} = 0 \tag{3.85}$$

This implies that $b_{n_0}\beta_{n_0,n-1}$ on the right hand side of (3.82) 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,

$$\begin{aligned} \lim_{n\to\infty} c_0 \sum_{k=n_0}^{n-1} a_k^2 \beta_{k+1,n-1} &= c_0 \lim_{n\to\infty} \sum_{k=n_0}^{\infty} a_k^2 \beta_{k+1,n-1} u(n-1-k) && (3.86) \\ &= c_0 \sum_{k=n_0}^{\infty} a_k^2 \lim_{n\to\infty} \beta_{k+1,n-1} u(n-1-k) && (3.87) \\ &= 0 && (3.88) \end{aligned}$$

since $\lim_{n\to\infty} \beta_{k,n-1} = 0$ from (3.85). We could exchange the limit and summation to obtain (3.87) from (3.86), because the sum on the right side of (3.86) is absolutely convergent. Finally, combining the results in (3.85) and (3.88) and the fact that $b_n$ is a non-negative sequence, we obtain $\lim_{n\to\infty} b_n = 0$ which concludes the proof. □

**Proof: Lemma 5** Note that $\boldsymbol{\theta}(n)$ can be written as

$$\begin{aligned} \boldsymbol{\theta}(n) &= \left(\boldsymbol{r}(n)\boldsymbol{r}^\top(n) - \boldsymbol{A} - \hat{\sigma}^2\boldsymbol{I}\right)\boldsymbol{c}(n) \\ &= \left(\boldsymbol{r}(n)\boldsymbol{r}^\top(n) - (\boldsymbol{A} + \sigma^2\boldsymbol{I}) - (\hat{\sigma}^2 - \sigma^2)\boldsymbol{I}\right)\boldsymbol{c}(n) \\ &= \left(\boldsymbol{r}(n)\boldsymbol{r}^\top(n) - (\boldsymbol{A} + \sigma^2\boldsymbol{I}) - x\boldsymbol{I}\right)\boldsymbol{c}(n) && (3.89) \end{aligned}$$

where $x = \hat{\sigma}^2 - \sigma^2$ is a zero mean random variable. Note that $x$ is given as $x = (\boldsymbol{v}^\top \boldsymbol{n})^2 - \sigma^2$ where $\boldsymbol{v}^\top \boldsymbol{s}_i = 0$ for all $i$. By steps similar to those in the proof of Lemma 3, we can show that $E\left[\|\boldsymbol{\theta}(n)\|^2 \mid \hat{\boldsymbol{\epsilon}}(n) = \boldsymbol{\epsilon}\right]$ equals

$$(\boldsymbol{\epsilon} + \hat{\boldsymbol{c}})^\top \left\{E\left[\boldsymbol{r}\boldsymbol{r}^\top \boldsymbol{r}\boldsymbol{r}^\top\right] - \left(\boldsymbol{A} + \sigma^2 \boldsymbol{I}\right)^2 - 2E\left[x\boldsymbol{r}\boldsymbol{r}^\top\right] + E\left[x^2\right]\boldsymbol{I}\right\}(\boldsymbol{\epsilon} + \hat{\boldsymbol{c}}) \tag{3.90}$$

From the proof of Lemma 3, we have

$$E\left[\boldsymbol{r}\boldsymbol{r}^\top \boldsymbol{r}\boldsymbol{r}^\top\right] - \left(\boldsymbol{A} + \sigma^2 \boldsymbol{I}\right)^2 = \boldsymbol{B} \tag{3.91}$$

where $\boldsymbol{B}$ is a positive definite matrix. We need to evaluate $E[x\boldsymbol{r}\boldsymbol{r}^\top]$ and $E[x^2]$. From (3.69) and the fact that $E[x] = 0$, we can see that $E[x\boldsymbol{r}\boldsymbol{r}^\top] = E[x\boldsymbol{n}\boldsymbol{n}^\top]$. Then it is not difficult to show that

$$E\left[x\boldsymbol{r}\boldsymbol{r}^\top\right] = E\left[\left\{(\boldsymbol{v}^\top \boldsymbol{n})^2 - \sigma^2\right\}\boldsymbol{n}\boldsymbol{n}^\top\right] = 2\sigma^4 \boldsymbol{v}\boldsymbol{v}^\top \tag{3.92}$$

Similarly, it can be shown that

$$E\left[x^2\right] = 2\sigma^4 \tag{3.93}$$

Inserting (3.91), (3.92) and (3.93) into (3.90) yields

$$E\left[\|\boldsymbol{\theta}(n)\|^2 \mid \hat{\boldsymbol{\epsilon}}(n) = \boldsymbol{\epsilon}\right] = (\boldsymbol{\epsilon} + \hat{\boldsymbol{c}})^\top \left\{\boldsymbol{B} - 4\sigma^4 \boldsymbol{v}\boldsymbol{v}^\top + 2\sigma^4 \boldsymbol{I}\right\}(\boldsymbol{\epsilon} + \hat{\boldsymbol{c}}) \tag{3.94}$$

Note that $\boldsymbol{P}\hat{\boldsymbol{\epsilon}}(n) \neq \boldsymbol{0}$, thus $\hat{\boldsymbol{\epsilon}} \in \mathcal{L}$ almost surely, and $(\boldsymbol{\epsilon}+\hat{\boldsymbol{c}})^\top \boldsymbol{v}\boldsymbol{v}^\top(\boldsymbol{\epsilon}+\hat{\boldsymbol{c}}) = [(\boldsymbol{\epsilon}+\hat{\boldsymbol{c}})^\top \boldsymbol{v}]^2 = 0$ when $\boldsymbol{P}\hat{\boldsymbol{\epsilon}}(n) \neq \boldsymbol{0}$. Thus, (3.94) can be written as

$$E\left[\|\boldsymbol{\theta}(n)\|^2 \mid \hat{\boldsymbol{\epsilon}}(n) = \boldsymbol{\epsilon}\right] = (\boldsymbol{\epsilon} + \hat{\boldsymbol{c}})^\top \left\{\boldsymbol{B} + 2\sigma^4 \boldsymbol{I}\right\}(\boldsymbol{\epsilon} + \hat{\boldsymbol{c}}) \tag{3.95}$$

Since $\boldsymbol{B}$ is positive definite, the desired result follows. □

# Chapter 4

# Stochastic Power Control for Cellular Radio Systems

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, 16, 36, 78] 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 [8, 12, 35, 73, 77, 79]. Variations of the ordinary power control problem can be found in references [19, 24, 25, 74].

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 signal to interference ratio, received interference power, or bit error rate. None of those quantities can be obtained perfectly, and deterministic convergence results are no longer valid when these deterministic variables are replaced with their random estimates.

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.

For a CDMA system based on IS-95, [6] simulates a stochastic power control algorithm that results from using random SIR estimates in an otherwise deterministic iteration. The mapping between an available observation and actual SIR needed in the power control updates is also determined by simulation. In this chapter, we will work with a simpler system model, propose an observation based power control algorithm, and prove its convergence analytically. Thus, the main contribution of this work is to present practical power control algorithms with provable convergence.

Starting from a simple extension of the deterministic power control algorithms given in [12,73], this chapter 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 users satisfy their SIR requirements. Conditions for obtaining lower and upper bounds on the MSE are identified.

In this chapter, 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 [41], 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 $\alpha$, 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 [5, 17, 18, 27, 43, 44].

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: the output of its own receiver filter (matched filter) at its assigned base station and its own channel gain to its assigned base station.

## 4.1 System Model

We consider the uplink of a wireless multi-cell CDMA system with a BPSK modulation scheme. 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 $l$th 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 $i$. The delay of user $j$ relative to the matched filter of user $i$ is represented by $d_{ij}$. In Figure 4.1, two possible situations are depicted. User $j$ has a positive delay relative to user $i$ and creates interference to the $l$th 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 $l$th 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 $j$: $\bar{\Gamma}_{ij}$, $\Gamma_{ij}$ and $\tilde{\Gamma}_{ij}$. If $d_{ij} \geq 0$ (case of user $j$ in Figure 4.1), then

$$\begin{aligned}
\bar{\Gamma}_{ij} &= \int_0^{d_{ij}} s_i(t)s_j(T_b - d_{ij} + t)dt \\
\Gamma_{ij} &= \int_{d_{ij}}^{T_b} s_i(t)s_j(t - d_{ij})dt \qquad (4.1) \\
\tilde{\Gamma}_{ij} &= 0
\end{aligned}$$

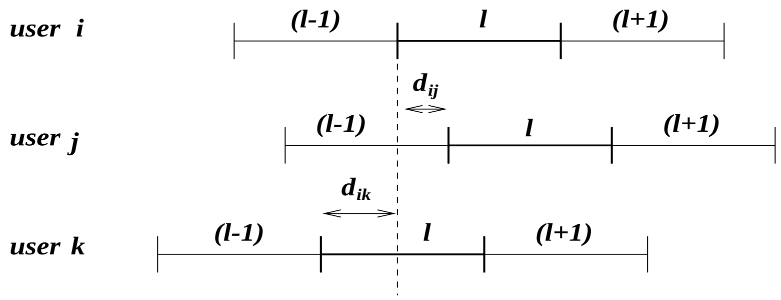


Figure 4.1: Asynchronous CDMA model.

and if $d_{ij} \leq 0$ (case of user $k$ in Figure 4.1),

$$\begin{aligned} \bar{\Gamma}_{ij} &= 0 \\ \Gamma_{ij} &= \int_0^{T_b+d_{ij}} s_i(t)s_j(t-d_{ij})dt \\ \tilde{\Gamma}_{ij} &= \int_{T_b+d_{ij}}^{T_b} s_i(t)s_j(t-T_b-d_{ij})dt \end{aligned} \tag{4.2}$$

Note from (4.1) and (4.2) that for each user either $\bar{\Gamma}_{ij}$ or $\tilde{\Gamma}_{ij}$ is equal to zero implying $\bar{\Gamma}_{ij}\tilde{\Gamma}_{ij} = 0$ 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

$$y_i(l) = \sum_{j=1}^{N} \sqrt{p_j}\sqrt{h_{ij}} \left(\bar{\Gamma}_{ij}b_j(l-1) + \Gamma_{ij}b_j(l) + \tilde{\Gamma}_{ij}b_j(l+1)\right) + n_i(l) \tag{4.3}$$

where $b_j(l)$ is the information bit of user $j$ ($\pm1$ equiprobably) in the $l$th bit interval and $n_i$ is a sample of additive white Gaussian noise (AWGN) having zero mean and $\sigma^2$ variance. Defining

$$\hat{b}_{ij}(l) = \bar{\Gamma}_{ij}b_j(l-1) + \Gamma_{ij}b_j(l) + \tilde{\Gamma}_{ij}b_j(l+1) \tag{4.4}$$

we can write (4.3) as

$$y_i(l) = \sum_{j=1}^{N} \sqrt{p_j}\sqrt{h_{ij}}\hat{b}_{ij}(l) + n_i(l) \tag{4.5}$$

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 $k = \pm1$.

## 4.2 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 $\gamma_i^*$ denote the SIR target of user $i$. The power control problem can be stated as follows:

$$\begin{array}{ll} \min & \sum_{i=1}^{N} p_i \\ \text{s.t.} & \frac{p_i h_{ii}}{\sum_{j\neq i} p_j h_{ij}\left(\bar{\Gamma}_{ij}^2+\Gamma_{ij}^2+\tilde{\Gamma}_{ij}^2\right)+\sigma^2} \geq \gamma_i^* \qquad \text{for } i=1,\ldots,N \end{array} \tag{4.6}$$

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

$$A_{ij} = \begin{cases} 0 & i=j \\ h_{ij}\left(\bar{\Gamma}_{ij}^2+\Gamma_{ij}^2+\tilde{\Gamma}_{ij}^2\right) & i\neq j \end{cases} \qquad \text{and} \qquad H_{ij} = \begin{cases} h_{ii} & i=j \\ 0 & i\neq j \end{cases} \tag{4.7}$$

the QoS requirements in (4.6) can be written as the vector inequality

$$\boldsymbol{p} \geq \boldsymbol{\Upsilon}\boldsymbol{H}^{-1}(\boldsymbol{A}\boldsymbol{p}+\sigma^2\boldsymbol{u}) \tag{4.8}$$

where $\mathbf{u} = [1\cdots 1]^\top$. We say that the set of SIR targets $\gamma_i^*$, $i = 1,\cdots,N$ are feasible if there is a non-negative finite vector $\boldsymbol{p}$ that satisfies (4.8).

It is not difficult to show that if the SIR targets $\gamma_i^*$, $i = 1,\cdots,N$ are feasible, then the power vector which satisfies all $N$ inequalities with equality in (4.6) minimizes the sum of the transmitted powers (see [73, 74] or the appendix of [76]). Therefore, if $\gamma_i^*$, $i = 1,\cdots,N$ are feasible, a power control algorithm finds (in the iterative case converges to) the solution of

$$\boldsymbol{p} = \boldsymbol{\Upsilon}\boldsymbol{H}^{-1}(\boldsymbol{A}\boldsymbol{p}+\sigma^2\boldsymbol{u}) \tag{4.9}$$

The existence of a non-negative solution to (4.9) is specified by the following two theorems on eigenvalues of non-negative matrices by Perron and Frobenius [46].

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

Thus, Equation (4.9) and Theorem 2 imply the following result.

**Lemma 6** *The SIR targets $\gamma_i^*$, $i = 1, \cdots, N$ are feasible iff $\rho < 1$ where $\rho$ is the Perron-Frobenius eigenvalues of matrix $\boldsymbol{\Upsilon}\boldsymbol{H}^{-1}\boldsymbol{A}$.*

For the remainder of this chapter, we will assume that the deterministic power control problem is feasible. From (4.9), we obtain

$$\left(\boldsymbol{I} - \boldsymbol{\Upsilon}\boldsymbol{H}^{-1}\boldsymbol{A}\right)\boldsymbol{p} = \sigma^2\boldsymbol{\Upsilon}\boldsymbol{H}^{-1}\boldsymbol{u} \tag{4.10}$$

Note that, if the relative delays, correlation coefficients and channel gains are known, then from (4.10), the optimal power vector can be found as

$$\bar{\boldsymbol{p}} = \sigma^2\boldsymbol{B}^{-1}\boldsymbol{\Upsilon}\boldsymbol{H}^{-1}\boldsymbol{u} \tag{4.11}$$

where

$$\boldsymbol{B} = \boldsymbol{I} - \boldsymbol{\Upsilon}\boldsymbol{H}^{-1}\boldsymbol{A} \tag{4.12}$$

with $\boldsymbol{I}$ the $N \times N$ identity matrix. However, using (4.11) 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 [73], iterative, distributed, deterministic power control algorithms were defined in general as,

$$\boldsymbol{p}(n+1) = \boldsymbol{T}(\boldsymbol{p}(n)) \tag{4.13}$$

When the current power vector is $\boldsymbol{p}(n)$, $T_j(\boldsymbol{p}(n))$, the $j$th component of $\boldsymbol{T}(\boldsymbol{p}(n))$, is the interference that user $j$ is required to overcome. If the SIR targets $\gamma_i^*$, $i = 1, \cdots, N$ are feasible, then the algorithm (4.13) 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 [73]. In the notation of this chapter, the interference function is

$$\boldsymbol{T}(\boldsymbol{p}) = \boldsymbol{\Upsilon}\boldsymbol{H}^{-1}(\boldsymbol{A}\boldsymbol{p} + \sigma^2\boldsymbol{u}) \tag{4.14}$$

When $\boldsymbol{T}(\boldsymbol{p})$ is of the form of (4.14), the iteration (4.13) is the power control algorithm of Foschini and Miljanic [12, Eqn. (18)]. Implicit in the iteration (4.13), 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 $\boldsymbol{T}(\boldsymbol{p})$.

## 4.3 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 $l$th bit interval in the $n$th L-bit window, $y_i(n,l)$. From (4.3), squaring the matched filter output $y_i(n,l)$ yields

$$v_i(n,l) = y_i^2(n,l) = \sum_{j=1}^{N} \left(\bar{\Gamma}_{ij}^2 + \Gamma_{ij}^2 + \tilde{\Gamma}_{ij}^2\right) h_{ij} p_j(n) + \sigma^2 + w_i(n,l) \tag{4.15}$$

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

$$\begin{aligned} w_i(n,l) \quad &= \quad \sum_{j=1}^{N} \sqrt{p_j}\sqrt{h_{ij}}\Gamma_{ij} b_j(l) \left(\bar{\Gamma}_{ij} b_j(l-1) + \tilde{\Gamma}_{ij} b_j(l+1)\right) \\ &\quad + \sum_{j=1}^{N} \sum_{k \neq j} \sqrt{p_j}\sqrt{p_k}\sqrt{h_{ij}}\sqrt{h_{ik}}\hat{b}_{ij}(l)\hat{b}_{ik}(l) + 2n_i(l) \sum_{j=1}^{N} \sqrt{p_j}\sqrt{h_{ij}}\hat{b}_j(l) \\ &\quad + \left(n_i^2(l) - \sigma^2\right) \end{aligned} \tag{4.16}$$

where we dropped the common index $n$ from the terms in (4.16) for convenience. In a more careful statement of (4.16) $p_j$, $b_j(l)$, $\hat{b}_{ij}(l)$ 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 $n$th $L$-bit interval as

$$v_i(n) = \frac{1}{L}\sum_{l=1}^{L} v_i(n,l) = \sum_{j=1}^{N} \left(\bar{\Gamma}_{ij}^2 + \Gamma_{ij}^2 + \tilde{\Gamma}_{ij}^2\right) h_{ij} p_j(n) + \sigma^2 + w_i(n) \tag{4.17}$$

where

$$w_i(n) = \frac{1}{L}\sum_{l=1}^{L} w_i(n,l) \tag{4.18}$$

Note that in deriving (4.17) we assumed that the power levels of the interfering users do not change in the $L$-bit window of the $i$th 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 $i$th elements, and using power vector $\boldsymbol{p}$, (4.17) becomes

$$\boldsymbol{v} = (\boldsymbol{A} + \boldsymbol{H})\boldsymbol{p} + \sigma^2 \boldsymbol{u} + \boldsymbol{w} \tag{4.19}$$

From (4.16), it follows that $E[w_i(l)] = 0$ since the transmitted bits $(b_j(l))$ are independent and equiprobably $\pm 1$, and the zero mean noise $n_i$ is independent of the bits and has variance $\sigma^2$. Hence,

$$E[\boldsymbol{w}] = \boldsymbol{0} \tag{4.20}$$

$$E[\boldsymbol{v}] = (\boldsymbol{A} + \boldsymbol{H})\boldsymbol{p} + \sigma^2 \boldsymbol{u} \tag{4.21}$$

From (4.21) 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 (4.21) to (4.14) we can express the interference $\boldsymbol{T}(\boldsymbol{p})$ in terms of $E[\boldsymbol{v}]$ as

$$\boldsymbol{T}(\boldsymbol{p}) = -\boldsymbol{\Upsilon}\boldsymbol{p} + \boldsymbol{\Upsilon}\boldsymbol{H}^{-1}E[\boldsymbol{v}] \tag{4.22}$$

Equation (4.22) effectively subtracts the signal component from the total received power to obtain the interference. In [73], the following deterministic power control algorithm was given.

$$\boldsymbol{p}(n+1) = (1-\epsilon)\boldsymbol{p}(n) + \epsilon \boldsymbol{T}(\boldsymbol{p}(n)) \tag{4.23}$$

The algorithm (4.23) 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 [73] 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 (4.22) into (4.23), the interference averaging algorithm becomes

$$\boldsymbol{p}(n+1) = (1-\epsilon)\boldsymbol{p}(n) + \epsilon\left(-\boldsymbol{\Upsilon}\boldsymbol{p}(n) + \boldsymbol{\Upsilon}\boldsymbol{H}^{-1}E[\boldsymbol{v}(n)]\right) \tag{4.24}$$

The deterministic iteration (4.24) 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 (4.24) by the unbiased estimate $\boldsymbol{v}$.

$$\boldsymbol{p}(n+1) = (1-\epsilon)\boldsymbol{p}(n) + \epsilon\left(-\boldsymbol{\Upsilon}\boldsymbol{p}(n) + \boldsymbol{\Upsilon}\boldsymbol{H}^{-1}\boldsymbol{v}(n)\right) \tag{4.25}$$

We note that (4.25) is a special case of the following more general algorithm

$$\boldsymbol{p}(n+1) = (1-a_n)\boldsymbol{p}(n) + a_n\left(-\boldsymbol{\Upsilon}\boldsymbol{p}(n) + \boldsymbol{\Upsilon}\boldsymbol{H}^{-1}\boldsymbol{v}(n)\right) \tag{4.26}$$

where the fixed step size $\epsilon$ is replaced with a variable step size sequence $a_n$ that may be a function of the iteration index $n$. Therefore, (4.25) corresponds to a special case of the algorithm (4.26) when $a_n = \epsilon$ for all $n$. The iteration (4.26) can be re-written in the form

$$\boldsymbol{p}(n+1) = \left[1 - a_n\left(\boldsymbol{I} + \boldsymbol{\Upsilon}\right)\right]\boldsymbol{p}(n) + a_n\boldsymbol{\Upsilon}\boldsymbol{H}^{-1}\boldsymbol{v}(n) \tag{4.27}$$

Before investigating the convergence properties of the stochastic power control algorithm (4.27), we write it componentwise by using the definitions of $\boldsymbol{H}$, $\boldsymbol{\Upsilon}$ and $\boldsymbol{v}(n)$ as

$$p_i(n+1) = \left[1 - a_n(1+\gamma_i^*)\right]p_i(n) + a_n\frac{\gamma_i^*}{h_{ii}}\left(\frac{1}{L}\sum_{l=1}^{L} y_i^2(n,l)\right) \tag{4.28}$$

Equation (4.28) defines the power update rule for user $i$. As seen from (4.28), 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 (4.28): the users power value in the previous iteration $p_i(n)$, its SIR target value $\gamma_i^*$, and step size sequence $a_n$ are trivially known by the user.

Also seen from (4.28) 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 (4.28). As we shall see, the convergence proof for (4.28) 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 (4.28) will be stochastic and will be specified in terms of the mean squared error at iteration $n$,

$$\mathrm{MSE}(n) = E\left[\|\boldsymbol{p}(n) - \bar{\boldsymbol{p}}\|^2\right] \tag{4.29}$$

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 $\epsilon$ 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 $\epsilon$ 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.

## 4.4 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 (4.27) (equivalent componentwise representation was given in (4.28)). Equation (4.27) can equivalently be written as

$$\boldsymbol{p}(n+1) = \boldsymbol{p}(n) - a_n\Big[\boldsymbol{p}(n) + \boldsymbol{\Upsilon}\boldsymbol{p}(n) - \boldsymbol{\Upsilon}\boldsymbol{H}^{-1}\boldsymbol{v}(n)\Big] \tag{4.30}$$

From (4.19), at time $n$, we have

$$\boldsymbol{v}(n) = (\boldsymbol{A} + \boldsymbol{H})\boldsymbol{p}(n) + \sigma^2\boldsymbol{u} + \boldsymbol{w}(n) \tag{4.31}$$

Applying (4.31) to (4.30), we obtain

$$\boldsymbol{p}(n+1) = \boldsymbol{p}(n) - a_n\Big[\boldsymbol{B}\boldsymbol{p}(n) - \sigma^2\boldsymbol{\Upsilon}\boldsymbol{H}^{-1} - \boldsymbol{\Upsilon}\boldsymbol{H}^{-1}\boldsymbol{w}(n)\Big] \tag{4.32}$$

where $\boldsymbol{B}$ was defined in (4.12). It will be mathematically convenient to define

$$\boldsymbol{\eta}(n) = \boldsymbol{\Upsilon}\boldsymbol{H}^{-1}\boldsymbol{w}(n) \tag{4.33}$$

Note that $\boldsymbol{\eta}(n)$ represents a normalized form of the random component of the noise contribution $\boldsymbol{w}(n)$. Also note from (4.20) that $E[\boldsymbol{\eta}(n)] = 0$. Inserting (4.33) to (4.32) and observing from (4.11) that $\boldsymbol{B}\bar{\boldsymbol{p}} = \sigma^2\boldsymbol{\Upsilon}\boldsymbol{H}^{-1}\boldsymbol{u}$, we obtain

$$\boldsymbol{p}(n+1) = \boldsymbol{p}(n) - a_n[\boldsymbol{B}\,(\boldsymbol{p}(n) - \bar{\boldsymbol{p}}) - \boldsymbol{\eta}(n)] \tag{4.34}$$

As stated in the previous section, we will prove the convergence in the mean squared sense. The norm used in (4.29) 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 4.A.

In Appendix 4.B, we prove the following lemma as a simple consequence of the Rayleigh quotient (see (4.72)).

**Lemma 7** *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 7 verifies that it is sufficient to prove convergence for a $G$-norm with a suitably chosen symmetric matrix $\boldsymbol{G}$. From (4.34), we see that convergence will depend on the properties of the matrix $\boldsymbol{B}$. In Appendix 4.B, we verify the following result.

**Lemma 8** *Matrix* $-\boldsymbol{B}$ *is stable iff the deterministic power control problem (4.6) 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 (4.34) yields

$$\boldsymbol{x}(n+1) = \boldsymbol{x}(n) - a_n[\boldsymbol{B}\boldsymbol{x}(n) - \boldsymbol{\eta}(n)] \tag{4.35}$$

Taking and squaring the $G$-norms of both sides of (4.35) we obtain,

$$\begin{aligned} \boldsymbol{x}(n+1)^\top \boldsymbol{G}\boldsymbol{x}(n+1) &= \boldsymbol{x}(n)^\top \boldsymbol{G}\boldsymbol{x}(n) - a_n \boldsymbol{x}(n)^\top \Big[\boldsymbol{G}\boldsymbol{B} + \boldsymbol{B}^\top \boldsymbol{G}\Big]\boldsymbol{x}(n) \\ &\quad + 2a_n \boldsymbol{x}(n)^\top \boldsymbol{G}\boldsymbol{\eta}(n) - 2a_n^2 \boldsymbol{x}(n)^\top \boldsymbol{B}^\top \boldsymbol{G}\boldsymbol{\eta}(n) \\ &\quad + a_n^2 \boldsymbol{x}(n)^\top \boldsymbol{B}^\top \boldsymbol{G}\boldsymbol{B}\boldsymbol{x}(n) + a_n^2 \boldsymbol{\eta}(n)^\top \boldsymbol{G}\boldsymbol{\eta}(n) \end{aligned} \tag{4.36}$$

By taking the conditional expectation of both sides of (4.36), conditioned on $\boldsymbol{x}(n) = \boldsymbol{x}$, and observing that $E[\boldsymbol{\eta}(n) \mid \boldsymbol{x}(n) = \boldsymbol{x}] = \boldsymbol{0}$, we obtain

$$\begin{aligned} E\Big[\|\boldsymbol{x}(n+1)\|_G^2 \mid \boldsymbol{x}(n) = \boldsymbol{x}\Big] &= \|\boldsymbol{x}\|_G^2 - a_n \boldsymbol{x}^\top \Big[\boldsymbol{G}\boldsymbol{B} + \boldsymbol{B}^\top \boldsymbol{G}\Big]\boldsymbol{x} + a_n^2 \boldsymbol{x}^\top \boldsymbol{B}^\top \boldsymbol{G}\boldsymbol{B}\boldsymbol{x} \\ &\quad + a_n^2 E\Big[\|\boldsymbol{\eta}(n)\|_G^2 \mid \boldsymbol{x}(n) = \boldsymbol{x}\Big] \end{aligned} \tag{4.37}$$

Our proof of convergence will proceed by bounding the individual terms on the right hand side of (4.37). As proven in Lemma 8, the feasibility of the deterministic power control problem guarantees the stability of $-\boldsymbol{B}$. By Theorem 3 in Appendix 4.A, stability of matrix $-\boldsymbol{B}$, in turn, guarantees the existence of a symmetric, positive definite matrix $\boldsymbol{G}$ as a solution of

$$\boldsymbol{G}\boldsymbol{B} + \boldsymbol{B}^\top \boldsymbol{G} = \boldsymbol{C} \tag{4.38}$$

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 (4.37) by using Lemma 11 in

Appendix 4.A and the fact that both $\boldsymbol{GB}+\boldsymbol{B}^\top\boldsymbol{G}$ and $\boldsymbol{B}^\top\boldsymbol{GB}$ are symmetric, positive definite matrices.

**Lemma 9** *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*

$$k_0 \|\boldsymbol{x}\|_G^2 \le \boldsymbol{x}^\top \left[\boldsymbol{GB}+\boldsymbol{B}^\top\boldsymbol{G}\right]\boldsymbol{x} \le k_0' \|\boldsymbol{x}\|_G^2 \tag{4.39}$$

$$k_1 \|\boldsymbol{x}\|_G^2 \le \boldsymbol{x}^\top \boldsymbol{B}^\top\boldsymbol{GB}\boldsymbol{x} \le k_1' \|\boldsymbol{x}\|_G^2 \tag{4.40}$$

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 9 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 4.B.

**Lemma 10** *There exist positive constants* $d_0$, $c_0$ *and* $c_1$, $0 < d_0 \le c_0 < \infty$, *such that*

$$d_0 \le E\left[\|\boldsymbol{\eta}(n)\|_G^2 \mid \boldsymbol{x}(n)=\boldsymbol{x}\right] \le c_0 + c_1 \|\boldsymbol{x}\|_G^2 \tag{4.41}$$

Using Lemmas 9 and 10, we obtain the following upper and lower bounds from (4.37)

$$E\left[\|\boldsymbol{x}(n+1)\|_G^2 \mid \boldsymbol{x}(n)=\boldsymbol{x}\right] \le \|\boldsymbol{x}\|_G^2 - a_n k_0 \|\boldsymbol{x}\|_G^2 + a_n^2\left\{k_1' \|\boldsymbol{x}\|_G^2 + c_0 + c_1 \|\boldsymbol{x}\|_G^2\right\} \tag{4.42}$$

$$E\left[\|\boldsymbol{x}(n+1)\|_G^2 \mid \boldsymbol{x}(n)=\boldsymbol{x}\right] \ge \|\boldsymbol{x}\|_G^2 - a_n k_0' \|\boldsymbol{x}\|_G^2 + a_n^2\left\{k_1 \|\boldsymbol{x}\|_G^2 + d_0\right\} \tag{4.43}$$

Taking the expectation of both sides of the final inequalities (4.42) and (4.43) with respect to $\boldsymbol{x}(n)$ and letting $b_n = E\left[\|\boldsymbol{x}(n)\|_G^2\right]$ we obtain,

$$b_{n+1} \le \left\{1 - k_0 a_n + (k_1' + c_1)a_n^2\right\} b_n + a_n^2 c_0 \tag{4.44}$$

$$b_{n+1} \ge \left\{1 - k_0'\epsilon + k_1 a_n^2\right\} b_n + a_n^2 d_0 \tag{4.45}$$

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 (4.44) and (4.45).

### 4.4.1 Convergence Results for Fixed $a_n$

We now consider the fixed step size stochastic iteration (4.25). By defining

$$\alpha_0 = 1 - k_0\epsilon + (k_1' + c_1)\epsilon^2 \qquad \alpha_1 = 1 - k_0'\epsilon + k_1\epsilon^2 \tag{4.46}$$

we can write the lower and upper bounds (4.44) and (4.45) on the non-negative sequence $b_n$ as,

$$\alpha_1 b_n + \epsilon^2 d_0 \leq b_{n+1} \leq \alpha_0 b_n + \epsilon^2 c_0 \tag{4.47}$$

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 $\epsilon$ is chosen such that $\mid \alpha_0 \mid < 1$ and $\mid \alpha_1 \mid < 1$.

We note that $\alpha_0$ and $\alpha_1$ are equal to 1 at $\epsilon = 0$. We also note that both $\alpha_0$ and $\alpha_1$ are locally decreasing as $\epsilon$ increases, since

$$\left.\frac{d\alpha_0}{d\epsilon}\right|_{\epsilon=0} = -k_0 < 0 \qquad \text{and} \qquad \left.\frac{d\alpha_1}{d\epsilon}\right|_{\epsilon=0} = -k_0' < 0 \tag{4.48}$$

This means that we can always choose a small nonzero $\epsilon$ 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

$$\frac{\epsilon^2 d_0}{1-\alpha_1} \leq \lim_{n\to\infty} b_n \leq \frac{\epsilon^2 c_0}{1-\alpha_0} \tag{4.49}$$

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 $\epsilon$, 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 7 implies that the limiting MSE goes to zero and the stochastic power control algorithm (4.25) converges to the unique optimal power vector $\bar{\boldsymbol{p}}$. However, as $\epsilon$ approaches zero, $\alpha_0$ and $\alpha_1$ approach 1 which slows the convergence rate. Thus it is undesirable to choose $\epsilon$ too small. Furthermore, it is also undesirable to choose $\epsilon$ 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 $\epsilon$ that gives rise to acceptable values of $\alpha_0$ and $\alpha_1$ 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 $n$.

### 4.4.2 Convergence Results for Varying $a_n$

In this section, we will examine the variable step size iteration (4.27). We will show that if the coefficient sequence $a_n$ satisfies the following two conditions

$$\sum_{n=1}^{\infty} a_n = \infty \qquad \text{and} \qquad \sum_{n=1}^{\infty} a_n^2 < \infty \tag{4.50}$$

then the power control algorithm (4.27) converges to the optimal power vector $\bar{\boldsymbol{p}}$ in the mean square sense. Note that $a_n = a/n$ satisfies conditions given in (4.50). It was shown in [43] 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 [43, pages 60–61] in the following derivation. We will need only the upper bound given in (4.44). Since $a_n$ is a monotonically decreasing sequence, there exists $n_0$ and $\delta$, $0 < \delta < 1$, such that for $n \geq n_0$,

$$\left(1 - k_0 a_n + (k_1' + c_1) a_n^2\right) \leq (1 - (1-\delta) k_0 a_n) \tag{4.51}$$

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 (4.44) can be further bounded as,

$$b_{n+1} \leq \{1-(1-\delta)k_0 a_n\} b_n + a_n^2 c_0 \tag{4.52}$$

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

$$b_n \leq b_{n_0}\beta_{n_0,n-1} + c_0 \sum_{k=n_0}^{n-1} a_k^2 \beta_{k+1,n-1} \tag{4.53}$$

where

$$\beta_{k,n} = \begin{cases} \prod_{j=k}^{n} [1 - a_j(1-\delta)k_0] & 0 \leq k \leq n \\ 1 & k > n \end{cases} \tag{4.54}$$

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

$$\beta_{k,n} = \exp\left(\sum_{j=k}^{n} \ln[1 - a_j(1-\delta)k_0]\right) \leq \exp\left\{-(1-\delta)k_0 \sum_{j=k}^{n} a_j\right\} \tag{4.55}$$

By the first condition in (4.50) 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 (4.53) 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,

$$\lim_{n\to\infty} c_0 \sum_{k=n_0}^{n-1} a_k^2 \beta_{k+1,n-1} = c_0 \lim_{n\to\infty} \sum_{k=n_0}^{\infty} a_k^2 \beta_{k+1,n-1} u(n-1-k) \tag{4.56}$$

$$= c_0 \sum_{k=n_0}^{\infty} a_k^2 \lim_{n\to\infty} \beta_{k+1,n-1} u(n-1-k) = 0 \tag{4.57}$$

since $\lim_{n\to\infty} \beta_{k,n-1} = 0$. We could exchange the limit and summation to obtain (4.57) from (4.56), because the sum on the right side of (4.56) is absolutely convergent. Finally, combining the result in (4.57) 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 7, we conclude that the algorithm converges to the unique global optimal power vector in the mean squared sense, i.e., $\lim_{n\to\infty} E\left[\|\boldsymbol{p}(n) - \bar{\boldsymbol{p}}\|^2\right] = 0$.

## 4.5 Discussion

Throughout this chapter 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 [20] to use larger SIR target values than needed allowing for a fade margin. In this chapter, 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 ($\gamma_i^*$, $i = 1, \cdots, N$) 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 (4.28) the modified power update equation for user $i$ can be written as

$$p_i(n+1) = [1 - a_n(1+\gamma_i^*)]\, p_i(n) + a_n \frac{\gamma_i^*}{\hat{h}_{ii}} \left( \frac{1}{L} \sum_{l=1}^{L} y_i^2(n,l) \right) \tag{4.58}$$

Let $\hat{\boldsymbol{H}}$ denote a diagonal matrix with $\hat{h}_{ii}$ as its $i$th diagonal element. The modified version of (4.30) reflecting the estimation error in the channel gains is given as

$$\boldsymbol{p}(n+1) = \boldsymbol{p}(n) - a_n \left[ \boldsymbol{p}(n) + \boldsymbol{\Upsilon} \boldsymbol{p}(n) - \boldsymbol{\Upsilon} \hat{\boldsymbol{H}}^{-1} \boldsymbol{v}(n) \right] \tag{4.59}$$

Note that the iteration (4.59) converges to a point where expected value of the term in the square brackets on the right hand side of (4.59) is equal to zero, since it is the fixed point of the iteration [41,43]. 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 [10] yields

$$E\left[ \frac{1}{\hat{h}_{ii}} \right] \geq \frac{1}{E\left[ \hat{h}_{ii} \right]} = \alpha_i \frac{1}{h_{ii}} \tag{4.60}$$

for some $\alpha_i \geq 1$. The equality in (4.60) 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 (4.21), the expected value of the term in the squared brackets on the right hand side of (4.59) is equal to

$$\boldsymbol{p} + \boldsymbol{\Upsilon}\boldsymbol{p} - \boldsymbol{\Upsilon}\boldsymbol{\Lambda}\boldsymbol{H}^{-1}\boldsymbol{A}\boldsymbol{p} - \boldsymbol{\Upsilon}\boldsymbol{\Lambda}\boldsymbol{p} - \sigma^2\boldsymbol{\Upsilon}\boldsymbol{\Lambda}\boldsymbol{H}^{-1}\boldsymbol{u} \tag{4.61}$$

Equating (4.61) to zero yields the fixed point solution

$$\left((\boldsymbol{I} + \boldsymbol{\Upsilon}(\boldsymbol{I} - \boldsymbol{\Lambda})) - \boldsymbol{\Upsilon}\boldsymbol{\Lambda}\boldsymbol{H}^{-1}\boldsymbol{A}\right)\boldsymbol{p} = \sigma^2\boldsymbol{\Upsilon}\boldsymbol{\Lambda}\boldsymbol{H}^{-1}\boldsymbol{u} \tag{4.62}$$

Defining

$$\tilde{\boldsymbol{\Upsilon}} = (\boldsymbol{I} + \boldsymbol{\Upsilon}(\boldsymbol{I} - \boldsymbol{\Lambda}))^{-1}\boldsymbol{\Upsilon}\boldsymbol{\Lambda} \tag{4.63}$$

equation (4.62) can be written equivalently as

$$\left(\boldsymbol{I} - \tilde{\boldsymbol{\Upsilon}}\boldsymbol{H}^{-1}\boldsymbol{A}\right)\boldsymbol{p} = \sigma^2\tilde{\boldsymbol{\Upsilon}}\boldsymbol{H}^{-1}\boldsymbol{u} \tag{4.64}$$

Comparing (4.10) and (4.64) we observe that $\tilde{\boldsymbol{\Upsilon}}$ is the matrix of modified SIR targets. With algorithm (4.58), the SIR of user $i$ converges to the $i$th diagonal element of $\tilde{\boldsymbol{\Upsilon}}$, $\tilde{\gamma_i}^*$

$$\tilde{\gamma_i}^* = \frac{\alpha_i\gamma_i^*}{1 + \gamma_i^*(1 - \alpha_i)} \tag{4.65}$$

instead of the originally intended SIR target $\gamma_i^*$.

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

$$E\left[\frac{1}{\hat{h}_{ii}}\right] = \frac{1}{2\xi_i}\ln\left(\frac{1 + \xi_i}{1 - \xi_i}\right)\frac{1}{h_{ii}} \tag{4.66}$$

and therefore $\alpha_i = 1/(2\xi_i)\ln((1 + \xi_i)/(1 - \xi_i))$. Note that $\alpha_i$ monotonically increases with $\xi_i$, the percentage error. Also note from (4.65) that $\tilde{\gamma_i}^* \geq \gamma_i^*$ for all $\alpha_i \geq 1$ and $\tilde{\gamma_i}^*$ increases with $\alpha_i$. 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.

### 4.6 Simulation Results

In our simulations, we consider a general multicell CDMA system on a rectangular grid. There are $M = 25$ base stations with $(x, y)$ coordinates $(1000i + 500, 1000j + 500)$ for $0 \leq i, j \leq 4$. The $x$ and $y$ coordinates of each user are independent uniformly distributed random variables between 0 and 5,000 meters. The experiments are conducted for number of users ($N$) between 200 and 500. Figure 4.2 shows the positions of users and the base stations with symbols $\times$ and $\circ$, respectively, for $N = 500$. Each user is assigned to its nearest base station. The path loss exponent used while calculating the channel gains of the users is taken to be $\alpha = 4$. At the beginning of the iterations, power vector is always initialized to zero. The simulations are over $10,000$ bits. For $L$ bit measurement averaging, the number of power control iterations is $10,000/L$.

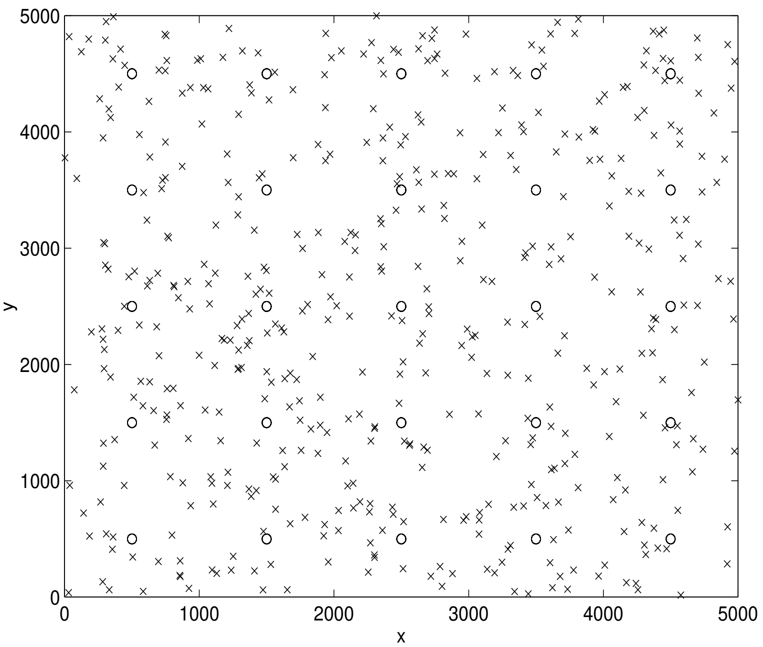


Figure 4.2: Simulation environment for $N = 500$. Symbols $\circ$ and $\times$ denote the base stations and the users, respectively.

We chose the processing gain to be $G = 150$ and a random signature sequence of length $G$ chips was assigned to each user. Although the convergence theorems permit individual SIR targets $\gamma_i^*$ for each user $i$, for the simulations we chose a common SIR target $\gamma_i^* = 4$ ($\approx 6$ dB) for all users. The AWGN noise power equaled $\sigma^2 = 10^{-13}$ W, corresponding roughly to a 1 MHz bandwidth.

First we investigate the performance of the stochastic power control algorithms for $N = 200$. The normalized squared error which we define as

$$\mathrm{NSE}(n) = \|\boldsymbol{p}(n) - \bar{\boldsymbol{p}}\|^2 \,/\, \|\bar{\boldsymbol{p}}\|^2 \tag{4.67}$$

is plotted as a function of iteration index in Figure 4.3. The curves of Figure 4.3 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 4.4 and 4.5 show the same performance criteria when averaging is implemented with $L = 10$ and $L = 100$, respectively.

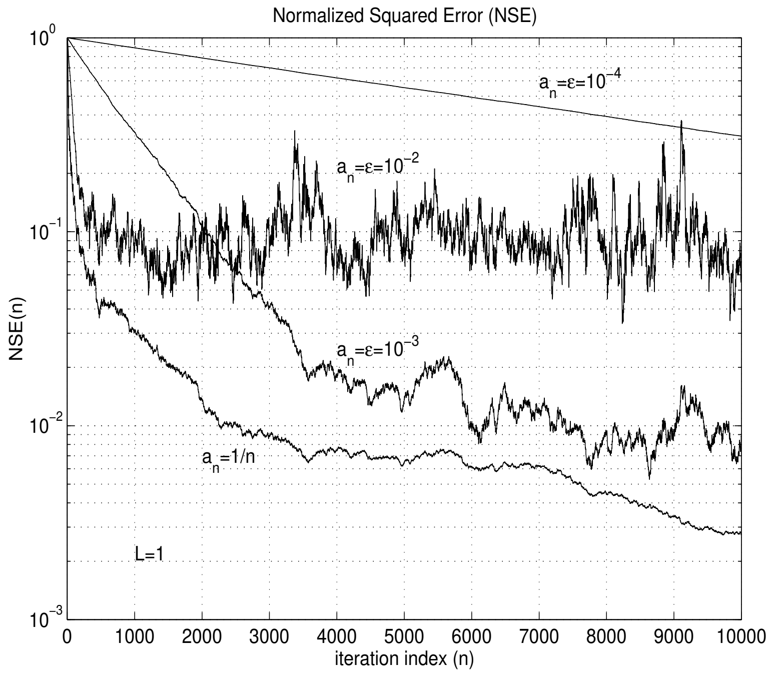


Figure 4.3: NSE as a function of $n$ for stochastic power control algorithms with $a_n = 1/n$ and $a_n = \epsilon$ for $\epsilon = 10^{-2}, 10^{-3}, 10^{-4}$. No averaging is used, $L = 1$.

We observe the trade-off between the convergence rate and the value of the limiting NSE: when $\epsilon$ is large, $\alpha_0$ and $\alpha_1$ 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 4.3. On the other hand, if $\epsilon$ is close to zero, then the limiting value of the NSE is smaller, but since $\alpha_0$ and $\alpha_1$ 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 the $\epsilon = 10^{-4}$ curve in

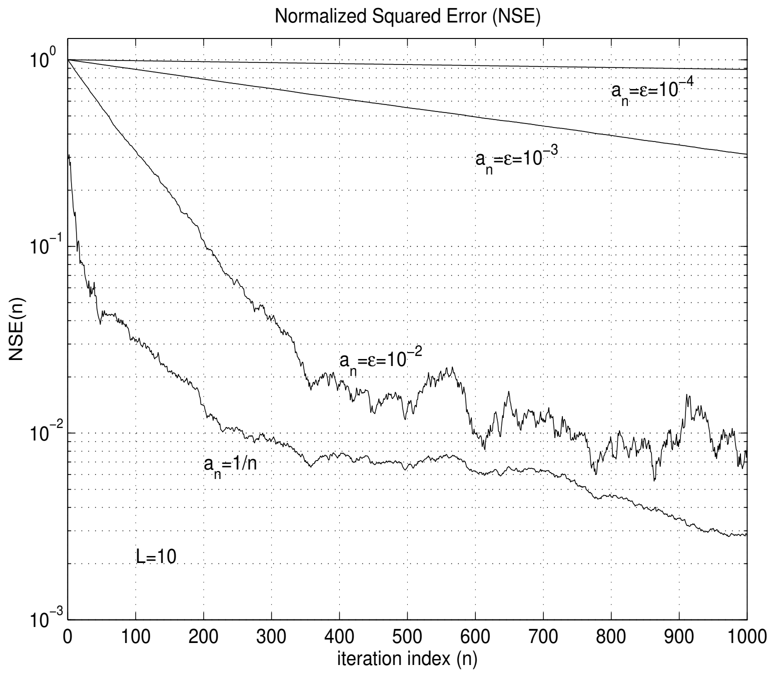


Figure 4.4: NSE as a function of $n$ for stochastic power control algorithms with $a_n = 1/n$ and $a_n = \epsilon$ for $\epsilon = 10^{-2}, 10^{-3}, 10^{-4}$. Averaging over $L = 10$ bits is implemented.

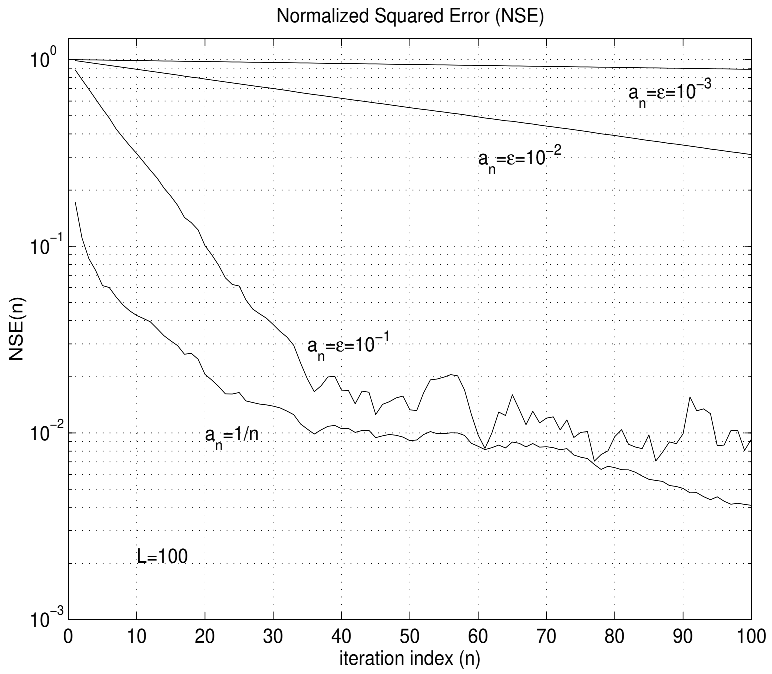


Figure 4.5: NSE as a function of $n$ for stochastic power control algorithms with $a_n = 1/n$ and $a_n = \epsilon$ for $\epsilon = 10^{-1}, 10^{-2}, 10^{-3}$. Averaging over $L = 100$ bits is implemented.

Figure 4.3. Also we observe from Figures 4.3, 4.4 and 4.5 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 stochastic power control algorithm with $a_n = 1/n$ and with $a_n = \epsilon$ for $\epsilon = 10^{-2}, 10^{-3}$ and $10^{-4}$ and plotted the average of SIRs of all users and average deviation of the SIRs of all users from the target SIR in Figures 4.6 and 4.7, respectively, as a function of the iteration index $n$. If $\gamma_i^{\mathrm{dB}}(n)$ and $\gamma_i^{*\mathrm{dB}}$ denote the the SIR of user $i$ at iteration $n$ and the target SIR of the same user in dB, the average SIR plotted in Figure 4.6 is calculated as

$$\frac{1}{N}\sum_{i=1}^{N}\gamma_i^{\mathrm{dB}}(n) \tag{4.68}$$

and the average deviation of the SIRs from the target SIR plotted in Figure 4.7 is calculated as

$$\sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(\gamma_i^{\mathrm{dB}}(n) - \gamma_i^{*\mathrm{dB}}\right)^2} \tag{4.69}$$

We observed that with $a_n = 1/n$, SIRs converge to the target SIR as expected; the average SIR goes to the target SIR (see Figure 4.6) and the deviation of SIRs from the target SIR decreases steadily as number of iterations grows (see Figure 4.7). For fixed $a_n = \epsilon$, we observed the trade-off between the convergence rate and oscillations around the convergence point. As $\epsilon$ increases, the convergence rate increases, the average SIR approaches to the target SIR faster, and the average deviation initially decreases faster, however, the SIR sequences oscillate around the target value with increasing amplitude; the average deviation curves level off and stop decreasing as $n$ increases. Two extreme cases worth comparing. For $\epsilon = 10^{-2}$, the average SIR increases towards the target SIR faster and average deviation decreases faster initially, but flattens after about 500 iterations. For $\epsilon = 10^{-4}$, the average SIR increases slowly towards the target value and the average deviation decreases slowly but steadily.

Figure 4.8 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 $\gamma^*$ and $\tilde{\gamma}^*$. From (4.66) $\alpha_i = 1.03$ for all $i$. With $\gamma^* = 4$ ($\approx$ 6 dB), the common

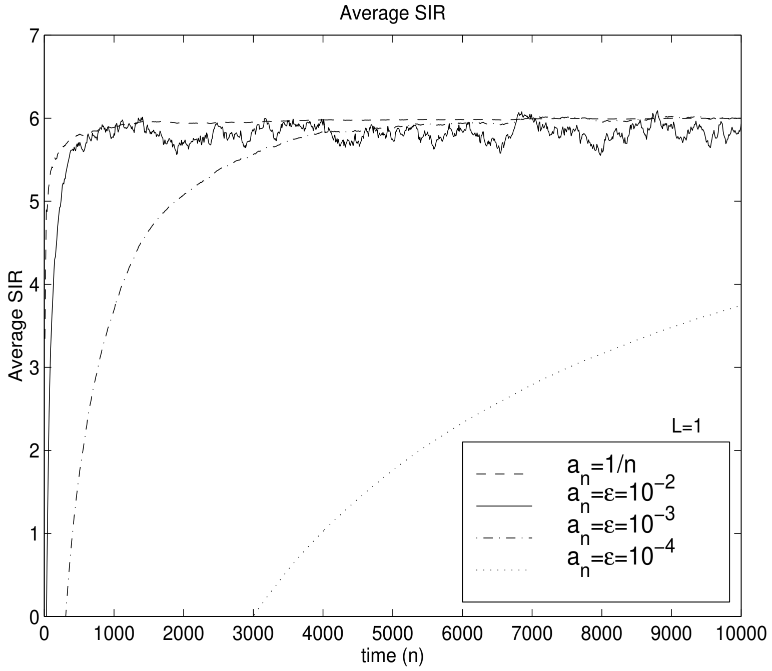


Figure 4.6: Average SIR as a function of $n$ for stochastic power control algorithms with $a_n = 1/n$ and $a_n = \epsilon$ for $\epsilon = 10^{-2}, 10^{-3}, 10^{-4}$. No averaging is used, $L = 1$.

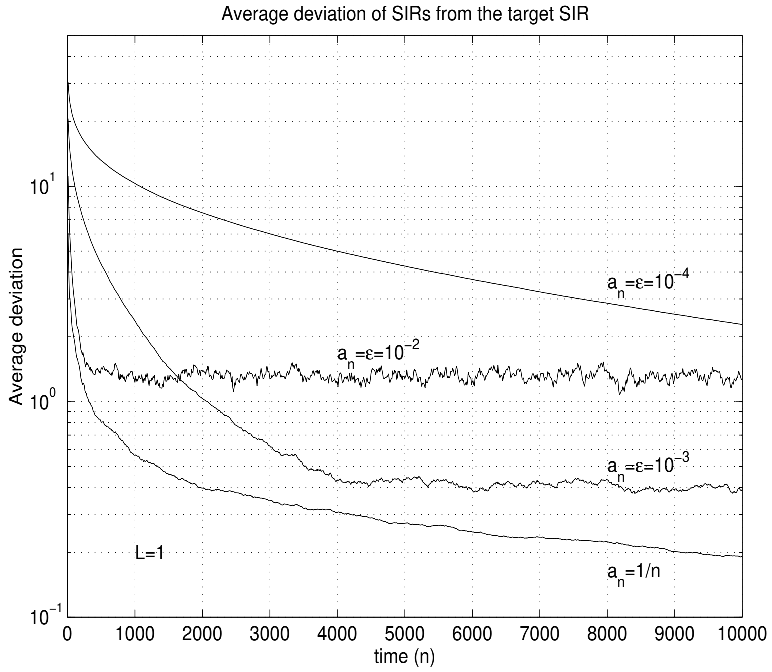


Figure 4.7: Average deviation of the SIRs from the target SIR as a function of $n$ for stochastic power control algorithms with $a_n = 1/n$ and $a_n = \epsilon$ for $\epsilon = 10^{-2}, 10^{-3}, 10^{-4}$. No averaging is used, $L = 1$.

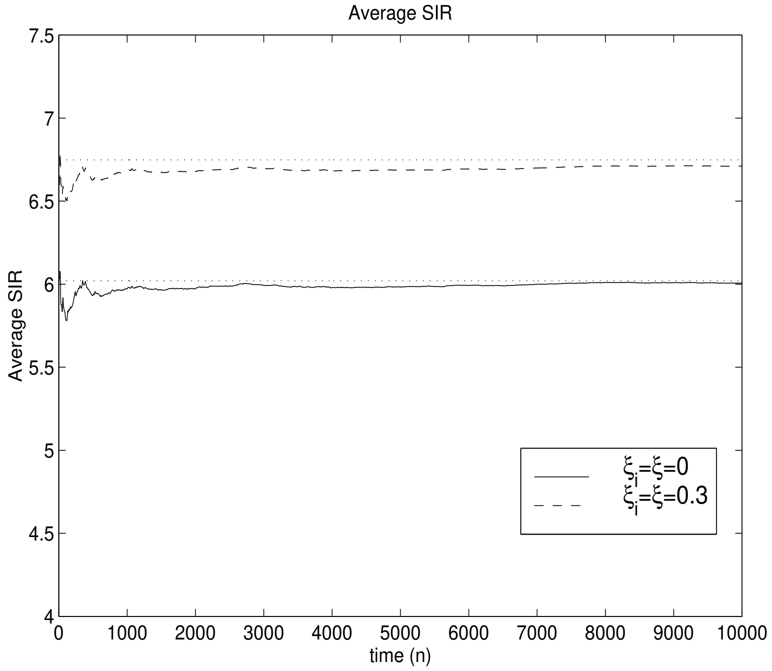


Figure 4.8: Average SIR of all users as a function of iteration index for the case with ($\xi_i = \xi = 0.3$) and without ($\xi_i = \xi = 0$) channel estimation error. Stochastic power control algorithm with $a_n = 1/n$ is used for $N = 200$ and $\gamma^* = 4$ ( $\approx 6$ dB). Converging point SIR $\tilde{\gamma}^* = 4.73$ ($\approx 6.75$ dB). Horizontal dotted lines show the levels $\gamma^*$ and $\tilde{\gamma}^*$ in dB.

convergence point SIR for all users is calculated to be $\hat{\gamma}^* = 4.73$ ($\approx 6.75$ dB) from (4.65). Other parameters of this experiment are the same as the previous one, i.e., $N = 200$, $\sigma = 10^{-13}$, $G = 150$. In Figure 4.8, the average SIR of all the users which is defined as in (4.68) is plotted for the stochastic power control algorithm with $a_n = 1/n$.

In a practical system, averaged matched filter outputs are fed back from the base stations to the mobiles and the term $\frac{1}{L}\sum_{l=1}^{L} y_i^2(n,l)$ in (4.28) needs to be quantized. The simulation results presented up to this point assumed infinite precision on this feedback (no quantization). Also in a practical system, a high value of measurement averaging ($L$) needs to be used to keep the number of power control bits per information bit small. In Figure 4.9, we present results for a practical version of the proposed algorithm. Figure 4.9 shows the average deviation of the SIRs from the SIR target for no quantization and for quantization with 2, 3 and 4 bits. It is seen that quantizing the average of matched filter outputs with 3 bits (8 quantization levels) gives quite satisfactory results and the performance of the proposed algorithm with 4 bit quantization

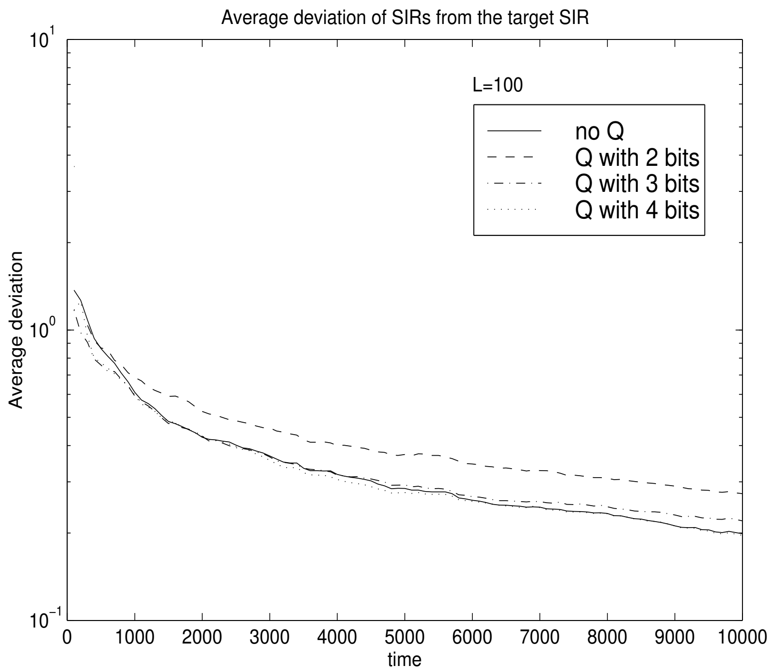


Figure 4.9: Average deviation of SIRs from the target SIR as a function of $n$ for stochastic power control algorithms with $a_n = 1/n$ for no quantization (infinite precision) and for quantization with 2, 3 and 4 bits. Averaging over $L = 100$ is implemented.

(16 quantization levels) is not distinguishable from the case where no quantization is applied (infinite precision on the feedback information). Since the averaging interval is $L = 100$ bits, with 3 (4) bits of quantization the ratio of information bits to power control bits is 100/3=33.3 (100/4=25).

In the current IS-95 CDMA system [66], 800 power control updates occur in each second. Every update uses a single bit to command the mobile unit to increase or decrease its transmitter power by a fixed amount. A 9.6 kb/s uplink connection has an effective data rate of 28.8 kb/s since the uplink data undergo a rate 1/3 convolutional encoding. Thus, the ratio of information bits to power control bits for the current IS-95 system is 28800/800=36 which is roughly the same as the ratio found above for the proposed algorithm with $L = 100$ and 3 bit quantization of the feedback information. It should be noted however that in the IS-95 system, the power control information is transmitted more frequently but with fewer bits at a time (1 bit) as opposed to the proposed algorithm where more power control bits are transmitted at a time (3, 4 bits) but less frequently. In the current IS-95 system the power control bits overwrite the

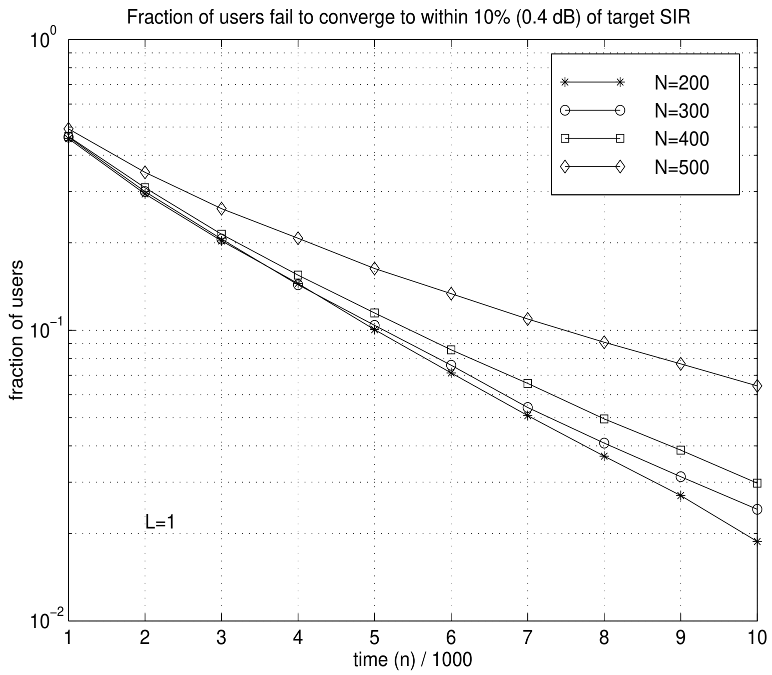


Figure 4.10: Fraction of users fail to converge to within 10% (0.4 dB) of the target SIR versus number of iterations divided by 1000, for the stochastic power control algorithm with $a_n = 1/n$. No averaging is used, $L = 1$. $N$ = number of users.

downlink data bits which are recovered by the error correction coding. If the same scheme is used in the proposed algorithm, a suitable coding scheme which can correct burst bit errors should be chosen, since a power control update will in general consist of a few bits as discussed above.

The rest of the simulations examine the stochastic power control algorithm 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 and positions of the mobiles. In Figure 4.10, we varied the number of users between 200 and 500 with increments of 100. We observed that the convergence rate decreases with increasing number of users.

We implemented the stochastic power control algorithm with $L = 1$, 10 and 100 bits of measurement averaging for $N = 200$ and $N = 400$ users. The fraction of users who fail to converge within 10% of their SIR target is plotted in Figure 4.11 for $N = 200$ and in Figure 4.12 for $N = 400$. In general, we observed that when the system is lightly loaded (case of $N = 200$), $M/L$ iterations of the stochastic power control algorithm

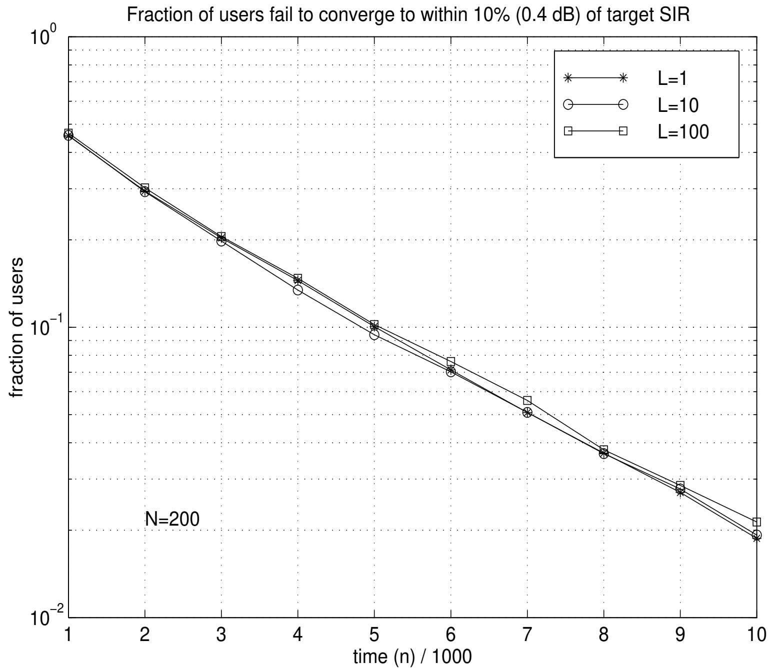


Figure 4.11: Fraction of users fail to converge to within 10% ( 0.4 dB ) of the SIR target versus time (n) / 1000, for the stochastic power control algorithm with $a_n = 1/n$. Number of users in the system is $N = 200$. Different curves correspond to $L = 1, 10, 100$ where $L$ is the number of bits over which the observations are averaged.

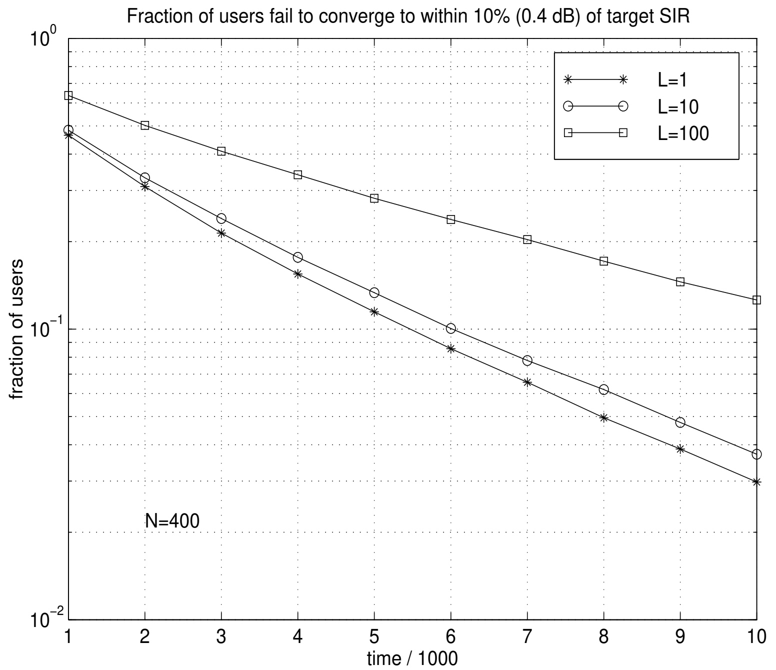


Figure 4.12: Fraction of users fail to converge to within 10% ( 0.4 dB ) of the SIR target versus time (n) / 1000, for the stochastic power control algorithm with $a_n = 1/n$. Number of users in the system is $N = 400$. Different curves correspond to $L = 1, 10, 100$ where $L$ is the number of bits over which the observations are averaged.

with averaging over $L$ bits performs as well as $M$ iterations of the stochastic power control algorithm without averaging. However, when the system is highly loaded (case of $N = 400$), averaging over a large number of bits $L$ slows the convergence.

## 4.7 Chapter Summary and Conclusions

Power control is necessary to cope with the near-far problem, to save battery life of mobile terminals, and to increase the capacity of CDMA system which is interference limited by decreasing the interference. In deterministic power control algorithms transmitter powers of the users are updated iteratively by using local variables in an autonomous (distributed) manner and the system power vector converges to the optimal one where all users satisfy their SIR based quality of service requirements while using the least power. The deterministic approach to power control assumes that either the SIR or the interference or the BER is measured perfectly. These parameters cannot be known exactly but can be estimated from the received signal at the base stations. This chapter introduces stochastic power control algorithms where the transmitter powers of the users are updated iteratively and autonomously by using readily available random observations. This chapter presents practical power control algorithms with provable convergence.

In the proposed power control algorithms each user uses the squared value of its matched filter's output and its own channel gain to its assigned base station to update its transmitter power. Since the parameters needed by the user to update its power are related to that user only the proposed algorithms are distributed.

Since the matched filter outputs are random, the proposed power control algorithms evolve stochastically and we prove the convergence of the algorithms in the MSE sense. Two classes of stochastic power control algorithms are proposed: 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 MSE for the first class of stochastic power control algorithms. We showed that we can always choose a small enough $\epsilon$ to have finite lower upper bounds on the MSE so long as the deterministic power control problem is feasible. We also showed that the algorithms converge in the MSE as $\epsilon$ goes to zero. 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.

We investigated the effects of averaging of the observations, quantization of the feedback information from the base stations and estimation errors of the channel gains on the convergence of the proposed algorithms.

## 4.A Appendix: Matrix Properties

The definition of a *stable* matrix given below can be found in [47, 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 [47, page 405] or [13, 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,*

$$\boldsymbol{S}^\top \boldsymbol{G} + \boldsymbol{G}\boldsymbol{S} = \boldsymbol{C} \tag{4.70}$$

The following simple theorem on vector norms can be found in [26, 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 [49, page 349].

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

$$\lambda_{\min}^M \|\boldsymbol{x}\|^2 \le \boldsymbol{x}^\top \boldsymbol{M}\boldsymbol{x} \le \lambda_{\max}^M \|\boldsymbol{x}\|^2 \tag{4.71}$$

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

$$\lambda_{\min}^G \|\boldsymbol{x}\|^2 \le \|\boldsymbol{x}\|_G^2 \le \lambda_{\max}^G \|\boldsymbol{x}\|^2 \qquad \text{and} \qquad \frac{1}{\lambda_{\max}^G} \|\boldsymbol{x}\|_G^2 \le \|\boldsymbol{x}\|^2 \le \frac{1}{\lambda_{\min}^G} \|\boldsymbol{x}\|_G^2 \tag{4.72}$$

We combine (4.72) and the Rayleigh quotient in the following useful results.

**Lemma 11** *For any symmetric positive definite matrix $\boldsymbol{G}$ and symmetric matrix $\boldsymbol{M}$*

$$\frac{\lambda_{\min}^{M}}{\lambda_{\max}^{G}}\,\|\boldsymbol{x}\|_G^2 \leq \boldsymbol{x}^\top \boldsymbol{M}\boldsymbol{x} \leq \frac{\lambda_{\max}^{M}}{\lambda_{\min}^{G}}\,\|\boldsymbol{x}\|_G^2 \tag{4.73}$$

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

$$\lambda_{\min}^{S^\top S}\,\|\boldsymbol{x}\|^2 \leq \|\boldsymbol{S}\boldsymbol{x}\|^2 \leq \lambda_{\max}^{S^\top S}\,\|\boldsymbol{x}\| \tag{4.74}$$

## 4.B Appendix: Additional Proofs

**Proof: Lemma 7** From (4.72), for any realization of the random variables, we have

$$\frac{1}{\lambda_{\max}^{G}}\,\|\boldsymbol{p}(n)-\bar{\boldsymbol{p}}\|_G^2 \leq \|\boldsymbol{p}(n)-\bar{\boldsymbol{p}}\|^2 \leq \frac{1}{\lambda_{\min}^{G}}\,\|\boldsymbol{p}(n)-\bar{\boldsymbol{p}}\|_G^2 \tag{4.75}$$

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

$$\frac{1}{\lambda_{\max}^{G}}\lim_{n\to\infty} E\Big[\|\boldsymbol{p}(n)-\bar{\boldsymbol{p}}\|_G^2\Big] \leq \lim_{n\to\infty} E\Big[\|\boldsymbol{p}(n)-\bar{\boldsymbol{p}}\|^2\Big] \leq \frac{1}{\lambda_{\min}^{G}}\lim_{n\to\infty} E\Big[\|\boldsymbol{p}(n)-\bar{\boldsymbol{p}}\|_G^2\Big] \tag{4.76}$$

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. □

**Proof: Lemma 8** Using $\lambda_i$ and $\mu_i$ 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 6 imply $|\mu_i| < 1$. Hence, $|Re\{\mu_i\}| < 1$ and the result follows. □

**Proof: Lemma 10** 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 10. Since $\boldsymbol{G}$ is symmetric positive definite, $\boldsymbol{\eta}(n)\boldsymbol{G}\boldsymbol{\eta}(n) \leq \lambda_{\max}^{G}\boldsymbol{\eta}(n)^\top\boldsymbol{\eta}(n)$ for all realizations of the random vector $\boldsymbol{\eta}(n)$. Therefore, we have

$$E\Big[\|\boldsymbol{\eta}(n)\|_G^2 \mid \boldsymbol{p}(n) = \boldsymbol{p}\Big] \leq \lambda_{\max}^{G} E\Big[\|\boldsymbol{\eta}(n)\|^2 \mid \boldsymbol{p}(n) = \boldsymbol{p}\Big] \tag{4.77}$$

From (4.33), $\|\boldsymbol{\eta}(n)\|^2 = \Big\|\boldsymbol{\Upsilon}\boldsymbol{H}^{-1}\boldsymbol{w}(n)\Big\|^2$ so that applying Corollary 1 and taking the conditional expectation yields

$$E\Big[\|\boldsymbol{\eta}(n)\|^2 \mid \boldsymbol{p}(n) = \boldsymbol{p}\Big] \leq \frac{(\gamma_{\max}^*)^2}{h_{\min}^2} E\Big[\|\boldsymbol{w}(n)\|^2 \mid \boldsymbol{p}(n) = \boldsymbol{p}\Big] \tag{4.78}$$

In deriving (4.78), 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 (4.77) and (4.78) yields,

$$E\Big[\|\boldsymbol{\eta}(n)\|^2_G \mid \boldsymbol{p}(n) = \boldsymbol{p}\Big] \leq \frac{(\gamma^*_{\max})^2 \lambda^G_{\max}}{h^2_{\min}} \sum_{i=1}^{N} E\Big[w_i^2(n) \mid \boldsymbol{p}(n) = \boldsymbol{p}\Big] \tag{4.79}$$

Using (4.18), $E[w_i^2(n)]$ can be expressed as

$$\begin{aligned} E\Big[w_i^2(n) \mid \boldsymbol{p}(n) = \boldsymbol{p}\Big] &= \frac{1}{L^2}\Bigg(\sum_{l=1}^{L} E\Big[w_i^2(n,l) \mid \boldsymbol{p}(n) = \boldsymbol{p}\Big] \\ &\quad + \sum_{l=1}^{L}\sum_{k\neq l} E[w_i(n,l)w_i(n,k) \mid \boldsymbol{p}(n) = \boldsymbol{p}]\Bigg) \end{aligned} \tag{4.80}$$

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 $k \neq l$. This requires the computation of cross correlations between $\hat{b}_{ij}(l)$ terms. Note from the definition of $\hat{b}_{ij}(l)$ given in (4.4) that

$$E\Big[\hat{b}_{ij}(l)\hat{b}_{ik}(m)\Big] = \begin{cases} 0 & j \neq k \\ 0 & j = k, |l-m| > 1 \\ \Gamma_{ij}(\bar{\Gamma}_{ij} + \tilde{\Gamma}_{ij}) & j = k, |l-m| = 1 \\ \bar{\Gamma}^2_{ij} + \Gamma^2_{ij} + \tilde{\Gamma}^2_{ij} & j = k, l = m \end{cases} \tag{4.81}$$

Note also that

$$E\Big[\hat{b}^4_{ij}(l)\Big] = \left(\bar{\Gamma}^2_{ij} + \Gamma^2_{ij} + \tilde{\Gamma}^2_{ij}\right)^2 + 4\Gamma^2_{ij}\left(\bar{\Gamma}^2_{ij} + \tilde{\Gamma}^2_{ij}\right) \tag{4.82}$$

and for $k \neq l$

$$E\Big[\hat{b}^2_{ij}(l)\hat{b}^2_{ij}(k)\Big] = \left(\bar{\Gamma}^2_{ij} + \Gamma^2_{ij} + \tilde{\Gamma}^2_{ij}\right)^2 \tag{4.83}$$

First we will evaluate $E[w_i^2(n,l) \mid \boldsymbol{p}(n) = \boldsymbol{p}]$. From (4.15),

$$E\Big[w_i^2(n,l) \mid \boldsymbol{p}(n) = \boldsymbol{p}\Big] = E\Big[v_i^2(n,l) \mid \boldsymbol{p}(n) = \boldsymbol{p}\Big] - \left(\sum_{j=1}^{N} p_j h_{ij}\left(\bar{\Gamma}^2_{ij} + \Gamma^2_{ij} + \tilde{\Gamma}^2_{ij}\right) + \sigma^2\right)^2 \tag{4.84}$$

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

$$v_i^2(n,l) = \underbrace{\sum_{j=1}^{N}\sum_{k=1}^{N}\sum_{s=1}^{N}\sum_{t=1}^{N} \sqrt{p_j}\sqrt{p_k}\sqrt{p_s}\sqrt{p_t}\sqrt{h_{ij}}\sqrt{h_{ik}}\sqrt{h_{is}}\sqrt{h_{it}}\hat{b}_{ij}(l)\hat{b}_{ik}(l)\hat{b}_{is}(l)\hat{b}_{it}(l)}_{T_1}$$

$$+4\underbrace{n_i^2(l)\sum_{j=1}^{N}\sum_{k=1}^{N}\sqrt{p_j}\sqrt{p_k}\sqrt{h_{ij}}\sqrt{h_{ik}}\hat{b}_{ij}(l)\hat{b}_{ik}(l)}_{T_2}+\underbrace{n_i^4(l)}_{T_3}$$

$$+4\underbrace{n_i(l)\sum_{j=1}^{N}\sum_{k=1}^{N}\sum_{s=1}^{N}\sqrt{p_j}\sqrt{p_k}\sqrt{p_s}\sqrt{h_{ij}}\sqrt{h_{ik}}\sqrt{h_{is}}\hat{b}_{ij}(l)\hat{b}_{ik}(l)\hat{b}_{is}(l)}_{T_4} \tag{4.85}$$

$$+2\underbrace{n_i^2(l)\sum_{j=1}^{N}\sum_{k=1}^{N}\sqrt{p_j}\sqrt{p_k}\sqrt{h_{ij}}\sqrt{h_{ik}}\hat{b}_{ij}(l)\hat{b}_{ik}(l)}_{T_5}+2\underbrace{n_i^3(l)\sum_{j=1}^{N}\sqrt{p_j}\sqrt{h_{ij}}\hat{b}_{ij}(l)}_{T_6}$$

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 (4.81) and (4.82),

$$\begin{aligned} E[T_1 \mid \boldsymbol{p}(n)=\boldsymbol{p}] &= 3\sum_{j=1}^{N}\sum_{k\neq j} p_j p_k h_{ij} h_{ik}(\bar{\Gamma}_{ij}^2+\Gamma_{ij}^2+\tilde{\Gamma}_{ij}^2)(\bar{\Gamma}_{ik}^2+\Gamma_{ik}^2+\tilde{\Gamma}_{ik}^2) \\ &\quad +\sum_{j=1}^{N} p_j^2 h_{ij}^2\left((\bar{\Gamma}_{ij}^2+\Gamma_{ij}^2+\tilde{\Gamma}_{ij}^2)^2+4\Gamma_{ij}^2(\bar{\Gamma}_{ij}^2+\tilde{\Gamma}_{ij}^2)\right) \end{aligned} \tag{4.86}$$

From (4.81) and the fact that $E[n_i^2(l)]=\sigma^2$,

$$E[T_2 \mid \boldsymbol{p}(n)=\boldsymbol{p}] = E[T_5 \mid \boldsymbol{p}(n)=\boldsymbol{p}] = \sigma^2\sum_{j=1}^{N} p_j h_{ij}(\bar{\Gamma}_{ij}^2+\Gamma_{ij}^2+\tilde{\Gamma}_{ij}^2) \tag{4.87}$$

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

$$E[T_4 \mid \boldsymbol{p}(n)=\boldsymbol{p}] = E[T_6 \mid \boldsymbol{p}(n)=\boldsymbol{p}] = 0 \tag{4.88}$$

And finally

$$E[T_3 \mid \boldsymbol{p}(n)=\boldsymbol{p}] = 3\sigma^4 \tag{4.89}$$

We obtain $E[v_i^2(n,l) \mid \boldsymbol{p}(n)=\boldsymbol{p}]$ by combining the results of (4.86), (4.87), (4.88) and (4.89) and insert the result in (4.84) to get

$$\begin{aligned} E\Big[w_i^2(n,l) \mid \boldsymbol{p}(n)=\boldsymbol{p}\Big] &= 2\sum_{j=1}^{N}\sum_{k\neq j} p_j p_k h_{ij} h_{ik}(\bar{\Gamma}_{ij}^2+\Gamma_{ij}^2+\tilde{\Gamma}_{ij}^2)(\bar{\Gamma}_{ik}^2+\Gamma_{ik}^2+\tilde{\Gamma}_{ik}^2) \\ &\quad +4\sum_{j=1}^{N} p_j^2 h_{ij}^2 \Gamma_{ij}^2(\bar{\Gamma}_{ij}^2+\tilde{\Gamma}_{ij}^2) \\ &\quad +4\sigma^2\sum_{j=1}^{N} p_j h_{ij}(\bar{\Gamma}_{ij}^2+\Gamma_{ij}^2+\tilde{\Gamma}_{ij}^2)+2\sigma^4 \end{aligned} \tag{4.90}$$

By similar manipulations, it can be shown that

$$E[w_i(n,l)w_i(n,l-1) \mid \boldsymbol{p}(n)=\boldsymbol{p}] = 2\sum_{j=1}^{N}\sum_{k\neq j} p_j p_k h_{ij} h_{ik} \Gamma_{ij}\Gamma_{ik}(\bar{\Gamma}_{ij}+\tilde{\Gamma}_{ij})(\bar{\Gamma}_{ik}+\tilde{\Gamma}_{ik}) \tag{4.91}$$

and for $|l-k|>1$

$$E[w_i(n,l)w_i(n,k) \mid \boldsymbol{p}(n)=\boldsymbol{p}] = 0 \tag{4.92}$$

By using the result (4.92) we observe that (4.80) simplifies to

$$\begin{aligned} E\Big[w_i^2(n) \mid \boldsymbol{p}(n)=\boldsymbol{p}\Big] \;&=\; \frac{1}{L^2}\Big\{ LE\Big[w_i^2(n,l) \mid \boldsymbol{p}(n)=\boldsymbol{p}\Big] \\ &\quad +2(L-1)E[w_i(n,l)w_i(n,l-1) \mid \boldsymbol{p}(n)=\boldsymbol{p}]\Big\} \end{aligned} \tag{4.93}$$

Inserting (4.90) and (4.91) into (4.93) yields

$$\begin{aligned} E\Big[w_i^2(n) \mid \boldsymbol{p}(n)=\boldsymbol{p}\Big] \;&=\; \frac{1}{L^2}\Big\{ 2L\sum_{j=1}^{N}\sum_{k\neq j} p_j p_k h_{ij} h_{ik}(\bar{\Gamma}_{ij}^2+\Gamma_{ij}^2+\tilde{\Gamma}_{ij}^2)(\bar{\Gamma}_{ik}^2+\Gamma_{ik}^2+\tilde{\Gamma}_{ik}^2) \\ &\quad +4L\sum_{j=1}^{N} p_j^2 h_{ij}^2 \Gamma_{ij}^2(\bar{\Gamma}_{ij}^2+\tilde{\Gamma}_{ij}^2) \\ &\quad +4L\sigma^2\sum_{j=1}^{N} p_j h_{ij}(\bar{\Gamma}_{ij}^2+\Gamma_{ij}^2+\tilde{\Gamma}_{ij}^2) + 2L\sigma^4 \\ &\quad +4(L-1)\sum_{j=1}^{N}\sum_{k\neq j} p_j p_k h_{ij} h_{ik}\Gamma_{ij}\Gamma_{ik}(\bar{\Gamma}_{ij}+\tilde{\Gamma}_{ij})(\bar{\Gamma}_{ik}+\tilde{\Gamma}_{ik})\Big\} \end{aligned} \tag{4.94}$$

Combining the second and the last terms in (4.94) yields

$$\begin{aligned} E\Big[w_i^2(n) \mid \boldsymbol{p}(n)=\boldsymbol{p}\Big] \;&=\; \frac{1}{L^2}\Big\{ 2L\sum_{j=1}^{N}\sum_{k\neq j} p_j p_k h_{ij} h_{ik}(\bar{\Gamma}_{ij}^2+\Gamma_{ij}^2+\tilde{\Gamma}_{ij}^2)(\bar{\Gamma}_{ik}^2+\Gamma_{ik}^2+\tilde{\Gamma}_{ik}^2) \\ &\quad +4\sum_{j=1}^{N} p_j^2 h_{ij}^2 \Gamma_{ij}^2(\bar{\Gamma}_{ij}^2+\tilde{\Gamma}_{ij}^2) + 4L\sigma^2\sum_{j=1}^{N} p_j h_{ij}(\bar{\Gamma}_{ij}^2+\Gamma_{ij}^2+\tilde{\Gamma}_{ij}^2) \\ &\quad +2L\sigma^4 + 4(L-1)\Big(\sum_{j=1}^{N} p_j h_{ij}\Gamma_{ij}(\bar{\Gamma}_{ij}+\tilde{\Gamma}_{ij})\Big)^2\Big\} \end{aligned} \tag{4.95}$$

Now we will derive an upper bound for (4.95). First we will include the terms corresponding to $j=k$ in the first double summation on the right hand side of (4.95). Since the included terms are nonnegative, this process yields an upper bound, and the double sum becomes equal to $\Big(\sum_j p_j h_{ij}(\bar{\Gamma}_{ij}^2+\Gamma_{ij}^2+\tilde{\Gamma}_{ij}^2)\Big)^2$. In order to derive an upper bound for the last term in (4.95), 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 $ab \le (a^2+b^2)/2$ 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 (4.95) 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

$$\begin{aligned} E\Big[w_i^2(n) \mid \boldsymbol{p}(n)=\boldsymbol{p}\Big] \;&\le\; \frac{3}{L}\left(\sum_{j=1}^{N} p_j h_{ij}(\bar{\Gamma}_{ij}^2+\Gamma_{ij}^2+\tilde{\Gamma}_{ij}^2)+\sigma^2\right)^2 \\ &\quad +\frac{4}{L^2}\sum_{j=1}^{N} p_j^2 h_{ij}^2 \Gamma_{ij}^2(\bar{\Gamma}_{ij}^2+\tilde{\Gamma}_{ij}^2) \end{aligned} \tag{4.96}$$

Note that the expression in the parenthesis on the right hand side of (4.96) is equal to the $i$th element of vector $((\boldsymbol{A}+\boldsymbol{H})\boldsymbol{p}+\sigma^2\boldsymbol{u})$. Inserting (4.96) into (4.79) yields

$$\begin{aligned} E\Big[\|\boldsymbol{\eta}(n)\|_G^2 \mid \boldsymbol{p}(n)=\boldsymbol{p}\Big] \;&\le\; \frac{(\gamma_{\max}^*)^2\lambda_{\max}^G}{h_{\min}^2 L}\Bigg\{3\left\|(\boldsymbol{A}+\boldsymbol{H})\boldsymbol{p}+\sigma^2\boldsymbol{u}\right\|^2 \\ &\quad +\frac{4}{L}\sum_{i=1}^{N}\sum_{j=1}^{N} p_j^2 h_{ij}^2\Gamma_{ij}^2(\bar{\Gamma}_{ij}^2+\tilde{\Gamma}_{ij}^2)\Bigg\} \end{aligned} \tag{4.97}$$

Let $\boldsymbol{F}$ be a diagonal matrix with its $j$th 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 (4.97) 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 4.A, we can conclude $\boldsymbol{p}^\top\boldsymbol{F}\boldsymbol{p}\le\lambda_{\max}^F\|\boldsymbol{p}\|^2$, where $\lambda_{\max}^F$ is the largest eigenvalue of $\boldsymbol{F}$. Then we can further bound (4.97) as

$$E\Big[\|\boldsymbol{\eta}(n)\|_G^2 \mid \boldsymbol{p}(n)=\boldsymbol{p}\Big] \le \theta\left\{3\left\|(\boldsymbol{A}+\boldsymbol{H})\boldsymbol{p}+\sigma^2\boldsymbol{u}\right\|^2+\frac{4}{L}\lambda_{\max}^F\|\boldsymbol{p}\|^2\right\} \tag{4.98}$$

where $\theta=(\gamma_{\max}^*)^2\lambda_{\max}^G/(h_{\min}^2L)$. Note that for any two vectors $\boldsymbol{a}$ and $\boldsymbol{b}$, $\|\boldsymbol{a}+\boldsymbol{b}\|^2\le 2(\|\boldsymbol{a}\|^2+\|\boldsymbol{b}\|^2)$. Applying this to the first term in (4.98) and noting that $\|\boldsymbol{u}\|^2=N$ yields

$$E\Big[\|\boldsymbol{\eta}(n)\|_G^2 \mid \boldsymbol{p}(n)=\boldsymbol{p}\Big] \le \theta\left\{6\left(\|(\boldsymbol{A}+\boldsymbol{H})\boldsymbol{p}\|^2+\sigma^2N\right)+\frac{4}{L}\lambda_{\max}^F\|\boldsymbol{p}\|^2\right\} \tag{4.99}$$

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

$$\begin{aligned} E\Big[\|\boldsymbol{\eta}(n)\|_G^2 \mid \boldsymbol{p}(n)=\boldsymbol{p}\Big] \;&\le\; 2\theta\left\{3(\mu\|\boldsymbol{p}\|^2+\sigma^2N)+\frac{2}{L}\lambda_{\max}^F\|\boldsymbol{p}\|^2\right\} \\ &=\; 2\theta\left\{\mu_1\|\boldsymbol{p}\|^2+3\sigma^2N\right\} \end{aligned} \tag{4.100}$$

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 (4.72) to (4.100) yields the desired result,

$$E\Big[\|\boldsymbol{\eta}(n)\|_G^2 \mid \boldsymbol{p}(n) = \boldsymbol{p}\Big] \leq c_0 + c_1 \|\boldsymbol{p} - \bar{\boldsymbol{p}}\|_G^2 \tag{4.101}$$

Now we prove the lower bound in Lemma 10. From Equations (4.33) and (4.72), we obtain,

$$\|\boldsymbol{\eta}(n)\|_G^2 \geq \lambda_{\min}^G \left\|\boldsymbol{\Upsilon}\boldsymbol{H}^{-1}\boldsymbol{w}(n)\right\|^2 \tag{4.102}$$

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

$$\|\boldsymbol{\eta}(n)\|_G^2 \geq \frac{\lambda_{\min}^G (\gamma_{\min}^*)^2}{h_{\max}^2} \|\boldsymbol{w}(n)\|^2 \tag{4.103}$$

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

$$E\Big[\|\boldsymbol{\eta}(n)\|_G^2 \mid \boldsymbol{p}(n) = \boldsymbol{p}\Big] \geq \frac{\lambda_{\min}^G (\gamma_{\min}^*)^2}{h_{\max}^2} \sum_{i=1}^{N} E\Big[w_i^2(n) \mid \boldsymbol{p}(n) = \boldsymbol{p}\Big] \tag{4.104}$$

We note from (4.95) that

$$E\Big[w_i^2(n) \mid \boldsymbol{p}(n) = \boldsymbol{p}\Big] \geq \frac{2\sigma^4}{L} \tag{4.105}$$

since all of the remaining terms are non-negative. Applying this result to (4.104) and expressing our conditioning yields the desired result

$$E\Big[\|\boldsymbol{\eta}(n)\|_G^2 \mid \boldsymbol{x}(n) = \boldsymbol{x}\Big] \geq \frac{2\lambda_{\min}^G (\gamma_{\min}^*)^2 N \sigma^4}{h_{\max}^2 L} = d_0 \tag{4.106}$$

□

# Chapter 5

# Combined Power Control and Interference Suppression

Power control and multiuser detection methods have been proposed and studied independently as means of overcoming the near-far problem of the CDMA scheme. The rationale of this chapter is that even when optimal power control is applied, the near-far problem can be further decreased by employing interference suppression through multiuser detection, and similarly even when the interference is suppressed optimally at a fixed power level (in the sense of maximizing the SIR via MMSE filtering for instance) the powers can be further optimized to decrease the MMSE. Thus, in this chapter we combine the power control and multiuser detection approaches to overcome the near-far effect. Associated with each user there are two parameters: the transmitter power of the user and the receiver filter of the user at its assigned base station. In this chapter we propose an iterative and distributed algorithm which controls both the transmitter powers and the receiver filters of the users.

At each iteration, the receiver filter coefficients of the users are updated to suppress the interference optimally and then the transmitter powers of the users are assigned so that each user creates the minimum possible interference to others while satisfying its quality of service requirement. For a fixed power vector, the filter coefficients that maximize the SIR are the MMSE multiuser detectors [30, 71]. Thus, at each iteration, the proposed algorithm first updates the receiver filters to be the MMSE filters for all users, then updates the power levels so that each user has just enough transmitter power to satisfy its SIR based quality of service requirement.

The implementation of this approach will require interference measurements at each receiver. We show that the resulting power control algorithm converges to a fixed point power vector where all the users satisfy their SIR based quality of service requirements and that the linear receiver converges to the MMSE multiuser detector. The fixed point power vector $\bar{\boldsymbol{p}}$ satisfies $\bar{\boldsymbol{p}} \leq \boldsymbol{p}'$ for any power vector $\boldsymbol{p}'$ for which there are filter coefficients that yield acceptable SIR for all users. In [24] a power control algorithm is proposed for a CDMA system with adaptive MMSE receivers. The algorithm in [24] and the algorithm in this chapter will converge to the same minimum power solution; however, the algorithm of [24] uses measurements of the minimum mean squared error which requires the knowledge of the information bits transmitted by the users and assumes that an adaptive MMSE receiver will adjust to changes in the transmitter powers.

## 5.1 System Model

We consider the uplink of a wireless cellular system with a fixed base station assignment of $N$ users to $M$ base stations. We assume a synchronous CDMA scheme and BPSK modulation in order to simplify the analysis of our algorithm. 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 $h_{ij}$.

The signature waveform of user $i$, denoted by $s_i(t)$, 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 baseband received signal, $r(t)$, in one bit interval at the front end of the receiver filters at the assigned base of user $i$ is given by,

$$r_i(t) = \sum_{j=1}^{N} \sqrt{p_j}\sqrt{h_{ij}} b_j s_j(t) + n(t) \tag{5.1}$$

where $b_j$ is the information bit transmitted by user $j$ ($+1$ or $-1$ with equal probability) and $n(t)$ is an additive white Gaussian noise (AWGN) process.

We define the chip waveform to be $\psi(t)$, $t \in [0, T_c]$ and 0 elsewhere, where $T_c$ is the chip duration. Thus $\{\psi(t - iT_c), i = 0, \ldots, G-1\}$, where $G = T_b/T_c$ is the processing gain, is a basis for the signal space. This allows us to represent both the signature sequences and the linear receiver filters of the users with $G$ dimensional vectors. We will use $\boldsymbol{s}_i$ and $\boldsymbol{c}_i$ to denote the pre-assigned unique signature sequence and the linear receiver filter of user $i$, respectively. In terms of signal vectors, the received signal at the assigned base station of user $i$ can be written as

$$\boldsymbol{r}_i = \sum_{j=1}^{N} \sqrt{p_j}\sqrt{h_{ij}}b_j\boldsymbol{s}_j + \boldsymbol{n} \tag{5.2}$$

where $\boldsymbol{n}$ is a Gaussian random vector with $E\Big[\boldsymbol{n}\boldsymbol{n}^\top\Big] = \sigma^2\boldsymbol{I}$.

## 5.2 Problem Definition

Let $\boldsymbol{c}_i$ denote the receiver filter for user $i$ at its assigned base station. The receiver filter output of user $i$ is

$$y_i = \sum_{j=1}^{N} \sqrt{p_j}\sqrt{h_{ij}}(\boldsymbol{c}_i^\top\boldsymbol{s}_j)b_j + \tilde{n}_i \tag{5.3}$$

where $\tilde{n}_i = \boldsymbol{c}_i^\top\boldsymbol{n}$ is a Gaussian random variable with zero mean and variance $\sigma^2\boldsymbol{c}_i^\top\boldsymbol{c}_i$. The signal to interference ratio (SIR) of user $i$ can be written as

$$\mathrm{SIR}_i = \frac{p_i h_{ii}(\boldsymbol{c}_i^\top\boldsymbol{s}_i)^2}{\sum_{j\neq i} p_j h_{ij}(\boldsymbol{c}_i^\top\boldsymbol{s}_j)^2 + \sigma^2(\boldsymbol{c}_i^\top\boldsymbol{c}_i)} \tag{5.4}$$

Our aim is to find optimal powers, $p_i$, and filter coefficients, $\boldsymbol{c}_i$ for $i = 1, \ldots, N$, such that the total transmitter power is minimized while each user $i$ satisfies a quality of service requirement $\mathrm{SIR}_i \geq \gamma_i^*$, where $\gamma_i^*$, called the *target* SIR, is the minimum acceptable level of SIR for user $i$. Therefore, we can state the problem mathematically as

$$\begin{array}{lll} \min & \sum_{i=1}^{N} p_i & \\ \text{s.t.} & p_i \geq \dfrac{\gamma_i^*}{h_{ii}} \dfrac{\sum_{j\neq i} p_j h_{ij}(\boldsymbol{c}_i^\top\boldsymbol{s}_j)^2 + \sigma^2\boldsymbol{c}_i^\top\boldsymbol{c}_i}{(\boldsymbol{c}_i^\top\boldsymbol{s}_i)^2} & i = 1, \ldots, N \\ & p_i \geq 0 & i = 1, \ldots, N \\ & \boldsymbol{c}_i \in \mathbf{R}^G & i = 1, \ldots, N \end{array} \tag{5.5}$$

The above problem statement is equivalent to the following one, where an inner optimization is inserted in the constraint set.

$$\begin{array}{lll} \min\limits_{\boldsymbol{p}} & \sum\limits_{i=1}^{N} p_i & \\ \text{s.t.} & p_i \geq \dfrac{\gamma_i^*}{h_{ii}} \min\limits_{\boldsymbol{c}_i\in\mathbf{R}^G} \dfrac{\sum_{j\neq i} p_j h_{ij}(\boldsymbol{c}_i^\top\boldsymbol{s}_j)^2 + \sigma^2(\boldsymbol{c}_i^\top\boldsymbol{c}_i)}{(\boldsymbol{c}_i^\top\boldsymbol{s}_i)^2}, & \forall i \\ & p_i \geq 0, & \forall i \end{array} \tag{5.6}$$

In (5.6) the outer optimization is defined over the power vector only, whereas the inner optimization problem assumes a fixed power vector and is defined over the filter coefficients of the individual users. Before describing the power control algorithm, we solve the inner optimization problem for the filter coefficients corresponding to a fixed power vector in the next section.

## 5.3 Derivation for the Filter Coefficients

We now derive the filter coefficients when the power vector is fixed and equal to $\boldsymbol{p}$. The inner optimization problem given in (5.6) can be written equivalently as,

$$\min_{\boldsymbol{c}_i} \frac{\boldsymbol{c}_i^\top \left(\sum_{j\neq i} p_j h_{ij} \boldsymbol{s}_j \boldsymbol{s}_j^\top + \sigma^2 \boldsymbol{I}\right) \boldsymbol{c}_i}{(\boldsymbol{c}_i^\top \boldsymbol{s}_i)^2} \tag{5.7}$$

We define the $G \times G$ matrix $\boldsymbol{A}_i$ which is a function of the powers of all the users, except the power of user $i$, as

$$\boldsymbol{A}_i = \sum_{j\neq i} p_j h_{ij} \boldsymbol{s}_j \boldsymbol{s}_j^\top + \sigma^2 \boldsymbol{I} \tag{5.8}$$

and another $G \times G$ matrix $\boldsymbol{B}_i$ as $\boldsymbol{B}_i = \boldsymbol{s}_i \boldsymbol{s}_i^\top$. This permits us to write Equation (5.7) as

$$\min_{\boldsymbol{c}_i} \frac{\boldsymbol{c}_i^\top \boldsymbol{A}_i \boldsymbol{c}_i}{\boldsymbol{c}_i^\top \boldsymbol{B}_i \boldsymbol{c}_i} \tag{5.9}$$

Since $\boldsymbol{A}_i$ is strictly positive definite due to the term $\sigma^2 \boldsymbol{I}$, it can be written as $\boldsymbol{A}_i = \boldsymbol{R}_i^\top \boldsymbol{R}_i$ for some non-singular matrix $\boldsymbol{R}_i$. We define the one-to-one (since $\boldsymbol{R}_i$ is invertible) transformation $\boldsymbol{x}_i = \boldsymbol{R}_i \boldsymbol{c}_i$ and write (5.9) in terms of $\boldsymbol{x}_i$ as

$$\min_{\boldsymbol{x}_i} \frac{\boldsymbol{x}_i^\top \boldsymbol{x}_i}{\boldsymbol{x}_i^\top \boldsymbol{R}_i^{-\top} \boldsymbol{B}_i \boldsymbol{R}_i^{-1} \boldsymbol{x}_i} \tag{5.10}$$

Defining a $G$ dimensional vector $\boldsymbol{u}_i$ as $\boldsymbol{u}_i = \boldsymbol{R}_i^{-\top} \boldsymbol{s}_i$, (5.10) can equivalently be written as

$$\min_{\boldsymbol{x}_i} \frac{\boldsymbol{x}_i^\top \boldsymbol{x}_i}{\boldsymbol{x}_i^\top \boldsymbol{u}_i \boldsymbol{u}_i^\top \boldsymbol{x}_i} \tag{5.11}$$

The eigenvector of matrix $\boldsymbol{u}_i \boldsymbol{u}_i^\top$ with the maximum eigenvalue attains the minimum of the objective function in (5.11) [49]. Note that the rank of $\boldsymbol{u}_i \boldsymbol{u}_i^\top$ is equal to 1. Therefore, $(G-1)$ eigenvalues of it are equal to zero and the remaining one is equal to $\boldsymbol{u}_i^\top \boldsymbol{u}_i$ with the corresponding eigenvector $\boldsymbol{u}_i$. Thus, the solution of (5.11) is obtained to be $\bar{\boldsymbol{x}}_i = \boldsymbol{u}_i$. Applying the inverse transformation $\bar{\boldsymbol{c}}_i = \boldsymbol{R}_i^{-1} \bar{\boldsymbol{x}}_i$ yields the solution of

(5.9)

$$\bar{\boldsymbol{c}}_i = \boldsymbol{A}_i^{-1}\boldsymbol{s}_i \tag{5.12}$$

and the minimum of the objective function in (5.9) is equal to $(\boldsymbol{s}_i^\top \boldsymbol{A}^{-1}\boldsymbol{s}_i)^{-1}$. This result is not so surprising since it is well known that for a fixed power vector $\boldsymbol{p}$, the MMSE filter coefficients maximize the SIR [30], and the optimization problem of (5.9) can be written using (5.4) as

$$\min_{\boldsymbol{c}_i} p_i h_{ii} \frac{1}{\mathrm{SIR}_i} \tag{5.13}$$

Equation (5.13) is equivalent to

$$\max_{\boldsymbol{c}_i} \mathrm{SIR}_i \tag{5.14}$$

since the power vector therefore $p_i$ is assumed fixed. The MMSE filter coefficients are given as [30]

$$\boldsymbol{c}_i^* = \frac{\sqrt{p_i}}{1 + p_i \boldsymbol{s}_i^\top \boldsymbol{A}_i^{-1}\boldsymbol{s}_i} \boldsymbol{A}_i^{-1}\boldsymbol{s}_i \tag{5.15}$$

Note that the MMSE solution $\boldsymbol{c}_i^*$ is just a scaled version of $\bar{\boldsymbol{c}}_i$ and the optimization problem of (5.9) is insensitive to the scaling of the vector $\boldsymbol{c}_i$. As a convention we will use the MMSE solution, $\boldsymbol{c}_i^*$, given in (5.15) as the solution of the inner optimization problem instead of $\bar{\boldsymbol{c}}$ of (5.12).

## 5.4 Power Control Algorithm

When we view (5.6) as a set of interference constraints on the power vector $\boldsymbol{p}$, we can define a power control algorithm in which each user $i$ iteratively attempts to compensate for the interference. We define

$$\begin{aligned} I_i(\boldsymbol{p}, \boldsymbol{c}_i) &= \frac{\gamma_i^*}{h_{ii}} \frac{\sum_{j\neq i} p_j h_{ij} (\boldsymbol{c}_i^\top \boldsymbol{s}_j)^2 + \sigma^2 (\boldsymbol{c}_i^\top \boldsymbol{c}_i)}{(\boldsymbol{c}_i^\top \boldsymbol{s}_i)^2} && (5.16) \\ T_i(\boldsymbol{p}) &= \min_{\boldsymbol{c}_i} I_i(\boldsymbol{p}, \boldsymbol{c}_i) && (5.17) \end{aligned}$$

and we propose the power control algorithm

$$\boldsymbol{p}(n+1) = \boldsymbol{T}(\boldsymbol{p}(n)) \tag{5.18}$$

where

$$\boldsymbol{T}(\boldsymbol{p}) = [T_1(\boldsymbol{p}), \cdots, T_N(\boldsymbol{p})]^\top \tag{5.19}$$

Each power control iteration (5.18) includes an optimization of the filter coefficients to maximally suppress the interference. In effect, we choose the filter coefficients to minimize the required transmitter power. This is analogous to integrated power control and base station assignment algorithms [19, 74] in which a user's base station assignment is iteratively chosen to minimize the transmitter power. In [73], power control algorithms of the form

$$\boldsymbol{p}(n+1) = \boldsymbol{I}(\boldsymbol{p}(n)) \tag{5.20}$$

are analyzed for standard interference functions $\boldsymbol{I}(\boldsymbol{p})$. The definition of standard interference functions and the theorem describing the convergence of (5.20) follow.

**Definition 2** *$\boldsymbol{I}(\boldsymbol{p})$ is a standard interference function if for all $\mathbf{p} \geq \mathbf{0}$ the following properties are satisfied.*

- *Positivity: $\boldsymbol{I}(\boldsymbol{p}) > \mathbf{0}$*
- *Monotonicity: If $\mathbf{p} \geq \mathbf{p}'$ then $\boldsymbol{I}(\boldsymbol{p}) \geq \boldsymbol{I}(\boldsymbol{p}')$*
- *Scalability: For all $\alpha > 1$, $\alpha \boldsymbol{I}(\boldsymbol{p}) > \boldsymbol{I}(\alpha \boldsymbol{p})$*

**Theorem 6** *If there exists $\mathbf{p}' \geq \boldsymbol{I}(\mathbf{p}')$, then for any initial power vector $\mathbf{p}(0)$, the sequence $\mathbf{p}(n) = \boldsymbol{I}(\boldsymbol{p}(n-1))$ converges to a unique fixed point $\bar{\mathbf{p}}$ such that $\bar{\mathbf{p}} \leq \mathbf{p}'$ for any $\mathbf{p}' \geq \boldsymbol{I}(\boldsymbol{p}')$.*

The condition that there exists $\boldsymbol{p}' \geq \boldsymbol{I}(\boldsymbol{p}')$ is simply a requirement that a feasible power vector exists. The fixed point $\bar{\boldsymbol{p}}$ is a minimum power solution in that $\bar{\boldsymbol{p}} \leq \boldsymbol{p}'$ for any feasible power vector $\boldsymbol{p}'$. Thus, we prove the convergence of the power control algorithm (5.18) by proving that the transformation $\boldsymbol{T}(\boldsymbol{p})$ is standard.

**Theorem 7** *$\boldsymbol{T}(\boldsymbol{p})$ is a standard interference function.*

**Proof: Theorem 7** From (5.16), for any fixed $\boldsymbol{c}_i$ we have $I_i(\boldsymbol{p}, \boldsymbol{c}_i) > 0$. Therefore, $T_i(\boldsymbol{p}) = \min_{\boldsymbol{c}_i} I_i(\boldsymbol{p}, \boldsymbol{c}_i) > 0$ and $\boldsymbol{T}(\boldsymbol{p})$ is positive. To prove monotonicity, we note for any fixed $\boldsymbol{c}_i$ that $\boldsymbol{p} \geq \boldsymbol{p}'$ implies $I_i(\boldsymbol{p}, \boldsymbol{c}_i) \geq I_i(\boldsymbol{p}', \boldsymbol{c}_i)$. If the minimum of $I_i(\boldsymbol{p}, \boldsymbol{c}_i)$ is achieved at $\boldsymbol{c}_i^*$, then,

$$
\begin{aligned}
T_i(\boldsymbol{p}) &= \min_{\boldsymbol{c}_i} I_i(\boldsymbol{p}, \boldsymbol{c}_i) &(5.21)\\
&= I_i(\boldsymbol{p}, \boldsymbol{c}_i^*) &(5.22)\\
&\geq I_i(\boldsymbol{p}', \boldsymbol{c}_i^*) &(5.23)\\
&\geq \min_{\boldsymbol{c}_i} I_i(\boldsymbol{p}', \boldsymbol{c}_i) = T_i(\boldsymbol{p}') &(5.24)
\end{aligned}
$$

For scalability, we note that for any fixed $\boldsymbol{c}_i$ and $\alpha > 1$ we have $\alpha I_i(\boldsymbol{p}, \boldsymbol{c}_i) > I_i(\alpha\boldsymbol{p}, \boldsymbol{c}_i)$. Assuming again that the minimum of $I_i(\boldsymbol{p}, \boldsymbol{c}_i)$ is achieved at $\boldsymbol{c}_i^*$, we have

$$
\begin{aligned}
\alpha T_i(\boldsymbol{p}) &= \min_{\boldsymbol{c}_i} \alpha I_i(\boldsymbol{p}, \boldsymbol{c}_i) &(5.25)\\
&= \alpha I_i(\boldsymbol{p}, \boldsymbol{c}_i^*) &(5.26)\\
&> I_i(\alpha\boldsymbol{p}, \boldsymbol{c}_i^*) &(5.27)\\
&\geq \min_{\boldsymbol{c}_i} I_i(\alpha\boldsymbol{p}, \boldsymbol{c}_i) = T_i(\alpha\boldsymbol{p}) &(5.28)
\end{aligned}
$$

□

Since $\boldsymbol{T}(\boldsymbol{p})$ is a standard interference function, the power control algorithm (5.18) converges to $\bar{\boldsymbol{p}} = \boldsymbol{T}(\bar{\boldsymbol{p}})$. The filter coefficients converge to $\bar{\boldsymbol{c}}_i = \arg\min_{\boldsymbol{c}_i} I_i(\bar{\boldsymbol{p}}, \boldsymbol{c}_i)$. Equivalently, the power control algorithm converges to a minimum power solution for the SIR target based power control problem with linear receiver filters; and the linear receiver filter converges to the MMSE multiuser detector.

## 5.5 Implementation of the Power Control Algorithm

The power control algorithm (5.18) is implicitly a two stage algorithm. First, we adjust the filter coefficients to be the MMSE coefficients for power vector $\boldsymbol{p}$. Second, we adjust the transmitter powers to meet the SIR constraints for the chosen filter coefficients. In this section, we describe how the iteration (5.18) may be implemented in practice.

We will denote the matrix $\boldsymbol{A}_i$ as $\boldsymbol{A}_i(\boldsymbol{p}(n))$ below in order to emphasize its dependency on the power vector. This matrix is calculated using (5.8) when $\boldsymbol{p}(n)$ is given. At iteration $n+1$, the MMSE filter $\hat{\boldsymbol{c}}_i$ is constructed by using the current power vector $\boldsymbol{p}(n)$ and then the power vector is updated using the new filter coefficients $\hat{\boldsymbol{c}}_i$. The resulting iterative algorithm for user $i$ is

$$\hat{\boldsymbol{c}}_i = \frac{\sqrt{p_i(n)}}{1 + p_i(n)\boldsymbol{s}_i^\top \boldsymbol{A}_i^{-1}(\boldsymbol{p}(n))\boldsymbol{s}_i} \boldsymbol{A}_i^{-1}(\boldsymbol{p}(n))\boldsymbol{s}_i \tag{5.29}$$

$$p_i(n+1) = \frac{\gamma_i^*}{h_{ii}} \frac{\left(\sum_{j \neq i} p_j(n) h_{ij} (\hat{\boldsymbol{c}}_i^\top \boldsymbol{s}_j)^2 + \sigma^2 \hat{\boldsymbol{c}}_i^\top \hat{\boldsymbol{c}}_i\right)}{(\hat{\boldsymbol{c}}_i^\top \boldsymbol{s}_i)^2} \tag{5.30}$$

Equations (5.29) and (5.30) represent a deterministic iteration of the transmitter powers and filter coefficients. If the SIR targets are feasible, then starting from any initial power vector $\boldsymbol{p}$ and filter coefficients $\boldsymbol{c}_1, \ldots, \boldsymbol{c}_N$, the algorithm converges deterministically to the unique minimum power fixed point.

The theoretical properties of the iteration using equations (5.29) and (5.30) are of little practical use if the quantities needed to perform the iteration cannot be determined. Moreover, from (5.29) and (5.30), it would appear that all transmitter powers $p_j$ and channel gains $h_{ij}$ are needed to obtain $\boldsymbol{A}_i$ and hence $\hat{\boldsymbol{c}}_i$. Fortunately, this is not the case. In particular, we can estimate $\boldsymbol{A}_i$ by sampling the received signal before the receiver filters and taking empirical averages. From (5.2), the mutual independence of the zero mean transmitted bits $\{b_n\}$ and the Gaussian noise $\boldsymbol{n}$ implies

$$E\left[\boldsymbol{r}_i \boldsymbol{r}_i^\top\right] = \boldsymbol{A}_i + p_i h_{ii} \boldsymbol{s}_i \boldsymbol{s}_i^\top \tag{5.31}$$

Therefore, $\boldsymbol{r}_i \boldsymbol{r}_i^\top - p_i h_{ii} \boldsymbol{s}_i \boldsymbol{s}_i^\top$ is an unbiased estimate for $\boldsymbol{A}_i$. If at the assigned base station of user $i$, the uplink gain $h_{ii}$ and transmitter power $p_i$ are known, $\boldsymbol{A}_i$ can be estimated by a sample average of $\boldsymbol{r}_i \boldsymbol{r}_i^\top$ over multiple bit intervals. For the adjusted filter coefficients $\hat{\boldsymbol{c}}_i$, equation (5.3) implies that the average squared filter output for user $i$ under power vector $\boldsymbol{p}(n)$ is

$$E\left[y_i^2(n)\right] = \sum_{j=1}^{N} p_j(n) h_{ij} (\hat{\boldsymbol{c}}_i^\top \boldsymbol{s}_j)^2 + \sigma^2 \hat{\boldsymbol{c}}_i^\top \hat{\boldsymbol{c}}_i \tag{5.32}$$

Thus, from (5.30), the power control iteration can be written as

$$p_i(n+1) = \frac{\gamma_i^*}{h_{ii}} \frac{1}{(\hat{\boldsymbol{c}}_i^\top \boldsymbol{s}_i)^2} \left(E\left[y_i^2(n)\right] - p_i(n) h_{ii} (\hat{\boldsymbol{c}}_i^\top \boldsymbol{s}_i)^2\right) \tag{5.33}$$

A simple measurement based power control algorithm can use a sample average of $y_i^2(n)$ over multiple bit intervals to estimate $E[y_i^2(n)]$. We have presented these simple estimation methods not because they perform particularly well but rather to emphasize that the information needed for user $i$ to implement the MMSE power control is available at the receiver for user $i$. Thus, distributed implementation is possible. We note that the simple estimation methods still require a user to estimate its own uplink gain $h_{ii}$. This can be done, perhaps roughly, using the downlink transmission of a base station pilot tone. Alternatively, estimating the uplink gain can be avoided by direct estimation of the SIR without separate estimates of the signal and interference components [3, 4]. This is also the motivation for the MMSE power control algorithm of [24] that uses measurements of the mean squared error.

Although we have verified that the proposed power control can be implemented in a distributed manner using local measurements, we note that substitution of stochastic measurements does not preserve the deterministic convergence properties. Furthermore, the direct substitution of measured estimates in a deterministic algorithm may not be the most desirable approach. We emphasize that the need for measurements, and the consequent difficulty of analyzing the effect of measurements, is a property of all power control algorithms, whether or not those algorithms adapt the filter coefficients for interference suppression.

In particular, it may be preferable to use separate iterative algorithms for

1. the adaptation of the filter coefficients to the MMSE coefficients $\hat{\boldsymbol{c}}_i$

2. the iterative transmitter power adjustments for fixed filter coefficients.

For the first step, iterative algorithms that converge to the MMSE filter coefficients are given in [30] or the blind adaptive multiuser detector [21] can be used to converge stochastically to the MMSE solution. For fixed filter coefficients, the second step is equivalent to the conventional power control problem reviewed in Chapter 4. An alternate approach is the stochastic power control algorithm [57] given in Chapter 4. In Chapter 4, the power vector was shown to converge stochastically to the optimal power vector by using the random outputs of the fixed receiver filters. Therefore, alternating between the blind adaptive MMSE detector and the stochastic power control algorithm would yield a stochastically converging power control algorithm. However, the convergence of this two step iteration may be slow. We believe a combined stochastic optimization of filter coefficients and transmitter powers may have better convergence properties and should be investigated.

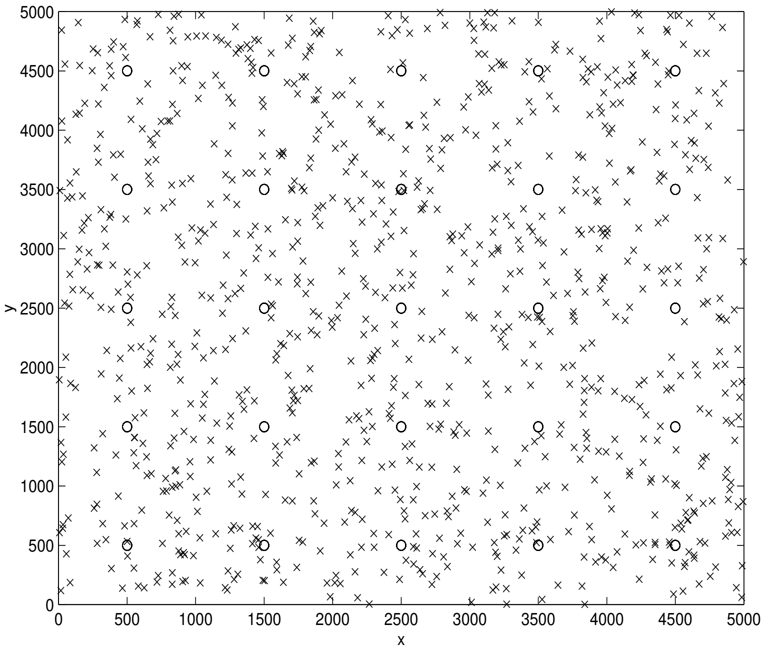


Figure 5.1: Simulation environment for $N = 1000$. Symbols o and x denote the base stations and the users, respectively.

## 5.6 Simulation Results

In our simulations we consider a multicell CDMA system on a rectangular grid. There are $M = 25$ base stations with $(x, y)$ coordinates $(1000i+500, 1000j+500)$ for $0 \leq i, j \leq 4$. The $x$ and $y$ coordinates of each user are independent uniformly distributed random variables between 0 and 5,000 meters. The experiments are conducted for $N = 250, 500$ and 1000 users. Figure 5.1 shows the positions of users and the base stations with symbols x and o, respectively, for $N = 1000$. Each user is assigned to its nearest base station. The path loss exponent used while calculating the channel gains of the users is taken to be $\alpha = 4$. At the beginning of the iterations, power vector is initialized to zero, and the filter coefficients are initialized to the signature sequences of the users (i.e., $p_i(0) = 0$ and $\boldsymbol{c}_i(0) = \boldsymbol{s}_i$).

We chose the processing gain to be $G = 150$ and a random signature sequence of

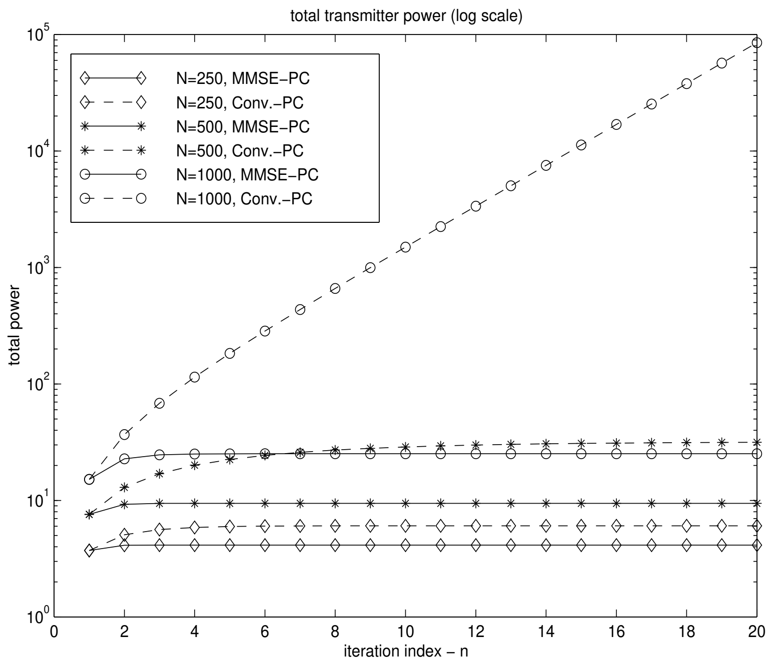


Figure 5.2: Total transmitter power for the conventional power control algorithm (Conv.-PC) and the MMSE power control algorithm (MMSE-PC) for $N = 250, 500$ and 1000.

length $G$ chips was assigned to each user. Although the convergence theorems permit individual SIR targets $\gamma_i^*$ for each user $i$, for the simulations we chose a common SIR target $\gamma_i^* = 4$ ($\approx$ 6 dB) for all users. The AWGN noise power equaled $\sigma^2 = 10^{-13}$, corresponding roughly to a 1 MHz bandwidth.

We compared the performance of the conventional power control algorithm which assumes a conventional detector structure composed of the filters matched to the signature sequences of the users, and the power control algorithm proposed in this chapter which optimizes the filter coefficients in addition to updating the powers. Since the filter coefficients are always chosen to be the MMSE detector, we call the proposed algorithm the MMSE power control. We compared the deterministic convergence of the conventional and MMSE power control algorithms.

Figure 5.2 shows in log scale the total transmitter power $\sum_{i=1}^{N} p_i$, as a function of the iteration index, for the MMSE and conventional power control algorithms. We observe that the MMSE power control outperforms the conventional power control in terms of total received power, and convergence rate. Using MMSE power control, the

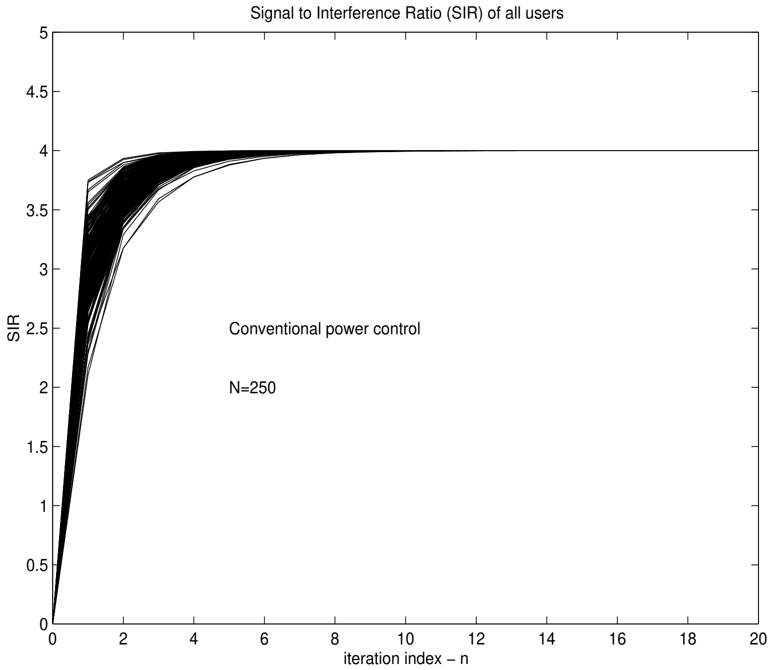


Figure 5.3: Signal to Interference Ratio (SIR) of all the users as a function of $n$ for the conventional power control algorithm. SIR target value $\gamma_i^* = 4$ ($\approx$ 6 dB) for $i = 1, \ldots, N$, number of users $N = 250$.

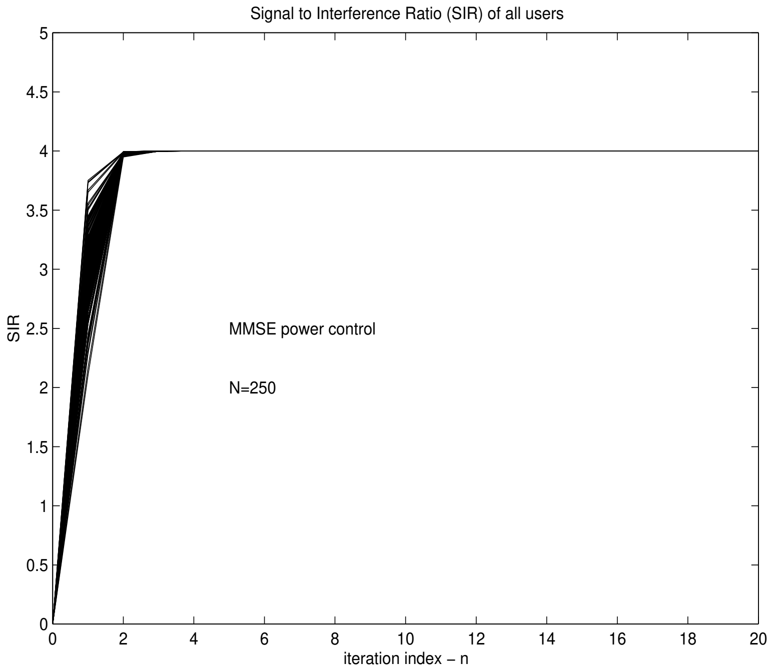


Figure 5.4: Signal to Interference Ratio (SIR) of all the users as a function of $n$ for the MMSE power control algorithm. SIR target value $\gamma_i^* = 4$ ($\approx$ 6 dB) for $i = 1, \ldots, N$, number of users $N = 250$.

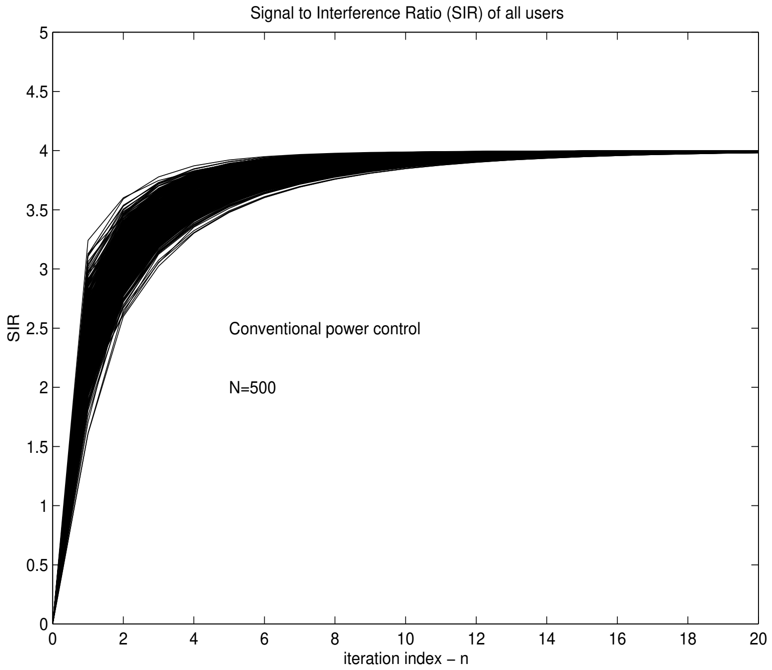


Figure 5.5: Signal to Interference Ratio (SIR) of all the users as a function of $n$ for the conventional power control algorithm. SIR target value $\gamma_i^* = 4$ ($\approx$ 6 dB) for $i = 1, \ldots, N$, number of users $N = 500$.

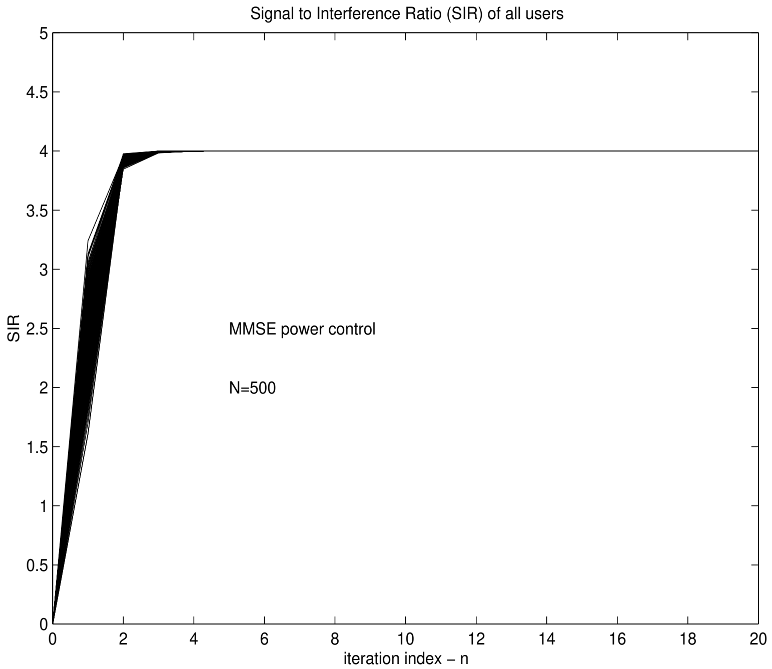


Figure 5.6: Signal to Interference Ratio (SIR) of all the users as a function of $n$ for the MMSE power control algorithm. SIR target value $\gamma_i^* = 4$ ($\approx$ 6 dB) for $i = 1, \ldots, N$, number of users $N = 500$.

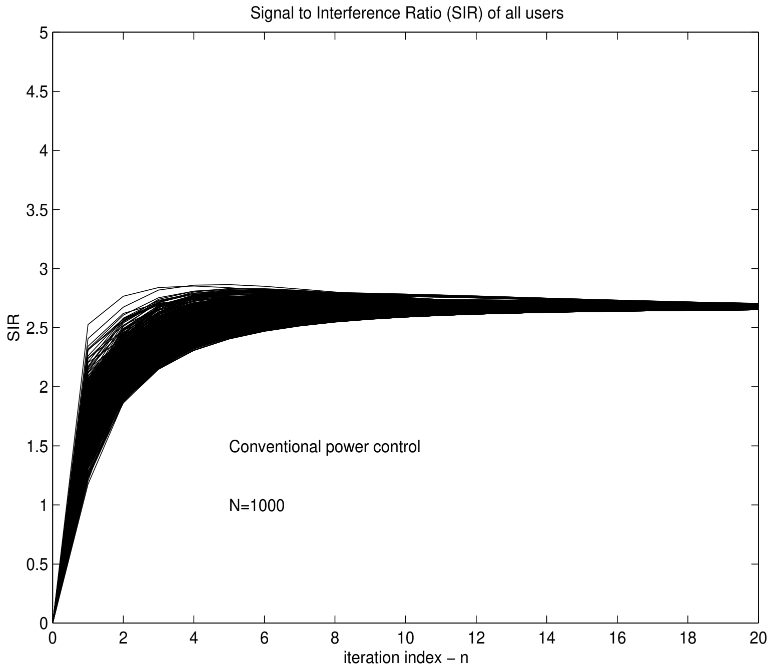


Figure 5.7: Signal to Interference Ratio (SIR) of all the users as a function of $n$ for the conventional power control algorithm. SIR target value $\gamma_i^* = 4$ ($\approx$ 6 dB) for $i = 1, \ldots, N$, number of users $N = 1000$.

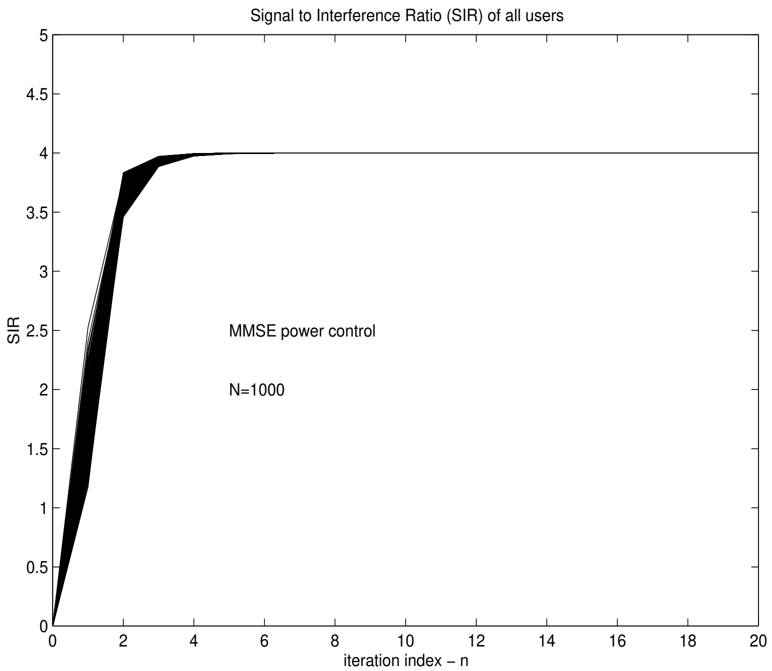


Figure 5.8: Signal to Interference Ratio (SIR) of all the users as a function of $n$ for the MMSE power control algorithm. SIR target value $\gamma_i^* = 4$ ($\approx$ 6 dB) for $i = 1, \ldots, N$, number of users $N = 1000$.

total transmitter power is less than that needed for the conventional detector. The savings in total transmit power increase with increasing number of users. Also, the MMSE power control algorithm converges to the optimal power vector faster than the conventional power control algorithm.

The steadily increasing transmitter power curve for conventional power control with $N = 1000$ in Figure 5.2 occurs because the conventional power control problem is infeasible. For this case, updating the receiver filter coefficients converted the infeasible conventional power control problem into a feasible problem.

In order to observe the convergence of the SIRs to the common target SIR, we plotted the SIRs of all of the users in Figures 5.3 and 5.4 for the conventional power control algorithm and the MMSE power control algorithm, respectively, for $N = 250$ users. Figures 5.5, 5.6 and Figures 5.7, 5.8 show the same graphs produced for $N = 500$ and $N = 1000$ users, respectively.

We observe from Figures 5.3, 5.4 and Figures 5.5, 5.6 that when the MMSE power control is used, the SIRs converge to the common target SIR faster than with the conventional power control algorithm. We again observe the infeasibility of the target SIR from Figure 5.7, by noting that the SIRs of the users converge to the values which are less than the target value ($\gamma_i^* = 4$). We note that the SIRs converge to the maximum achievable common SIR target with fixed system parameters such as channel gains, cross correlations between the signature sequences; see [16, 79].

## 5.7 Chapter Summary and Conclusions

Power control and multiuser detection methods have been proposed and studied independently as means of overcoming the near-far problem of the CDMA scheme. In this chapter we combined the power control and multiuser detection approaches to cope with the near-far problem more effectively. We proposed an iterative and distributed algorithm which controls both the transmitter powers and the receiver filters of the users. At each iteration, the proposed algorithm first updates the receiver filters of the users to be the MMSE filters and then updates transmitter powers of the users to overcome the residual interference.

We showed that the proposed algorithm converges to a minimum power solution where all the users satisfy their SIR based quality of service requirements; and that the linear receiver filters converge to the MMSE multiuser detectors.

We observed that the proposed algorithm, which is named MMSE power control algorithm, is superior in terms of the total transmitter power and convergence rate when compared with the conventional power control algorithm. With MMSE power control, the same system performance is achieved with less total transmitter power, increasing the capacity of the CDMA system when compared with the conventional power control. Since MMSE power control can convert a power control problem that is infeasible with conventional power control into a feasible one, it increases the system capacity by allowing the SIR objectives of the users to be higher, or by increasing the number of users supportable at a fixed SIR level.

# Chapter 6

# Iterative Construction of Optimum Sequence Sets to Maximize the Capacity of CDMA Systems

In Chapter 5, we considered joint optimization of a multicell CDMA system over the space of transmitter powers and receiver filters for a fixed set of signature sequences. We assumed a given set of signature sequences and proposed an algorithm which updated transmitter powers and linear receiver filters of the users to minimize the total transmitter power subject to SIR based quality of service requirements. We can observe from Chapter 5 that the required transmitter powers can be further decreased if signature sequences of the users could also be optimized instead of being assumed to be given and fixed. In this chapter, we will concentrate on maximization of the capacity of a single cell CDMA system by signature sequence selection.

The maximization of the capacity of a single cell CDMA system has been studied in two different contexts in [42,65]. In [42] and [65], signature sequence sets that maximize the *information theoretic* capacity and *network* capacity, respectively, of CDMA systems were identified. In both cases, it was shown that if the number of users is less than the processing gain then orthogonal signature sequences were optimal. In addition, it was shown that if the number of users is larger than the processing gain then signature sequences satisfying the Welch bound [70] on the total squared correlation with equality (called WBE sequences) should be used in order to maximize the capacity of a single cell synchronous CDMA system.

This chapter presents an algorithm which iteratively updates the signature sequences of the users in a distributed fashion, starting from an initial set of signature sequences. We prove that the algorithm converges to a set of orthogonal signature sequences if the number of users is less than the processing gain, and to a WBE set if the number of users is larger than the processing gain. We prove the convergence of the algorithm under very mild conditions on the initial set of signature sequences at which the algorithm is started. At each step, the algorithm replaces one signature sequence from the set with generalized normalized MMSE receiver corresponding to that signature sequence. Since the MMSE filter can be obtained distributedly for each user, the proposed algorithm is amenable to distributed implementation.

## 6.1 System Model

In this Chapter, we consider the uplink of a single cell synchronous CDMA system with $N$ users. The signature waveform of user $i$, denoted by $s_i(t)$, 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 received base band signal at the base station is

$$r(t) = \sum_{i=1}^{N} \sqrt{p_i} b_i s_i(t) + n(t) \tag{6.1}$$

where $p_i$ and $b_i$ are the received power and the information bit of user $i$, and $n(t)$ is the additive white Gaussian noise (AWGN) with zero mean and power spectral density $\sigma^2$.

We define the chip waveform to be $\psi(t)$, $t \in [0, T_c]$ and 0 elsewhere, where $T_c$ is the chip duration. Thus $\{\psi(t - iT_c), i = 0, \ldots, G-1\}$, where $G = T_b/T_c$ is the processing gain, is a basis for the signal space. This allows us to represent the signature sequences of the users with $G$ dimensional vectors. We will use the $G$-dimensional vector $\boldsymbol{s}_i$ to denote the signature sequence of user $i$. We define an $G \times N$ matrix $\boldsymbol{S}$ which contains the signature sequences of all users in its columns, i.e., $\boldsymbol{S} = [\boldsymbol{s}_1, \boldsymbol{s}_2, \cdots, \boldsymbol{s}_N]$ and a $N \times N$ diagonal matrix $\boldsymbol{P}$ whose $i$th diagonal element is $p_i$. In terms of signal vectors, the received signal at the base station can be written as

$$\boldsymbol{r} = \sum_{i=1}^{N} \sqrt{p_i} b_i \boldsymbol{s}_i + \boldsymbol{n} \tag{6.2}$$

where $\boldsymbol{n}$ is a Gaussian random vector with $E\Big[\boldsymbol{n}\boldsymbol{n}^\top\Big] = \sigma^2 \boldsymbol{I}_G$, where $\boldsymbol{I}_G$ is the $G \times G$ identity matrix.

## 6.2 Capacity of a Single Cell CDMA System

The capacity of a single cell synchronous CDMA system has been studied in two different contexts. The *information theoretic* capacity of a CDMA channel is given as [42,61]

$$\mathcal{C} = \bigcap_{J \subseteq \{1,\cdots,N\}} \left\{ (R_1, \cdots, R_N) : 0 \leq \sum_{j \in J} R_j \leq \frac{1}{2} \log \left[ \det \left( \boldsymbol{I}_G + \sigma^{-2} \boldsymbol{S}_J \boldsymbol{P}_J \boldsymbol{S}_J^\top \right) \right] \right\} \tag{6.3}$$

where $\boldsymbol{S}_J$ is the $G \times |J|$ matrix which is obtained from $\boldsymbol{S}$ by striking out the columns of $\boldsymbol{S}$ whose indices do not belong to the subset $J$, and $\boldsymbol{P}_J$ is a $|J| \times |J|$ diagonal matrix which is obtained by striking out the rows and columns of $\boldsymbol{P}$ whose indices do not belong to the subset $J$. In other words, $\boldsymbol{S}_J$ and $\boldsymbol{P}_J$ are the matrices of signature sequences and received powers of the users in the subset $J$. Note that for any two matrices $\boldsymbol{A}_{K \times M}$ and $\boldsymbol{B}_{M \times K}$

$$\det(\boldsymbol{I}_K + \boldsymbol{AB}) = \det(\boldsymbol{I}_M + \boldsymbol{BA}) \tag{6.4}$$

and the determinant in (6.3) could be written as

$$\det\left(\boldsymbol{I}_G + \sigma^{-2} \boldsymbol{S}_J \boldsymbol{P}_J \boldsymbol{S}_J^\top\right) = \det\left(\boldsymbol{I}_{|J|} + \sigma^{-2} \boldsymbol{P}_J \boldsymbol{S}_J^\top \boldsymbol{S}_J\right) = \det\left(\boldsymbol{I}_{|J|} + \sigma^{-2} \boldsymbol{P}_J \boldsymbol{\Gamma}_J\right) \tag{6.5}$$

as was done in [61]. Here $\boldsymbol{\Gamma}_J$ represents the cross correlation matrix of the signature sequences of the users whose indices are in the subset $J$.

An important measure of overall information capacity of a multiaccess channel is the *sum* capacity,

$$C_{\text{sum}} = \max_{(R_1,\cdots,R_N) \in \mathcal{C}} \sum_{i=1}^{N} R_i \tag{6.6}$$

From (6.3), the sum capacity of a CDMA channel is

$$C_{\text{sum}} = \frac{1}{2} \log \left[ \det \left( \boldsymbol{I}_G + \sigma^{-2} \boldsymbol{SPS}^\top \right) \right] \tag{6.7}$$

When the powers of the users are the same, $p_i = p$ for all $i$, (6.7) reduces to

$$C_{\text{sum}} = \frac{1}{2} \log \left[ \det \left( \boldsymbol{I}_G + \frac{p}{\sigma^2} \boldsymbol{SS}^\top \right) \right] \tag{6.8}$$

Using (6.4), the sum capacity expression in (6.8) can also be written as

$$C_{\text{sum}} = \frac{1}{2} \log \left[ \det \left( \boldsymbol{I}_N + \frac{p}{\sigma^2} \boldsymbol{S}^\top \boldsymbol{S} \right) \right] \tag{6.9}$$

Thus, the equivalent expressions in (6.8) and (6.9) determine the information theoretic sum capacity of a CDMA system with equal received powers. Clearly, the sum capacity depends on the signature sequence set being used. It was shown in [42] that the sum capacity is maximized if the signature sequences are chosen such that, if $N \leq G$

$$\boldsymbol{S}^{\top}\boldsymbol{S} = \boldsymbol{I}_N \tag{6.10}$$

and if $N > G$

$$\boldsymbol{S}\boldsymbol{S}^{\top} = \frac{N}{G}\boldsymbol{I}_G \tag{6.11}$$

The signature sequence sets satisfying (6.10) contain $N$ orthogonal signature sequences in $G$ dimensional vector space. When $N \leq G$, $N$ such sequences can be found. Note that there are infinitely many such sets in the continuous $\mathcal{R}^G$ space. The sequence sets satisfying (6.11) are named Welch Bound Equality (WBE) sequence sets in [42] because they satisfy the Welch's bound on the sum of the squares of the cross correlations of unit energy sequences with equality. A brief introduction on Welch's bound will follow in the next section.

In [65], *network* capacity of a CDMA system is defined in terms of the maximum number of admissible users, given the common quality of service measure defined in terms of the SIR, and the processing gain $G$. For a given processing gain ($G$) and a common SIR target ($\beta$), $N$ users are said to be admissible if there exist positive powers $p_i$ and signature sequences $\boldsymbol{s}_i$ such that all users have SIRs that are at least as large as the target SIR ($\mathrm{SIR}_i \geq \beta$, for $i = 1, \cdots, N$). The network capacity of a CDMA system was found for two kinds of linear receiver structures in [65]: matched filters and MMSE filters [30, 71]. Given the powers and the signature sequences of the users, the MMSE filter for the $i$th user, $\boldsymbol{c}_i$, is found by minimizing the MSE between the information bit and the filter output of user $i$

$$\mathrm{MSE}_i = E\Big[(\boldsymbol{r}^{\top}\boldsymbol{c}_i - b_i)^2\Big] = \boldsymbol{c}_i^{\top}\boldsymbol{B}\boldsymbol{c}_i - 2\sqrt{p_i}\boldsymbol{c}_i^{\top}\boldsymbol{s}_i + 1 \tag{6.12}$$

where

$$\boldsymbol{B} = \sum_{j=1}^{N} p_j \boldsymbol{s}_j \boldsymbol{s}_j^{\top} + \sigma^2 \boldsymbol{I}_G = \boldsymbol{S}\boldsymbol{P}\boldsymbol{S}^{\top} + \sigma^2 \boldsymbol{I}_G \tag{6.13}$$

The MMSE filter for the $i$th user is given as

$$\boldsymbol{c}_i = \arg\min_{\boldsymbol{c}_i} \text{MSE}_i = \sqrt{p_i}\boldsymbol{B}^{-1}\boldsymbol{s}_i \tag{6.14}$$

Using the matrix inversion formula

$$\left(\boldsymbol{M} + \boldsymbol{u}\boldsymbol{v}^\top\right)^{-1} = \boldsymbol{M}^{-1} - \frac{\boldsymbol{M}^{-1}\boldsymbol{u}\boldsymbol{v}^\top\boldsymbol{M}^{-1}}{1 + \boldsymbol{v}^\top\boldsymbol{M}^{-1}\boldsymbol{u}} \tag{6.15}$$

the MMSE solution can also be written as

$$\boldsymbol{c}_i = \frac{\sqrt{p_i}}{1 + p_i\boldsymbol{s}_i^\top\boldsymbol{A}_i^{-1}\boldsymbol{s}_i}\boldsymbol{A}_i^{-1}\boldsymbol{s}_i \tag{6.16}$$

where

$$\boldsymbol{A}_i = \sum_{j\neq i}^{N} p_j\boldsymbol{s}_j\boldsymbol{s}_j^\top + \sigma^2\boldsymbol{I}_G \tag{6.17}$$

The MMSE filter is also called the *optimum* linear receiver in [65] since the MMSE filter maximizes the SIR for a given received power vector.

It was found in [65] that the network capacity of a CDMA system with MMSE receivers is limited by the inequality

$$N < G\left(1 + \frac{1}{\beta}\right) \tag{6.18}$$

Note that (6.7) determines the capacity of a single cell CDMA system in terms of the maximum number of information bits that can be delivered by all of the terminals on the uplink to the base station; and (6.18) determines the capacity of a single cell CDMA system in terms of the maximum number of users that can communicate with the base station at a certain reliability (SIR $\geq$ target SIR, or equivalently bit error rate (BER) $\leq$ target BER) level on the uplink. It was found in [65] that for a single cell CDMA system, the network capacity with MMSE receivers is maximized if the signature sequence set is chosen to satisfy (6.10) if $N \leq G$ and (6.11) if $N > G$, and if the received powers of the users are chosen to be the same for all users. Note that when the signature sequences satisfy (6.10), the MMSE receivers reduce to the matched filters. Similarly, when the received powers of the users are the same, the matrix $\boldsymbol{B}$ defined in (6.13) becomes $\boldsymbol{B} = p\boldsymbol{S}\boldsymbol{S}^\top + \sigma^2\boldsymbol{I}_G$, and from (6.14), if the the signature sequence set satisfies (6.11), the MMSE receivers for all users reduce to the scaled versions of the matched filters. It was also shown in [65] that when the matched filters are used as the receivers, the network capacity is again given by (6.18) and the maximum is achieved by equal powers and signature sequence sets satisfying (6.10) and (6.11) for $N \leq G$ and $N > G$, respectively. As was noted in [65] this result is expected because once the signature set is chosen to satisfy (6.10) and (6.11), for $N \leq G$ and $N > G$, respectively, the MMSE receivers are scaled matched filters. An important consequence of this work is that in the joint optimization of signature sequences, powers and receiver filters, once the optimum signature sequences are used, there is no need for an optimization over the receiver filters; single user detectors (matched filters) are optimum. Therefore, the joint power/signature sequence/receiver filter optimization problem actually has only two "degrees of freedom": the powers and the signature sequences of the users.

Although it is known that the sequences which satisfy (6.10) for $N \leq G$ and (6.11) for $N > G$ maximize the capacity of a single cell CDMA system whether it is defined as information theoretic capacity or as network capacity, no algorithmic scheme is known to construct these sequence sets for general $N$ and $G$. Note that when $N \leq G$, a simple Gram-Schmidt orthogonalization procedure would yield $N$ orthogonal vectors starting with $N$ linearly independent vectors. For $N > G$, no such simple construction scheme exists. In the following section we will first describe some useful properties of WBE sequences and make a simple observation which will be the basis of the iterative construction scheme we will propose in Section 6.4.

## 6.3 WBE Sequences and a Simple Observation

Welch developed lower bounds for the $2k$th power of the maximum correlation among a set of $N$ unit energy sequences [70]. These lower bounds on the $2k$th powers of the maximum correlation were actually obtained from a lower bound on the sum of the $2k$th powers of the correlations of a sequence set.

$$\sum_{i=1}^{N}\sum_{j=1}^{N}(\boldsymbol{s}_i^\top \boldsymbol{s}_j)^{2k} \geq \frac{N^2}{\begin{pmatrix} G+k-1 \\ k \end{pmatrix}} \tag{6.19}$$

For $k = 1$, Welch's bound reduces to

$$\sum_{i=1}^{N}\sum_{j=1}^{N}(\boldsymbol{s}_i^\top \boldsymbol{s}_j)^2 \geq \frac{N^2}{G} \tag{6.20}$$

For a simpler derivation of the bound in (6.20) see [31, 32]. Note that sequence sets satisfying (6.11) satisfy the bound in (6.20) with equality. Note also that this bound is loose for $N < G$, because in this case the best we can do in terms of the total squared correlation is $N$ (by choosing $N$ orthogonal vectors) and the bound $N^2/G = N(N/G) < N$ cannot be achieved.

Consider a set of vectors, $\boldsymbol{s}_i$, $i = 1, \cdots, N$ used as signature sequences by $N$ users. Assume that the signals of the users at the base station are received by matched filters. Without loss of generality we will assume that $p_i = p$, for all $i$. In this case, the mean squared error for the $i$th user is

$$\text{MSE}_i = E\left[\left(\boldsymbol{r}^\top \boldsymbol{s}_i - b_i\right)^2\right] = \boldsymbol{s}_i^\top \boldsymbol{B} \boldsymbol{s}_i - 2\sqrt{p}\boldsymbol{s}_i^\top \boldsymbol{s}_i + 1 \tag{6.21}$$

where $\boldsymbol{r}$ is the received signal given in (6.2) and $\boldsymbol{B} = p\sum_{j=1}^{N} \boldsymbol{s}_j \boldsymbol{s}_j^\top + \sigma^2 \boldsymbol{I}_G$. The total mean squared error in the system is

$$\begin{aligned}
\text{MSE} &= \sum_{i=1}^{N} \text{MSE}_i \\
&= \sum_{i=1}^{N} \boldsymbol{s}_i^\top \left(p\sum_{j=1}^{N} \boldsymbol{s}_j \boldsymbol{s}_j^\top + \sigma^2 \boldsymbol{I}_G\right) \boldsymbol{s}_i - 2\sqrt{p}\sum_{i=1}^{N} \boldsymbol{s}_i^\top \boldsymbol{s}_i + N \\
&= p\sum_{i=1}^{N}\sum_{j=1}^{N}(\boldsymbol{s}_i^\top \boldsymbol{s}_j)^2 - (2\sqrt{p} - \sigma^2)\sum_{i} \boldsymbol{s}_i^\top \boldsymbol{s}_i + N
\end{aligned} \tag{6.22}$$

Defining total squared correlation (TSC) as

$$\text{TSC} = \sum_{i=1}^{N}\sum_{j=1}^{N}(\boldsymbol{s}_i^\top \boldsymbol{s}_j)^2 \tag{6.23}$$

we observe that

$$\text{MSE} = p\text{TSC} - (2\sqrt{p} - \sigma^2)\sum_{i} \boldsymbol{s}_i^\top \boldsymbol{s}_i + N \tag{6.24}$$

Thus when the signature sequences are restricted to be of unit energy

$$\text{MSE} = p\text{TSC} + (1 + \sigma^2 - 2\sqrt{p})N \tag{6.25}$$

and minimizing TSC subject to $\boldsymbol{s}_i^\top \boldsymbol{s}_i = 1$ is equivalent to minimizing MSE subject to the same condition. Since the sequence sets satisfying (6.10) for $N \leq G$, and (6.11) for $N > G$ minimize the TSC, they minimize the MSE as well. In other words, orthogonal sequences for $N \leq G$ and WBE sequences for $N > G$ are the *global optimal* solutions of the following two equivalent problems

$$\begin{array}{ll} \min & \text{TSC} \\ \text{s. t.} & \boldsymbol{s}_i^\top \boldsymbol{s}_i = 1 \quad \forall i \end{array} \qquad\qquad \begin{array}{ll} \min & \text{MSE} \\ \text{s. t.} & \boldsymbol{s}_i^\top \boldsymbol{s}_i = 1 \quad \forall i \end{array} \tag{6.26}$$

In the following, we will show that if we replace any one of the $N$ sequences in a set with the normalized MMSE filter corresponding to that sequence, the resulting set of sequences will have smaller a TSC. This simple observation will constitute the basis for the iterative algorithm we will propose in the next section. Although working with MSE would yield the same result, we will choose to work with TSC.

We first separate the terms that depend on a particular signature sequence, $\boldsymbol{s}_k$, in the TSC expression

$$\text{TSC} = (\boldsymbol{s}_k^\top \boldsymbol{s}_k)^2 + 2\boldsymbol{s}_k^\top \left( \sum_{j \neq k} \boldsymbol{s}_j \boldsymbol{s}_j^\top \right) \boldsymbol{s}_k + \sum_{i \neq k} \sum_{j \neq k} (\boldsymbol{s}_i^\top \boldsymbol{s}_j)^2 \tag{6.27}$$

Since we will always restrict ourselves to unit energy signature sequences, i.e., $\boldsymbol{s}_k^\top \boldsymbol{s}_k = 1$, we can add $2a^2 \boldsymbol{s}_k^\top \boldsymbol{s}_k$ and subtract $2a^2$ in the above expression to get

$$\text{TSC} = (\boldsymbol{s}_k^\top \boldsymbol{s}_k)^2 + 2\boldsymbol{s}_k^\top \left( \sum_{j \neq k} \boldsymbol{s}_j \boldsymbol{s}_j^\top + a^2 \boldsymbol{I}_G \right) \boldsymbol{s}_k + \sum_{i \neq k} \sum_{j \neq k} (\boldsymbol{s}_i^\top \boldsymbol{s}_j)^2 - 2a^2 \tag{6.28}$$

Let us replace the signature sequence of a user $k$ with another unit energy vector $\boldsymbol{c}_k$ where

$$\boldsymbol{c}_k = \frac{\boldsymbol{A}_k^{-1} \boldsymbol{s}_k}{(\boldsymbol{s}_k^\top \boldsymbol{A}_k^{-2} \boldsymbol{s}_k)^{1/2}} \tag{6.29}$$

where $\boldsymbol{A}_k = \sum_{j \neq k} \boldsymbol{s}_j \boldsymbol{s}_j^\top + a^2 \boldsymbol{I}_G$. Thus we map the set of signature sequences $\boldsymbol{S}$ to another set of signatures $\boldsymbol{S}'$ where

$$\boldsymbol{S} = [\boldsymbol{s}_1, \boldsymbol{s}_2, \cdots, \boldsymbol{s}_{k-1}, \boldsymbol{s}_k, \boldsymbol{s}_{k+1}, \cdots, \boldsymbol{s}_N] \quad \text{and} \quad \boldsymbol{S}' = [\boldsymbol{s}_1, \boldsymbol{s}_2, \cdots, \boldsymbol{s}_{k-1}, \boldsymbol{c}_k, \boldsymbol{s}_{k+1}, \cdots, \boldsymbol{s}_N]$$

Note that $\boldsymbol{c}_k$ is the normalized MMSE filter for user $k$. Also note that this MMSE filter is a generalized one; $\boldsymbol{c}_k$ is the normalized MMSE filter for user $k$ in a CDMA system where the variance of the AWGN is $a^2$, and each other user $j$ transmits with signature sequence $\boldsymbol{s}_j$ and has received power $p_j = 1$. Note that if we worked with $p$TSC in stead of TSC, and chose $a = \sigma$, then $\boldsymbol{A}_k$ would have been $\boldsymbol{A}_k = p\sum_{j\neq k}\boldsymbol{s}_j\boldsymbol{s}_j^\top + \sigma^2\boldsymbol{I}_G$, and the normalized MMSE filter $\boldsymbol{c}_k$ given in (6.29) would have corresponded to the actual normalized MMSE filter in the channel described by (6.2) when $p_j = p$ for all $j$. It will be apparent in what follows that any generalized normalized MMSE filter will be as good as any other in terms of constructing optimum sequences.

Let us denote the total squared error of set $\boldsymbol{S}$ with TSC and total squared error of set $\boldsymbol{S}'$ with TSC$'$. Then

$$\begin{aligned} \text{TSC} &= (\boldsymbol{s}_k^\top\boldsymbol{s}_k)^2 + 2\boldsymbol{s}_k^\top\boldsymbol{A}_k\boldsymbol{s}_k + C - 2a^2 \\ \text{TSC}' &= (\boldsymbol{c}_k^\top\boldsymbol{c}_k)^2 + 2\frac{\boldsymbol{s}_k^\top\boldsymbol{A}_k^{-1}\boldsymbol{s}_k}{\boldsymbol{s}_k^\top\boldsymbol{A}_k^{-2}\boldsymbol{s}_k} + C - 2a^2 \end{aligned} \tag{6.30}$$

where

$$C = \sum_{i\neq k}\sum_{j\neq k}(\boldsymbol{s}_i^\top\boldsymbol{s}_j)^2 \tag{6.31}$$

represents the squared correlation terms that are not affected by the update: $\boldsymbol{s}_k \to \boldsymbol{c}_k$. Note that, by the nature of the mapping in (6.29), $\boldsymbol{c}_k^\top\boldsymbol{c}_k = \boldsymbol{s}_k^\top\boldsymbol{s}_k = 1$. The following theorem verifies that replacing a particular sequence with its normalized MMSE receiver cannot increase the total squared correlation of the set. In other words, in terms of the TSC (or MSE) criteria, $\boldsymbol{S}'$ is a "better" set of signature sequences than $\boldsymbol{S}$.

**Theorem 8** *Given a set of unit energy vectors (signature sequences), $\boldsymbol{s}_i$, $i = 1, \cdots, N$, if any one of the vectors, say $\boldsymbol{s}_k$, is replaced with the normalized MMSE receiver for that vector, $\boldsymbol{c}_k$, the total squared correlation decreases: TSC$'$ $\leq$ TSC. The inequality is satisfied with equality, i.e., TSC$'$ = TSC, iff $\boldsymbol{c}_k = \boldsymbol{s}_k$.*

A proof of Theorem 8 is provided in the Appendix in Section 6.A.

## 6.4 An Iterative Algorithm

We observed in the previous section that given a set of unit energy vectors, if any one of these vectors is replaced with the corresponding normalized MMSE vector, then the

TSC of the set decreases. We start with $N$ unit length vectors $\boldsymbol{S}(0) = [\boldsymbol{s}_1(0), \cdots, \boldsymbol{s}_N(0)]$ at time 0. At iteration $n+1$ the algorithm replaces the vectors $\boldsymbol{S}(n) = [\boldsymbol{s}_1(n), \cdots, \boldsymbol{s}_N(n)]$ with their corresponding normalized MMSE filters one by one, and yields $\boldsymbol{S}(n+1) = [\boldsymbol{s}_1(n+1), \cdots, \boldsymbol{s}_N(n+1)]$. A complete iteration includes $N$ intermediate steps. At the $k$th intermediate step in iteration $n+1$, first $k-1$ vectors have already been updated, i.e., the current vector set has the form of $\boldsymbol{S}_{k-1}(n+1) = [\boldsymbol{s}_1(n+1), \cdots, \boldsymbol{s}_{k-1}(n+1), \boldsymbol{s}_k(n), \boldsymbol{s}_{k+1}(n), \cdots, \boldsymbol{s}_N(n)]$, and the $k$th vector is updated according to

$$\boldsymbol{s}_k(n+1) = \frac{\boldsymbol{A}_k^{-1}(n+1)\boldsymbol{s}_k(n)}{[\boldsymbol{s}^\top(n)\boldsymbol{A}_k^{-2}(n+1)\boldsymbol{s}_k(n)]^{1/2}} \tag{6.32}$$

to yield the vector set $\boldsymbol{S}_k(n+1) = [\boldsymbol{s}_1(n+1), \cdots \boldsymbol{s}_{k-1}(n+1), \boldsymbol{s}_k(n+1), \boldsymbol{s}_{k+1}(n), \cdots, \boldsymbol{s}_N(n)]$. The matrix $\boldsymbol{A}_k(n+1)$ in (6.32) is given as

$$\boldsymbol{A}_k(n+1) = \sum_{j<k} \boldsymbol{s}_j(n+1)\boldsymbol{s}_j^\top(n+1) + \sum_{j>k} \boldsymbol{s}_j(n)\boldsymbol{s}_j^\top(n) + a^2\boldsymbol{I}_G \tag{6.33}$$

## 6.5 Convergence of the Proposed Algorithm

Let $\mathrm{TSC}(n)$ denote the TSC at the end of iteration $n$ (for set $\boldsymbol{S}(n)$) and $\mathrm{TSC}_k(n)$ denote the TSC after the $k$th step in iteration $n$ (for set $\boldsymbol{S}_k(n)$). As a consequence of Theorem 8 we have

$$\mathrm{TSC}_{k+1}(n) \leq \mathrm{TSC}_k(n) \qquad \text{for } k = 0, \cdots, N-1; n \geq 0 \tag{6.34}$$

By recursive application of (6.34) we have

$$\mathrm{TSC}(n+1) \leq \mathrm{TSC}(n) \tag{6.35}$$

Note that for $N \leq G$, TSC is lower bounded by $N$, and in general (including the case $N > G$) TSC is lower bounded by $N^2/G$ from Welch's bound. Therefore the monotonically decreasing sequence $\mathrm{TSC}(n)$ should converge to a finite number. For $N \leq G$, if $\mathrm{TSC}(n)$ converges to $N$, then $\boldsymbol{S}(n)$ converges to a set of $N$ orthogonal sequences, and for $N > G$ if $\mathrm{TSC}(n)$ converges to $N^2/G$ then $\boldsymbol{S}(n)$ converges to a set of WBE sequences. Note that

$$\mathrm{TSC}(n+1) = \mathrm{TSC}_N(n+1) \leq \mathrm{TSC}_{N-1}(n+1) \leq \cdots \leq \mathrm{TSC}_1(n+1) \leq \mathrm{TSC}(n) \tag{6.36}$$

Thus, at the fixed point where $\mathrm{TSC}(n+1) = \mathrm{TSC}(n)$ we should have

$$\mathrm{TSC}(n+1) = \mathrm{TSC}_N(n+1) = \mathrm{TSC}_{N-1}(n+1) = \cdots = \mathrm{TSC}_1(n+1) = \mathrm{TSC}(n) \quad (6.37)$$

From Theorem 8, this occurs iff $\boldsymbol{s}_1(n+1) = \boldsymbol{s}_1(n), \boldsymbol{s}_2(n+1) = \boldsymbol{s}_2(n), \cdots, \boldsymbol{s}_N(n+1) = \boldsymbol{s}_N(n)$. Thus at the fixed point we must have $\boldsymbol{S}(n+1) = \boldsymbol{S}(n)$.

Let us denote the fixed point set of vectors as $\boldsymbol{S} = [\boldsymbol{s}_1, \cdots, \boldsymbol{s}_N]$. Let $\boldsymbol{B} = \boldsymbol{S}\boldsymbol{S}^\top + a^2\boldsymbol{I}_G$. From two alternative notations for the MMSE operation in (6.14) and (6.16), we will choose to use the one given in (6.14). The fixed point $\boldsymbol{S}$ satisfies

$$\boldsymbol{B}\boldsymbol{s}_i = \phi_i \boldsymbol{s}_i \qquad \text{for } i = 1, \cdots, N \quad (6.38)$$

In the following two subsections, we will investigate the properties of the fixed point for the cases $N \leq G$ and $N > G$ separately. We will identify the conditions under which we can guarantee that the fixed point is actually a set of orthogonal vectors for $N \leq G$ and a set of WBE sequences for $N > G$. We will need the following Lemma.

**Lemma 12** *Ranks of $\boldsymbol{S}_{k+1}(n)$ and $\boldsymbol{S}_k(n)$ are the same for all $k$ and $n$.*

A proof of this Lemma is provided in the Appendix. Note that this Lemma is valid whether $N \leq G$ or $N > G$. The important implication of this Lemma is that the rank of the signature sequence set at any iteration will be the same as the rank of the initial signature sequence set with which the algorithm is started. This means that for $N \leq G$, if the algorithm is started with a set of $N$ linearly independent signature sequences, the set of sequences at each iteration will be linearly independent, and the matrix $\boldsymbol{S}^\top(n)\boldsymbol{S}(n)$ will be invertible for all $n$, and at the fixed point $\boldsymbol{S}^\top\boldsymbol{S}$ will be invertible. Similarly, for $N > G$, if $G$ of $N$ initial signature sequences are chosen such that they are linearly independent, the rank of the signature sequence set will be equal to $G$ at every iteration. Thus, $\boldsymbol{S}(n)\boldsymbol{S}^\top(n)$ will be invertible for any $n$, and at the fixed point $\boldsymbol{S}\boldsymbol{S}^\top$ will be invertible.

### 6.5.1 Properties of the Fixed Point: $N \leq G$

In this section we will consider the case $N \leq G$. Multiplying the equality in (6.38) for user $i$ with $\boldsymbol{s}_j^\top$ from left we obtain

$$\boldsymbol{s}_j^\top \boldsymbol{B} \boldsymbol{s}_i = \phi_i \boldsymbol{s}_j^\top \boldsymbol{s}_i \qquad \text{for } i = 1, \cdots, N;\ j = 1, \cdots, N \tag{6.39}$$

Inserting $\boldsymbol{B} = \boldsymbol{S}\boldsymbol{S}^\top + a^2 \boldsymbol{I}_G$ into (6.39) and writing it in a matrix form we get

$$\boldsymbol{S}^\top \left( \boldsymbol{S}\boldsymbol{S}^\top + a^2 \boldsymbol{I}_G \right) \boldsymbol{S} = \boldsymbol{S}^\top \boldsymbol{S} \boldsymbol{\Phi} \tag{6.40}$$

where $\boldsymbol{\Phi}$ is an $N \times N$ diagonal matrix with $\phi_i$ as its $i$th diagonal element. Equation (6.40) can be written as

$$\left( \boldsymbol{S}^\top \boldsymbol{S} \right)^2 + a^2 \boldsymbol{S}^\top \boldsymbol{S} = \boldsymbol{S}^\top \boldsymbol{S} \boldsymbol{\Phi} \tag{6.41}$$

If the algorithm is started with $N$ linearly independent vectors $\boldsymbol{S}(0)$, then by Lemma 12 the set of vectors at the fixed point $\boldsymbol{S}$ will be linearly independent and $\boldsymbol{S}^\top \boldsymbol{S}$ will be invertible. Multiplying both sides of (6.41) with $(\boldsymbol{S}^\top \boldsymbol{S})^{-1}$ yields

$$\boldsymbol{S}^\top \boldsymbol{S} + a^2 \boldsymbol{I}_N = \boldsymbol{\Phi} \tag{6.42}$$

Thus, $\boldsymbol{S}^\top \boldsymbol{S}$ is a diagonal matrix. By the nature of the algorithm $\boldsymbol{s}_i^\top \boldsymbol{s}_i = 1$ for all $i$, and the diagonal elements of $\boldsymbol{S}^\top \boldsymbol{S}$ are all unity. Therefore, at the fixed point for $N \leq G$ we have $\boldsymbol{S}^\top \boldsymbol{S} = \boldsymbol{I}_N$. Note from (6.42) that at the fixed point all $\phi_i$ have the same value.

Therefore, for $N \leq G$, if the algorithm is started with $N$ linearly independent signature sequences then the proposed algorithm converges to a set of signature sequences $\boldsymbol{S}$ such that $\boldsymbol{S}^\top \boldsymbol{S} = \boldsymbol{I}_N$.

### 6.5.2 Properties of the Fixed point: $N > G$

Now we will consider the case $N > G$. Multiplying (6.38) for user $i$ with $\boldsymbol{s}_j^\top$ from left, and multiplying the same equation for user $j$ with $\boldsymbol{s}_i^\top$ from left, we obtain the following two equations

$$\boldsymbol{s}_j^\top \boldsymbol{B} \boldsymbol{s}_i = \phi_i \boldsymbol{s}_j^\top \boldsymbol{s}_i \tag{6.43}$$

$$\boldsymbol{s}_i^\top \boldsymbol{B} \boldsymbol{s}_j = \phi_j \boldsymbol{s}_i^\top \boldsymbol{s}_j \tag{6.44}$$

Since $\boldsymbol{B}$ is symmetric, we must have

$$\phi_i \boldsymbol{s}_j^\top \boldsymbol{s}_i = \phi_j \boldsymbol{s}_i^\top \boldsymbol{s}_j \tag{6.45}$$

Thus, if $\boldsymbol{s}_j^\top \boldsymbol{s}_i \neq 0$, then $\phi_i = \phi_j$. If the signature sequence set at the fixed point is not split into two or more orthogonal subsets, then we are guaranteed to have $\phi_i = \phi$ for all $i$. Note that some vectors may be orthogonal to each other but as long as we do not have a splitting of the signature set into two or more orthogonal subspaces, we can still conclude that all $\phi_i$ have the same value. As we will see below we can guarantee that the signature sequence set does not get split into orthogonal subsets at the fixed point by imposing quite mild conditions on the initial set of signature sequences, $\boldsymbol{S}(0)$. Without going into the details of these conditions, let us assume that they are satisfied and $\phi_i = \phi$ for all $i$. Thus we have

$$\boldsymbol{s}_j^\top \boldsymbol{B} \boldsymbol{s}_i = \phi \boldsymbol{s}_j^\top \boldsymbol{s}_i \qquad \text{for } i = 1, \cdots, N;\ j = 1, \cdots, N \tag{6.46}$$

Similar to the case of $N \leq G$, inserting $\boldsymbol{B} = \boldsymbol{S}\boldsymbol{S}^\top + a^2 \boldsymbol{I}_G$ into (6.46) and writing it in a matrix form we get

$$\boldsymbol{S}^\top \left( \boldsymbol{S}\boldsymbol{S}^\top + a^2 \boldsymbol{I}_G \right) \boldsymbol{S} = \phi \boldsymbol{S}^\top \boldsymbol{S} \tag{6.47}$$

Multiplying both sides of (6.47) from left by $\boldsymbol{S}$ and from right by $\boldsymbol{S}^\top$ yields

$$\boldsymbol{S}\boldsymbol{S}^\top \left( \boldsymbol{S}\boldsymbol{S}^\top + a^2 \boldsymbol{I}_G \right) \boldsymbol{S}\boldsymbol{S}^\top = \phi \boldsymbol{S}\boldsymbol{S}^\top \boldsymbol{S}\boldsymbol{S}^\top \tag{6.48}$$

Rewriting (6.48) we obtain

$$\left( \boldsymbol{S}\boldsymbol{S}^\top \right)^3 + a^2 \left( \boldsymbol{S}\boldsymbol{S}^\top \right)^2 = \phi \left( \boldsymbol{S}\boldsymbol{S}^\top \right)^2 \tag{6.49}$$

Note that $\boldsymbol{S}\boldsymbol{S}^\top$ is invertible if the rank of $\boldsymbol{S}$ is $G$. Note from Lemma 12 that the rank of the signature sequence set remains the same as the number of iterations increases. Thus, if $G$ of the $N$ signature sequences in $\boldsymbol{S}(0)$ are chosen linearly independent $\boldsymbol{S}\boldsymbol{S}^\top$ will be guaranteed to be invertible. Multiplying both sides of (6.49) by $(\boldsymbol{S}\boldsymbol{S}^\top)^{-2}$ yields

$$\boldsymbol{S}\boldsymbol{S}^\top + a^2 \boldsymbol{I}_G = \phi \boldsymbol{I}_G \tag{6.50}$$

Thus, $\boldsymbol{S}\boldsymbol{S}^\top = \alpha \boldsymbol{I}_G$. Note that $\text{trace}(\boldsymbol{S}\boldsymbol{S}^\top) = \text{trace}(\boldsymbol{S}^\top \boldsymbol{S}) = N$, since $\boldsymbol{s}_i^\top \boldsymbol{s}_i = 1$ for all $i$ by the nature of the algorithm. Since $\text{trace}(\alpha \boldsymbol{I}_G) = \alpha G$, we have $\alpha = N/G$.

Therefore, for $N > G$, if the algorithm is started with $N$ signature sequences, $G$ of which are linearly independent, and if the initial set of vectors are chosen such that the signal space cannot be partitioned into two or more orthogonal subspaces then the proposed algorithm converges to a set of signature sequences $\boldsymbol{S}$ such that $\boldsymbol{S}\boldsymbol{S}^\top = (N/G)\boldsymbol{I}_G$. The following Lemma verifies that for $N > G$, a signature set which is not split into orthogonal subsets cannot be split into orthogonal subsets in one iteration. Recursive application of this Lemma implies that if the initial signature set is not split into orthogonal subsets then it will never be split into orthogonal subsets.

**Lemma 13** *For $N > G$, $\boldsymbol{S}(n)$ can be partitioned into two or more orthogonal subspaces iff $\boldsymbol{S}(n-1)$ can be partitioned into two or more orthogonal subspaces.*

A proof of this Lemma is provided in the Appendix in Section 6.A.

## 6.6 Optimum Orthogonal Partitioning

As we observed, for $N > G$, if we start the algorithm with two orthogonal subsets then the algorithm does not converge to a WBE sequence set. Although TSC in this case is still a monotonically decreasing function of the number of iterations, since the algorithm is not guaranteed to converge to a WBE set, the TSC does not decrease to $N^2/G$. Thus, we concluded that in general orthogonal partitioning is suboptimal.

In this Section, we will investigate the conditions under which orthogonal splitting can actually be as good as non-splitting. Note that the processing gain is proportional to the available bandwidth. For a fixed bandwidth $W$, we can either have a signature sequence set with processing gain $G$, or we can have two orthogonal signature sequence sets with processing gains $G_1$ and $G_2 = G - G_1$ corresponding to bandwidth partitions $W_1$ and $W_2 = W - W_1$. Let us denote the number of users in the first band having a processing gain $G_1$ with $N_1$. Thus the number of users in the second band with a processing gain $G_2 = G - G_1$ is $N_2 = N - N_1$. From our observations above the best signature sequence assignment is attained when all of the bandwidth is used and either $N$ orthogonal vectors (if $N \leq G$), or a set of $N$ WBE sequences (if $N > G$) are assigned to the users. Below we will investigate the effects of orthogonal partitioning for the cases $N \leq G$ and $N > G$ separately. In both cases we will denote the TSC achieved by not partitioning the available bandwidth by $\text{TSC}_1$ and the TSC achieved by bandwidth partitioning by $\text{TSC}_2$.

If $N \leq G$, then we have to consider two cases: (i) $N_1 \leq G_1$ and $N_2 \leq G_2$, and (ii) $N_1 \leq G_1$ and $N_2 > G_2$. We will not list the third possibility where $N_1 > G_1$ and $N_2 \leq G_2$, because its result will be similar to that of (ii). In case (i), we can assign $N_1$ orthogonal signature sequences to the users in the first band and $N_2$ orthogonal signature sequences to the users in the second band. In this case, $\text{TSC}_1 = N$, and $\text{TSC}_2 = N_1 + (N - N_1) = N$. Thus, $\text{TSC}_1 = \text{TSC}_2$ and both assignments are equally good. In case (ii), we can assign $N_1$ orthogonal signature sequences to the users in the first band and $N_2$ WBE sequences for the users in the second band. Then,

$$\text{TSC}_2 = N_1 + \frac{(N - N_1)^2}{(G - G_1)} = N_1 + (N - N_1)\frac{(N - N_1)}{(G - G_1)} \tag{6.51}$$

Since $(N - N_1)/(G - G_1) > 1$, $\text{TSC}_2 > \text{TSC}_1$, and orthogonal partitioning is strictly suboptimal.

If $N > G$, then we have to consider two cases: (i) $N_1 > G_1$ and $N_2 > G_2$, and (ii) $N_1 > G_1$ and $N_2 \leq G_2$. In case (i) we can assign a set of $N_1$ WBE sequences to the users in the first band and $N_2$ WBE sequences to the users in the second band. In this case, $\text{TSC}_1 = N^2/G$ and

$$\text{TSC}_2 = \frac{N_1^2}{G_1} + \frac{(N - N_1)^2}{(G - G_1)} \tag{6.52}$$

It is not difficult to show that $\text{TSC}_1 \leq \text{TSC}_2$ in general, and that the inequality is satisfied with equality iff

$$\frac{N_1}{G_1} = \frac{N}{G} \tag{6.53}$$

In case (ii), we can assign a set of $N_1$ WBE sequences to the users in the first band and $N_2$ orthogonal sequences to the users in the second band. In this case,

$$\text{TSC}_2 = \frac{N_1^2}{G_1} + (N - N_1) > \frac{N_1^2}{G_1} + (N - N_1)\frac{(N - N_1)}{(G - G_1)} \geq \text{TSC}_1 \tag{6.54}$$

The first inequality is obtained from $(N - N_1)/(G - G_1) < 1$, and the second one is obtained by using the result of case (i). Thus in this case $\text{TSC}_2 > \text{TSC}_1$ and orthogonal partitioning is strictly suboptimal.

In the above derivation we observed that only in two cases orthogonal partitioning of the signature sequence set yields equally good solutions as not partitioning: if $N \leq G$, any partitioning is as good as the optimal one as long as both spaces have enough dimensionality to assign orthogonal signature sequences to the users; and if $N > G$, we must have the "equal loading" condition of (6.53) to have a solution that is as good as the optimal one.

The advantage of orthogonal partitioning is that the update of a signature sequence is not affected (does not depend) on the signature sequences in the orthogonal subset. Therefore less (local) information is needed for the update of each signature sequence. The results in this section may be useful in the design of signature sequences for a multicell system with a finite number of cells.

## 6.7 Simulation Results

In this section, we present some simple simulation results to verify our analysis. We take the processing gain to be $G = 10$. The initial signature sequences are created randomly. In all figures, $N$ updates take place between iterations $n$ and $(n+1)$. In each update the $i$th user's signature sequence is replaced with the corresponding normalized MMSE filter, for $i = 1, \cdots, N$. Figures 6.1 and 6.2 show the minimum and maximum eigenvalues of the matrix $\boldsymbol{S}^\top(n)\boldsymbol{S}(n)$, and TSC($n$), respectively, as a function of the iteration index for number of users $N = 5, 10$. As expected, the minimum and maximum eigenvalues of $\boldsymbol{S}^\top(n)\boldsymbol{S}(n)$ converge to 1 implying that the matrix converges to $\boldsymbol{S}^\top\boldsymbol{S} = \boldsymbol{I}_N$, and the TSC converges to $N$ since $N \leq G$.

Figure 6.3 and 6.4 show the minimum and maximum eigenvalues of the matrix $\boldsymbol{S}(n)\boldsymbol{S}^\top(n)$ and TSC($n$), respectively, for number of users $N = 20, 30, 40, 50$. As expected, the minimum and maximum eigenvalues of the matrix $\boldsymbol{S}(n)\boldsymbol{S}^\top(n)$ converge to $N/G$ implying that the matrix converges to $\boldsymbol{S}\boldsymbol{S}^\top = (N/G)\boldsymbol{I}_G$, and the TSC converges to $N^2/G$, since in these cases $N > G$.

In the following experiment we observe the effects of orthogonal partitioning investigated analytically in Section 6.6. The number of users is $N = 20$ and the processing

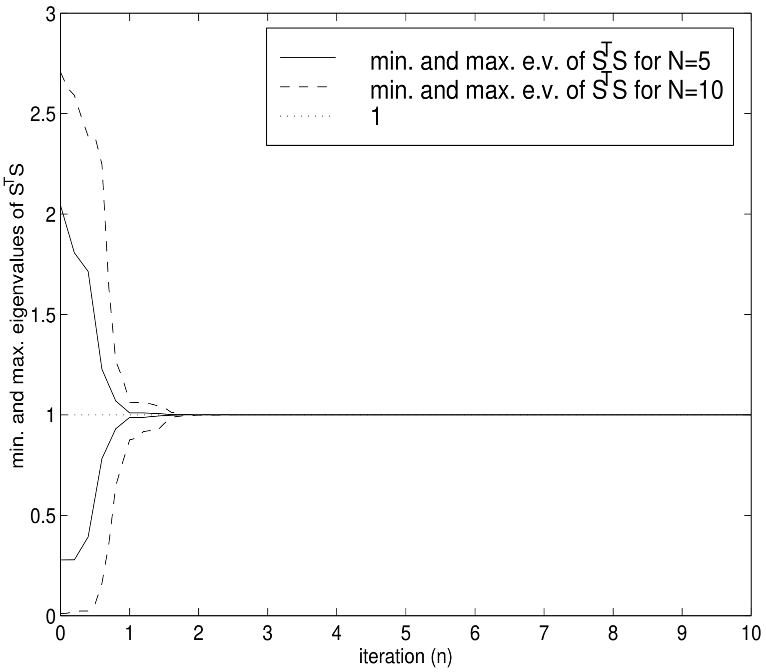


Figure 6.1: Minimum and maximum eigenvalues of matrix $\boldsymbol{S}^\top(n)\boldsymbol{S}(n)$.

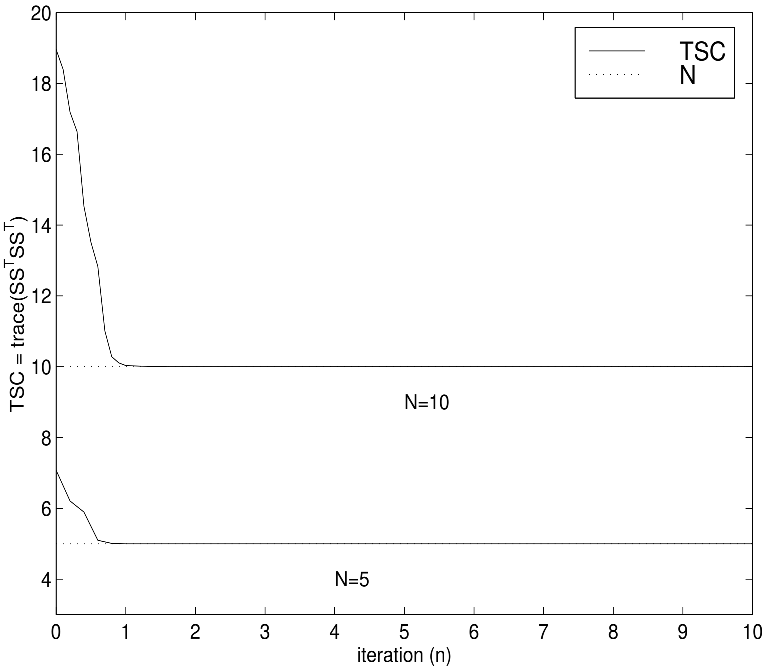


Figure 6.2: Total squared correlation, $\mathrm{TSC}(n)$.

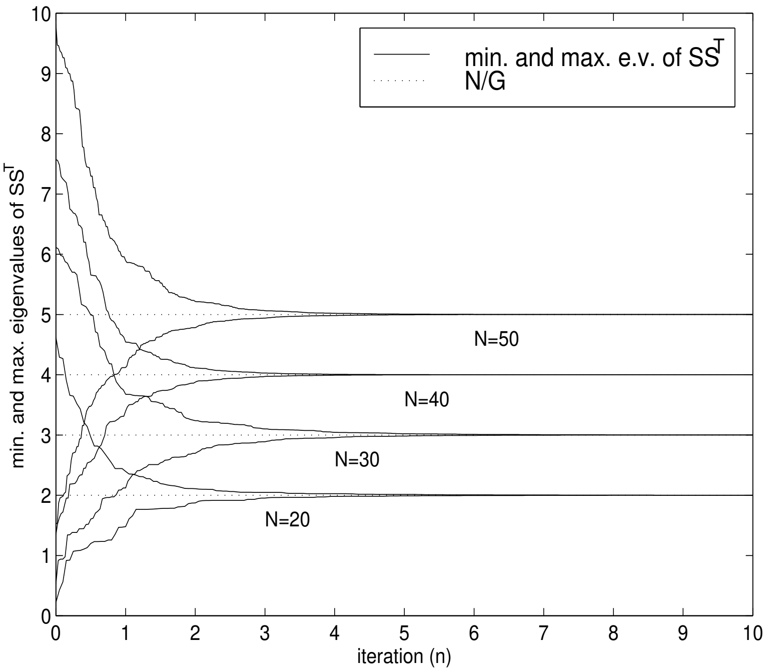


Figure 6.3: Minimum and maximum eigenvalues of matrix $\boldsymbol{S}(n)\boldsymbol{S}^{\top}(n)$.

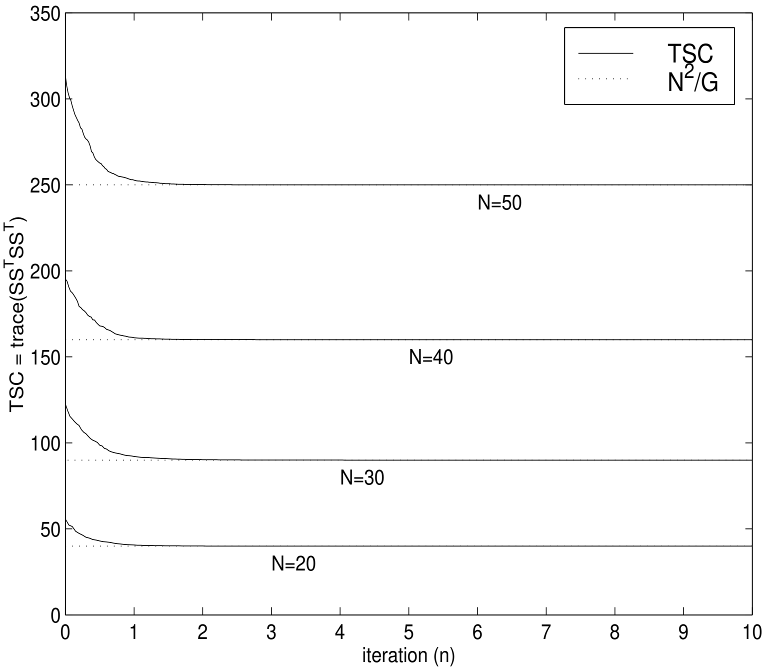


Figure 6.4: Total squared correlation, TSC$(n)$.

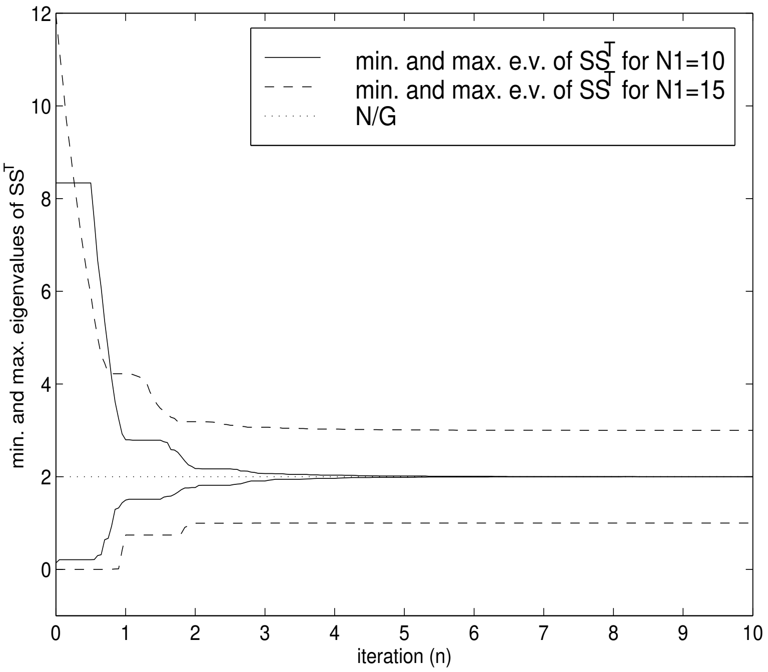


Figure 6.5: Minimum and maximum eigenvalues of matrix $\boldsymbol{S}(n)\boldsymbol{S}^{\top}(n)$.

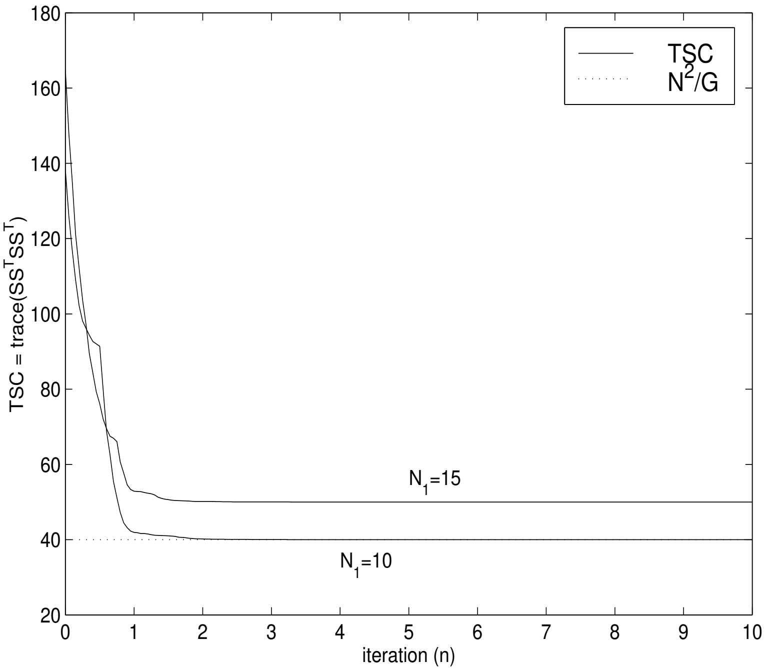


Figure 6.6: Total squared correlation, $\mathrm{TSC}(n)$.

gain is $G = 10$. Figures 6.5 and 6.6 show the minimum and maximum eigenvalues of the matrix $\boldsymbol{S}(n)\boldsymbol{S}^\top(n)$ and TSC$(n)$, respectively. Two kinds of orthogonal partitioning are examined. In the first one, the users are partitioned into two orthogonal subspaces according to $N_1 = N_2 = 10$ and $G_1 = G_2 = 5$. Since this an "equal loading" case $(N_1/N = G_1/G = 1/2)$, the whole signature sequence set converges to a WBE sequence set: the matrix converges to $\boldsymbol{S}\boldsymbol{S}^\top = (N/G)\boldsymbol{I}_G = 2\boldsymbol{I}_G$ and the TSC converges to $N^2/G = 40$; see the solid curves in Figure 6.5 and the curve corresponding to $N_1 = 10$ in Figure 6.6. In the second orthogonal partitioning, the users are partitioned into two orthogonal subspaces with $N_1 = 15$ and $N_2 = 5$ and again $G_1 = G_2 = 5$. Note that this partitioning does not yield equal loads to both subsets and the matrix $\boldsymbol{S}(n)\boldsymbol{S}^\top(n)$ does not converge to a multiple of the identity matrix (see dashed lines in Figure 6.5), and the TSC$(n)$ does not decrease down to $N^2/G = 40$, but converges to 50. Note that the signature sequences of $N_1 = 15$ users in the first subspace converge to a WBE sequence set with $N_1 = 15$ and $G_1 = 5$, achieving a TSC of $N_1^2/G_1 = 15^2/5 = 45$, and the signature sequences of $N_2 = 5$ users converge to $N_2$ orthogonal sequences yielding a TSC of $N_2 = 5$. Thus the combined TSC is 45+5=50 which can be seen in Figure 6.6.

## 6.8 Chapter Summary and Conclusions

Recently it was shown that in order to maximize the capacity of a single cell synchronous CDMA system, if $N \leq G$ then $N$ orthonormal sequences, and if $N > G$ then WBE sequences should be used as signature sequences. In this chapter we proposed an algorithm which iteratively updates the signature sequences of the users one at a time and converges to the optimum signature sequence sets.

The algorithm replaces one user's signature sequence with a generalized normalized MMSE filter of that user at each update. Since the MMSE receiver can be constructed in a distributed fashion, the proposed algorithm can be implemented distributedly. In this case each user updates its own signature sequence by making some local measurements and the whole signature sequence set converges to the optimum one.

## 6.A Appendix: Additional Proofs

**Proof: (Theorem 8)** In order to prove TSC$'$ $\leq$ TSC we have to show that

$$\frac{\boldsymbol{s}_k^\top \boldsymbol{A}_k^{-1} \boldsymbol{s}_k}{\boldsymbol{s}_k^\top \boldsymbol{A}_k^{-2} \boldsymbol{s}_k} \leq \boldsymbol{s}_k^\top \boldsymbol{A}_k \boldsymbol{s}_k \tag{6.55}$$

In order to prove (6.55) we will prove the following two inequalities

$$\boldsymbol{s}_k^\top \boldsymbol{A}_k^{-1} \boldsymbol{s}_k \leq \left(\boldsymbol{s}_k^\top \boldsymbol{A}_k^{-2} \boldsymbol{s}_k\right)^{1/2} \tag{6.56}$$

$$1 = \boldsymbol{s}_k^\top \boldsymbol{s}_k \leq \left(\boldsymbol{s}_k^\top \boldsymbol{A}_k \boldsymbol{s}_k\right)^{1/2} \left(\boldsymbol{s}_k^\top \boldsymbol{A}_k^{-1} \boldsymbol{s}_k\right)^{1/2} \tag{6.57}$$

Note that combining (6.56) and (6.57) will yield the necessary result in (6.55). Also note that the inequality in (6.55) will be satisfied with equality iff the inequalities in (6.56) and (6.57) are satisfied with equality. Since $\boldsymbol{A}_k$ is symmetric and positive definite we can write it as $\boldsymbol{A}_k = \boldsymbol{Q}\boldsymbol{\Lambda}\boldsymbol{Q}^\top$ where $\boldsymbol{\Lambda}$ is a diagonal matrix which consists of the eigenvalues of matrix $\boldsymbol{A}_k$ as its diagonal elements and $\boldsymbol{Q}$ consists of the orthonormal eigenvectors of the same matrix as its columns. Since $\boldsymbol{A}_k^{-1} = \boldsymbol{Q}\boldsymbol{\Lambda}^{-1}\boldsymbol{Q}^\top$ and $\boldsymbol{A}_k^{-2} = \boldsymbol{Q}\boldsymbol{\Lambda}^{-2}\boldsymbol{Q}^\top$, inequality in (6.56) can be written as

$$\boldsymbol{s}_k^\top \boldsymbol{Q}\boldsymbol{\Lambda}^{-1}\boldsymbol{Q}^\top \boldsymbol{s}_k \leq \left(\boldsymbol{s}_k^\top \boldsymbol{Q}\boldsymbol{\Lambda}^{-2}\boldsymbol{Q}^\top \boldsymbol{s}_k\right)^{1/2} \tag{6.58}$$

Let us define $\boldsymbol{x} = \boldsymbol{Q}^\top \boldsymbol{s}_k$. Since $\boldsymbol{Q}^\top\boldsymbol{Q} = \boldsymbol{I}_G$, $\boldsymbol{x}^\top\boldsymbol{x} = \boldsymbol{s}_k^\top \boldsymbol{s}_k = 1$. Inequality (6.58) can be written as

$$\sum_i \frac{1}{\lambda_i} x_i^2 \leq \left(\sum_i \frac{1}{\lambda_i^2} x_i^2\right)^{1/2} \tag{6.59}$$

where $\lambda_i$ is the $i$th eigenvalue of $\boldsymbol{A}_k$ (and the $(i,i)$th element of $\boldsymbol{\Lambda}$). Let us define two vectors $\boldsymbol{u}$ and $\boldsymbol{v}$, with elements $u_i = \frac{1}{\lambda_i} x_i$ and $v_i = x_i$. From the Cauchy-Schwarz inequality

$$\boldsymbol{u}^\top \boldsymbol{v} = \sum_i \frac{1}{\lambda_i} x_i^2 \leq \|\boldsymbol{u}\| \, \|\boldsymbol{v}\| = \left(\sum_i \frac{1}{\lambda_i^2} x_i^2\right)^{1/2} \left(\sum_i x_i^2\right)^{1/2} \tag{6.60}$$

Then (6.59) follows from (6.60) since $\boldsymbol{x}^\top\boldsymbol{x} = 1$. Note that this inequality is satisfied with equality iff $\boldsymbol{u} = \alpha\boldsymbol{v}$, i.e., $\boldsymbol{\Lambda}^{-1}\boldsymbol{x} = \alpha\boldsymbol{x}$. Inserting $\boldsymbol{x} = \boldsymbol{Q}^\top \boldsymbol{s}_k$, this condition can equivalently be stated as $\boldsymbol{A}_k^{-1}\boldsymbol{s}_k = \alpha\boldsymbol{s}_k$. Thus, the inequality in (6.56) is satisfied with equality if the normalized MMSE operation on $\boldsymbol{s}_k$ yields $\boldsymbol{s}_k$.

Similarly, inequality (6.57) can be written as

$$1 = \sum_i x_i^2 \leq \left(\sum_i \lambda_i x_i^2\right)^{1/2} \left(\sum_i \frac{1}{\lambda_i} x_i^2\right)^{1/2} \tag{6.61}$$

Now let us define vectors $\bar{\boldsymbol{u}}$ and $\bar{\boldsymbol{v}}$ as $\bar{u}_i = \sqrt{\lambda_i} x_i$ and $\bar{v}_i = \frac{1}{\sqrt{\lambda_i}} x_i$. Again from the Cauchy-Schwarz inequality

$$\bar{\boldsymbol{u}}^\top \bar{\boldsymbol{v}} = \sum_i x_i^2 \leq \|\bar{\boldsymbol{u}}\| \, \|\bar{\boldsymbol{v}}\| = \left(\sum_i \lambda_i x_i^2\right)^{1/2} \left(\sum_i \frac{1}{\lambda_i} x_i^2\right)^{1/2} \tag{6.62}$$

Note again that the inequality is satisfied with equality iff $\bar{\boldsymbol{u}} = \beta \bar{\boldsymbol{v}}$, i.e., $\boldsymbol{\Lambda}^{1/2}\boldsymbol{x} = \beta \boldsymbol{\Lambda}^{-1/2}\boldsymbol{x}$, equivalently $\boldsymbol{\Lambda}^{-1}\boldsymbol{x} = \beta \boldsymbol{x}$ as in the previous case. Thus, as in the previous inequality the equality is achieved if the normalized MMSE operation on $\boldsymbol{s}_k$ yields $\boldsymbol{s}_k$. □

**Proof: (Lemma 12)** We will prove that after one normalized MMSE operation the the rank of a set of vectors cannot decrease. As in the proof of Theorem 8 we denote the previous set of vectors as $\boldsymbol{S} = [\boldsymbol{s}_1, \cdots, \boldsymbol{s}_k, \cdots, \boldsymbol{s}_N]$ and the set after the update as $\boldsymbol{S}' = [\boldsymbol{s}_1, \cdots, \boldsymbol{c}_k, \cdots, \boldsymbol{s}_N]$, i.e., the $k$th user's signature sequence is updated. We will again disregard the normalization step in the update. We will assume the MMSE update given in (6.16)

$$\boldsymbol{c}_k = \boldsymbol{A}_k^{-1} \boldsymbol{s}_k \tag{6.63}$$

where $\boldsymbol{A}_k = \sum_{j \neq k} \boldsymbol{s}_j \boldsymbol{s}_j^\top + a^2 \boldsymbol{I}_G$. Thus

$$\left(\sum_{j \neq k} \boldsymbol{s}_j \boldsymbol{s}_j^\top + a^2 \boldsymbol{I}_G\right) \boldsymbol{c}_k = \boldsymbol{s}_k \tag{6.64}$$

and finally

$$\boldsymbol{c}_k = \frac{1}{a^2} \boldsymbol{s}_k - \frac{1}{a^2} \sum_{j \neq k} (\boldsymbol{s}_j^\top \boldsymbol{c}_k) \boldsymbol{s}_j \tag{6.65}$$

Therefore, $\boldsymbol{c}_k$ is a linear combination of all of the signature sequences. This means that the matrix $\boldsymbol{S}'$ is actually obtained from matrix $\boldsymbol{S}$ by elementary column operations which does not change the rank of the matrix. Thus, $\mathrm{rank}(\boldsymbol{S}') = \mathrm{rank}(\boldsymbol{S})$. □

**Proof: (Lemma 13)** We will show that for $N > G$ if at the end of one full iteration (where all of the vectors in the current set are updated) we have two orthogonal subsets then we must have two orthogonal subsets in the previous iteration. Let us denote the signature sequence set at the end of iteration $n$ with $\boldsymbol{S}(n)$. Let us assume that $\boldsymbol{S}(n)$ has two orthogonal subsets $\tilde{\boldsymbol{S}}(n)$ and $\bar{\boldsymbol{S}}(n)$. The signature sets corresponding to sets $\tilde{\boldsymbol{S}}(n)$ and $\bar{\boldsymbol{S}}(n)$ in the previous iteration are $\tilde{\boldsymbol{S}}(n-1)$ and $\bar{\boldsymbol{S}}(n-1)$. Let us assume that the sets $\tilde{\boldsymbol{S}}(n)$ and $\tilde{\boldsymbol{S}}(n-1)$ have $N_1$ signature sequences and sets $\bar{\boldsymbol{S}}(n)$ and $\bar{\boldsymbol{S}}(n-1)$ have $N_2$ signature sequences, where $N_1 + N_2 = N$. Let us also denote the ranks of the matrices as $\text{rank}(\tilde{\boldsymbol{S}}(n-1)) = \tilde{G}_{n,0}$ and $\text{rank}(\bar{\boldsymbol{S}}(n-1)) = \bar{G}_{n,0}$; $\text{rank}(\tilde{\boldsymbol{S}}(n)) = \tilde{G}_{n,N}$ and $\text{rank}(\bar{\boldsymbol{S}}(n)) = \bar{G}_{n,N}$.

Note that for any two sets $U$ and $V$, $\text{rank}(U) + \text{rank}(V) \geq \text{rank}(U \cup V)$ in general and the inequality is satisfied with equality iff $U$ and $V$ are linearly independent sets. We call two sets to be linearly independent if no one vector in one set can be written as a linear superposition of the vectors in the other set. In other words, two sets are linearly independent if they span nonintersecting subspaces except the origin. Note that any two orthogonal sets are also linearly independent. Thus, $\tilde{\boldsymbol{S}}(n)$ and $\bar{\boldsymbol{S}}(n)$ are linearly independent and we have $\tilde{G}_{n,N} + \bar{G}_{n,N} = G$. This follows from the assumption that the initial signature sequence set $\boldsymbol{S}(0)$ was chosen such that it had rank $G$ and from Lemma 12, rank of the set never decreases and the rank of the whole signature sequence set is $G$ (for the case of $N < G$, the rank of the set is always $N$) at any iteration.

Since $N > G$ we must have either $N_1 > \tilde{G}_{n,N}$ or $N_2 > \bar{G}_{n,N}$ or both. We will assume without loss of generality that $N_1 > \tilde{G}_{n,N}$. Again without loss of generality we assume that first the vectors in set $\tilde{\boldsymbol{S}}(n-1)$ are updated one by one, then the vectors in set $\bar{\boldsymbol{S}}(n-1)$ are updated again one at a time. Thus the set $\tilde{\boldsymbol{S}}(n-1)$ includes users indexed from 1 to $N_1$, and the set $\bar{\boldsymbol{S}}(n-1)$ includes the users indexed from $(N_1+1)$ to $N$. We will consider the step where the vector, $\boldsymbol{s}_i(n-1)$, $1 \leq i \leq N_1$, is updated to yield $\boldsymbol{s}_i(n)$. With the convention introduced in Section 6.4, new sets after this one update are denoted as $\tilde{\boldsymbol{S}}_i(n)$ and $\bar{\boldsymbol{S}}_i(n)$ and the ranks of the sets as $\tilde{G}_{n,i}$ and $\bar{G}_{n,i}$. As a simple consequence of Lemma 12 we note that the rank of any subset in $\boldsymbol{S}(n-1)$ cannot decrease after one update, i.e., $\tilde{G}_{n,i} \geq \tilde{G}_{n,k}$ for $i \geq k$. Although the rank of the whole signature sequence set does not change after one update, the rank of a subset of a signature sequences has a chance of increasing because a contribution to the rank of the subset from the complementary set is possible.

Since the sets $\tilde{\boldsymbol{S}}(n)$ and $\bar{\boldsymbol{S}}(n)$ are linearly independent, $\tilde{G}_{n,N} + \bar{G}_{n,N} = G$, and since the rank of a subset cannot decrease $\tilde{G}_{n,N} \geq \tilde{G}_{n,N-1}$ and $\bar{G}_{n,N} \geq \bar{G}_{n,N-1}$. Note that $\tilde{G}_{n,N-1} + \bar{G}_{n,N-1} \geq G$, from the fact that for any two arbitrary sets, the sum of the ranks of the sets is less than the rank of the union of the sets. Combining all these we get, $\tilde{G}_{n,N-1} + \bar{G}_{n,N-1} = G$; thus the sets in the previous step are linearly independent too. Also we have $\tilde{G}_{n,N} = \tilde{G}_{n,N-1}$ and $\bar{G}_{n,N} = \bar{G}_{n,N-1}$. Applying this recursively we note that $\tilde{G}_{n,m} + \bar{G}_{n,m} = G$, therefore the sets $\tilde{\boldsymbol{S}}_m(n)$ and $\bar{\boldsymbol{S}}_m(n)$ for $m = 0, \cdots, N$ and $n \geq 0$ are linearly independent. Also note that $\tilde{G}_{n,m} = \tilde{G}_{n,N}$ and $\bar{G}_{n,m} = \bar{G}_{n,N}$ for $m = 0, \cdots, N$.

Let us now concentrate on the update of signature sequence $\boldsymbol{s}_i(n-1)$ from set $\tilde{\boldsymbol{S}}(n-1)$. Ignoring the normalization step of the update, $\boldsymbol{s}_i(n)$ is given by $\boldsymbol{s}_i(n) = \boldsymbol{A}_i^{-1}(n)\boldsymbol{s}_i(n-1)$ and hence $\boldsymbol{A}_i(n)\boldsymbol{s}_i(n) = \boldsymbol{s}_i(n-1)$ where

$$\begin{aligned}\boldsymbol{A}_i(n) &= \sum_{k<i} \boldsymbol{s}_k(n)\boldsymbol{s}_k^\top(n) + \sum_{i<k\leq N_1} \boldsymbol{s}_k(n-1)\boldsymbol{s}_k^\top(n-1) \\ &\quad + \sum_{k>N_1} \boldsymbol{s}_k(n-1)\boldsymbol{s}_k^\top(n-1) + a^2\boldsymbol{I}_G \end{aligned} \tag{6.66}$$

Thus

$$\begin{aligned}\boldsymbol{s}_i(n) &= \frac{1}{a^2}\boldsymbol{s}_i(n-1) - \frac{1}{a^2}\sum_{k<i}(\boldsymbol{s}_k^\top(n)\boldsymbol{s}_i(n))\boldsymbol{s}_k(n) \\ &\quad -\frac{1}{a^2}\sum_{i<k\leq N_1}(\boldsymbol{s}_k^\top(n-1)\boldsymbol{s}_i(n-1))\boldsymbol{s}_k(n-1) \\ &\quad -\frac{1}{a^2}\sum_{k>N_1}(\boldsymbol{s}_k^\top(n-1)\boldsymbol{s}_i(n))\boldsymbol{s}_k(n-1)\end{aligned} \tag{6.67}$$

Therefore, if $\boldsymbol{s}_i^\top(n)\boldsymbol{s}_k \neq 0$ for some $k > N_1$ [i.e., $\boldsymbol{s}_k(n-1) \in \bar{\boldsymbol{S}}(n-1)$], the rank of $\tilde{\boldsymbol{S}}(n-1)$, $\tilde{G}_{n,0}$ increases, because some vectors in $\bar{\boldsymbol{S}}(n-1)$ contribute to $\boldsymbol{s}_i(n)$. Since $\tilde{\boldsymbol{S}}(n-1)$ and $\bar{\boldsymbol{S}}(n-1)$ are linearly independent, this contribution increases the rank of the first subset. Note that by definition $\tilde{G}_{n,0} = \mathrm{rank}(\tilde{\boldsymbol{S}}(n-1)) \leq \min\{N_1, G\}$ and since by assumption $\tilde{G}_{n,0} < N_1$ and by definition $\tilde{G}_{n,0} < G$, $\tilde{G}_{n,0}$, the rank of the first subset at step 0, can, in principle, increase. But this implies $\tilde{G}_{n,i} > \tilde{G}_{n,i-1}$, and recursively $\tilde{G}_{n,N} > \tilde{G}_{n,0}$. But above we showed that $\tilde{G}_{n,0} = \tilde{G}_{n,m} = \tilde{G}_{n,N}$ for $0 \leq m \leq N$. Thus for the rank of the first subset to remain the same, every new vector generated in the first subset should be orthogonal to all of the vectors in the second subset. Thus after $N_1$ updates the sets $\tilde{\boldsymbol{S}}_{N_1}(n)$ and $\bar{\boldsymbol{S}}_{N_1}(n)$ will be orthogonal.

Now let us consider the last update that yields $\tilde{\boldsymbol{S}}_{N_1}(n)$. Note that before the last vector is updated $\tilde{\boldsymbol{S}}_{N_1-1}(n)$ has $N_1 - 1$ vectors which are orthogonal to every vector in set $\bar{\boldsymbol{S}}(n-1)$. Note that $\bar{\boldsymbol{S}}(n-1)$ did not change because only the vectors in first subset are being updated yet. The last vector to be updated in $\tilde{\boldsymbol{S}}(n-1)$, $\boldsymbol{s}_{N_1}(n-1)$, and the normalized MMSE filter corresponding to that, $\boldsymbol{s}_{N_1}(n)$, are related by (ignoring the normalization step) $\boldsymbol{s}_{N_1}(n) = \boldsymbol{A}_{N_1}^{-1}(n)\boldsymbol{s}_{N_1}(n-1)$, or equivalently as $\boldsymbol{A}_{N_1}(n)\boldsymbol{s}_{N_1}(n) = \boldsymbol{s}_{N_1}(n-1)$ where

$$\boldsymbol{A}_{N_1}(n) = \sum_{k<N_1} \boldsymbol{s}_k(n)\boldsymbol{s}_k^\top(n) + \sum_{k>N_1} \boldsymbol{s}_k(n-1)\boldsymbol{s}_k^\top(n-1) + a^2\boldsymbol{I}_G \tag{6.68}$$

Since $\boldsymbol{s}_{N_1}^\top(n)\boldsymbol{s}_j(n-1) = 0$, for $j > N_1$, from the above discussion, we have $\tilde{\boldsymbol{A}}_{N_1}(n)\boldsymbol{s}_{N_1}(n) = \boldsymbol{s}_{N_1}(n-1)$ where

$$\tilde{\boldsymbol{A}}_{N_1}(n) = \sum_{k<N_1} \boldsymbol{s}_k(n)\boldsymbol{s}_k^\top(n) + a^2\boldsymbol{I}_G \tag{6.69}$$

Thus, for $j > N_1$,

$$\boldsymbol{s}_{N_1}^\top(n-1)\boldsymbol{s}_j(n-1) = \boldsymbol{s}_{N_1}^\top(n)\tilde{\boldsymbol{A}}_{N_1}(n)\boldsymbol{s}_j(n-1) = a^2\boldsymbol{s}_{N_1}^\top(n)\boldsymbol{s}_j(n-1) = 0 \tag{6.70}$$

Thus, the last vector in $\tilde{\boldsymbol{S}}(n-1)$, $\boldsymbol{s}_{N_1}(n-1)$, must be orthogonal to all the vectors in $\bar{\boldsymbol{S}}(n-1)$. Going back recursively, we find that all the vectors in $\tilde{\boldsymbol{S}}(n-1)$ must be orthogonal to all the vectors in $\bar{\boldsymbol{S}}(n-1)$.

Therefore, for $N > G$, if the signature sequence set has two linearly independent subsets (in our case this is a result of two orthogonal subsets), then in the previous iteration we must have two orthogonal subsets. Applying recursively, we conclude that, for the set at any iteration to have two orthogonal subsets, we must have started with two orthogonal subsets. Note that this proof can be generalized to cases where there are more than two orthogonal subsets. The crucial step of the generalization is that we can choose the subset which has more vectors than its rank to work with, and combine all other orthogonal subsets to yield a single subset which is orthogonal to this first subset which has more users than its rank. □

# Chapter 7

# Conclusion

This thesis developed new algorithms and practical versions of existing algorithms to lessen the effects of non-orthogonal signaling in CDMA systems. The optimum detection scheme, which removes the near-far effect, is exponentially complex making it intractable for a reasonable number of transmitters. This thesis has identified tractable cases where the detection problem can be solved in polynomial time. The tractability of the optimum multiuser detection requires a careful design of signature sequences to be used. It is shown in this thesis that if the signature sequences are chosen such that all cross correlations are negative, then the information bits of the users can be detected optimally with a polynomial computational complexity in the number of users.

The suboptimum multiuser detection schemes perform worse than the optimum multiuser detector but have polynomial complexity. In a wireless channel where users come and go, and have time varying parameters such as channel gains (due to mobility) it is important to have blind adaptive implementations of these near-far resistant multiuser detectors. This thesis presents a blind adaptive implementation of the well known decorrelating multiuser detector. The proposed algorithm works with available stochastic observations and recursively updates the receiver filters of the users. The algorithm converges to the well known decorrelating detector for a synchronous CDMA channel and to the well known one-shot asynchronous decorrelating detector in the asynchronous CDMA channel, in the mean squared error sense. The algorithm does not require any parameter of any interfering user to construct the decorrelating detector of a desired user. Only the desired user's signature sequence and variance of the receiver noise are needed by the algorithm.

Power control is a method which assigns users with transmitter power levels such that each user creates as little interference to other users as possible and satisfies its own quality of service requirement defined in terms of the SIR. Power control algorithms developed to date and their convergence proofs have assumed that certain parameters, such as the SIR and the interference power needed by the algorithms, are known exactly. In a real wireless system, these parameters should be extracted from the received signal at the base stations. This extraction yields random estimates for the necessary parameters. If these random estimates are substituted directly into the deterministic algorithms, the convergence can no longer be guaranteed. This thesis introduces a class of power control algorithms which use readily available observations such as matched filter outputs. The proposed algorithms are proved to converge to the optimal power vector in the mean squared error sense. The proposed algorithms are distributed in the sense that each user needs the output of its own matched filter at its assigned base station and its own channel gain to its assigned base station to update its transmitter power.

This thesis combines the power control and multiuser detection approaches to combat the near-far problem of CDMA systems more effectively. This thesis presents an algorithm which iteratively updates the linear receiver filters and the transmitter powers of the users. The proposed algorithm converges to a power vector and a set of receiver filters for the users which are the solutions of the joint optimization of the communication between the users and the base stations in the sense of minimizing total transmitter power subject to SIR based quality of service requirements. At each iteration, the proposed algorithm first updates the receiver filters of the users to be the MMSE multiuser receivers given the power vector, then updates the powers of the users to meet the SIR requirements. The proposed algorithm is distributed because for both the receiver and the power update, the parameters needed by each user are available at the input and the output of the receiver filter of that user.

Finally this thesis introduces an iterative and distributed interference avoidance method where each user updates its signature sequence (transmit waveform) to minimize the interference it creates to other users. The algorithm converges to a set of signature sequences which maximizes the capacity whether it is defined in terms of information theoretic capacity or network capacity. At each iteration only one user's signature sequence is updated. Each user updates its signature sequence to be a generalized normalized MMSE receiver filter corresponding to its signature sequence. Since the MMSE receiver can be implemented distributedly, the proposed algorithm can be implemented in a distributed fashion as well.

## 7.1 Suggestions for Further Study

Although the tractability of the optimum multiuser detection under certain conditions is proved in Chapter 2, no concrete receiver structure for these cases is presented. What the FIFO preflow-push algorithm that can be used to solve the network flow version of the optimum multiuser detection corresponds to in terms of a receiver implementation should be investigated.

The combined power control and interference suppression algorithm of Chapter 5 is still a deterministic algorithm. As discussed in Section 5.5, a stochastic algorithm which updates the filter coefficients and the power levels of the users based on real random observations is desirable.

The iterative signature sequence set construction algorithm of Chapter 6 is currently for a single cell synchronous CDMA system. The algorithm must be generalized to asynchronous multicell systems with multipath effect. Also the algorithm of Chapter 6 is a deterministic one, a stochastic algorithm where each user updates its signature sequence based on real random observations is desirable.

# References


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

[2] R. K. Ahuja, T. L. Magnanti, and J. B. Orlin. *Network Flows: Theory, Algorithms and Applications.* 1993.

[3] M. Andersin, N. Mandayam, and R. Yates. Subspace-based estimation of the signal-to-interference ratio for TDMA cellular systems. In *Proc. IEEE Vehicular Technology Conference, VTC-96*, pages 1155–1159, 1996.

[4] M.D. Austin and G.L. Stuber. In-service signal quality estimation for TDMA cellular systems. In *Proceedings of the International Symposium on Personal Indoor Mobile Radio Communications PIMRC'95*, pages 836–840, 1995.

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

[6] L. F. Chang and S. Ariyavisitakul. Performance of a CDMA radio communications system with feed-back power control and multipath dispersion. In *IEEE Global Telecommunications Conference*, pages 1017–1021, December 1991.

[7] D. S. Chen and S. Roy. An adaptive multiuser receiver for CDMA systems. *IEEE Journal on Selected Areas in Communications*, 12(5):808–816, June 1994.

[8] 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.

[9] G. R. Cooper and R. W. Nettleton. A spread-spectrum technique for high-capacity mobile communications. *IEEE Transactions on Vehicular Technology*, 27(4):264–275, November 1978.

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

[11] A. Duel-Hallen. Decorrelating decision-feedback multiuser detector for synchronous code-division multiple-access channels. *IEEE Transactions on Communications*, 41(2):285–290, February 1993.

[12] 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.

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

[14] M. R. Garey and D. S. Johnson. *Computers and Intractability: a guide to the theory of NP-completeness.* W. H. Freeman, 1979.

[15] K. S. Gilhousen, I.M. Jacobs, R. Padovani, A. J. Viterbi, Jr. L. A. Weaver, and III C.E. Wheatley. On the capacity of a cellular CDMA system. *IEEE Transactions on Vehicular Technology*, 40(2):303–312, May 1991.

[16] 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.

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

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

[19] 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.

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

[21] M. Honig, U. Madhow, and S. Verdú. Blind adaptive multiuser detection. *IEEE Transactions on Information Theory*, 41(4):944–960, July 1995.

[22] J. Y. N. Hui. Throughput analysis for code division multiple accessing for the spread spectrum channel. *IEEE Journal on Selected Areas in Communications*, 2(4):482–486, July 1984.

[23] P. L. Hammer (Ivanescu). Some network flow problems solved with pseudo-boolean programming. *Operations Research*, 13:388–399, 1965.

[24] 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.

[25] 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.

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

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

[28] R. Lupas and S. Verdú. Linear multiuser detectors for synchronous code-division multiple-access channels. *IEEE Transactions on Information Theory*, 35(1):123–136, January 1989.

[29] R. Lupas and S. Verdú. Near-far resistance of multiuser detectors in asynchronous channels. *IEEE Transactions on Communications*, 38(4):496–508, April 1990.

[30] U. Madhow and M. L. Honig. MMSE interference suppression for direct-sequence spread-spectrum CDMA. *IEEE Transactions on Communications*, 42(12):3178–3188, December 1994.

[31] J. L. Massey. On welch's bound for the correlation of a sequence set. In *IEEE International Symposium on Information Theory (ISIT'91)*, page 385, 1991.

[32] J. L . Massey and T. Mittelholzer. Welch's bound and sequence sets for code-division multiple-access systems. *Sequences II: Methods in Communication, Security and Computer Science.* R. Capocelli, A. De Santis and U. Vaccaro, Eds. New York: Springer-Verlag, 1991.

[33] H. J. Meyerhoff. Method for computing the optimum power balancing in multibeam satellite. *COMSAT Technical Review*, 4(1), Spring 1974.

[34] S. L. Miller. An adaptive direct-sequence code-division multiple-access receiver for multiuser interference rejection. *IEEE Transactions on Communications*, 43(2/3/4):1746–1755, February/March/April 1995.

[35] 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.

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

[37] J. C. Picard and H. D. Ratliff. Minimum cuts and related problems. *Networks*, 5:357–370, 1974.

[38] R. L. Pickholtz, D. L. Schilling, and L. B. Milstein. Theory of spread-spectrum communications - a tutorial. *IEEE Transactions on Communications*, 30(5), May 1982.

[39] J. G. Proakis. *Digital Communications.* 1989.

[40] P. B. Rapajic and B. S. Vucetic. Adaptive receiver structures for asynchronous CDMA systems. *IEEE Journal on Selected Areas in Communications*, 12(4):685–697, May 1994.

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

[42] M. Rupf and J. L . Massey. Optimum sequence multisets for synchronous code-division multiple-access channels. *IEEE Transactions on Information Theory*, IT-40(4):1261–1266, July 1994.

[43] 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.

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

[45] R. A. Scholtz. The spread spectrum concept. *IEEE Transactions on Communications*, 25(8):748–755, August 1977.

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

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

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

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

[50] S. Ulukus and R. D. Yates. Adaptive power control and MMSE interference suppression. *ACM Journal of Wireless Networks, Special Issue on Multiuser Detection in Wireless Communications*. To appear.

[51] S. Ulukus and R. D. Yates. A blind adaptive decorrelating detector for CDMA systems. *IEEE Journal on Selected Areas in Communications*. To appear.

[52] S. Ulukus and R. D. Yates. Power control using stochastic measurements. In *34th Annual Allerton Conference on Communications, Control and Computing*, pages 845–854, October 1996.

[53] S. Ulukus and R. D. Yates. Adaptive power control with MMSE multiuser detectors. In *IEEE International Conference on Communications, ICC'97*, June 1997.

[54] S. Ulukus and R. D. Yates. Adaptive power control with SIR maximizing multiuser detectors. In *Conference on Information Sciences and Systems, CISS'97*, March 1997.

[55] S. Ulukus and R. D. Yates. A blind adaptive decorrelating detector for CDMA systems. In *Global Telecommunications Conference, GLOBECOM'97*, pages 664–668, November 1997.

[56] S. Ulukus and R. D. Yates. Optimum multiuser detection is tractable for synchronous CDMA systems using m-sequences. *IEEE Communications Letters*, 2(4):89–91, April 1998.

[57] S. Ulukus and R. D. Yates. Stochastic power control for cellular radio systems. *IEEE Transactions on Communications*, 46(6):784–798, June 1998.

[58] M. K. Varanasi. Asymptotically optimum multiuser decision feedback detection. In *IEEE International Symposium on Information Theory*, page 275, 1997.

[59] M. K. Varanasi and B. Aazhang. Multistage detection in asynchronous code-division multiple-access communications. *IEEE Transactions on Communications*, 38(4):509–519, April 1990.

[60] M. K. Varanasi and B. Aazhang. Near-optimum detection in synchronous code-division multiple-access systems. *IEEE Transactions on Communications*, 39(5):725–736, May 1991.

[61] S. Verdú. Capacity region of Gaussian CDMA channels: The symbol-synchronous case. In *24th Allerton Conference on Communication, Control and Computing*, pages 1025–1034, October 1986.

[62] S. Verdú. Minimum probability of error for asynchronous gaussian multiple-access channels. *IEEE Transactions on Information Theory*, 32:85–96, January 1986.

[63] S. Verdú. Computational complexity of optimum multiuser detection. *Algorithmica*, 4:303–312, 1989.

[64] S. Verdú. Multiuser detection. *Advances in Statistical Signal Processing*, 2:369–409, 1993.

[65] P. Viswanath, V. Anatharam, and D. Tse. Optimal sequences, power control and capacity of spread-spectrum systems with multiuser receivers. Preprint. Submitted to IEEE Transactions on Information Theory.

[66] A. J. Viterbi. Spread spectrum communications – myths and realties. *IEEE Communications Magazine*, pages 11–18, May 1979.

[67] A. J. Viterbi and A. M. Viterbi. Erlang capacity of power controlled CDMA system. *IEEE Journal on Selected Areas in Communications*, 11(6):892–900, August 1993.

[68] A. J. Viterbi, A. M. Viterbi, K. S. Gilhausen, and E. Zahavi. Soft handoff extends CDMA cell coverage and increases reverse link capacity. *IEEE Journal on Selected Areas in Communications*, 12(8):1281–1288, October 1994.

[69] X. Wang and H. V. Poor. Blind adaptive interference suppression for CDMA communications based on eigenspace tracking. In *Conference on Information Sciences and Systems, CISS'97*, pages 468–473, March 1997.

[70] L. R. Welch. Lower bounds on the maximum cross correlation of signals. *IEEE Transactions on Information Theory*, IT-20(3):397–399, May 1974.

[71] Z. Xie, R. T. Short, and C. K. Rushforth. A family of suboptimum detectors for coherent multiuser communications. *IEEE Journal on Selected Areas in Communications*, 8(4):683–690, May 1990.

[72] B. Yang. Projection approximation subspace tracking. *IEEE Transactions on Signal Processing*, 43(1):95–107, January 1995.

[73] 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.

[74] 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.

[75] K. W. Yip and T. S. Ng. Code phase assignment – a technique for high capacity indoor mobile DS-CDMA communications. In *IEEE Vehicular Tech. Conf., VTC'94*, pages 1586–1590, 1994.

[76] 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.

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

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

[79] 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.

# Vita

## Şennur Ulukuş

**1991** B.S. in Electrical and Electronics Engineering, Bilkent University, Ankara, Turkey.

**1991-93** Teaching/Research Assistant, Department of Electrical and Electronics Engineering, Bilkent University, Ankara, Turkey.

**1993** M.S. in Electrical and Electronics Engineering, Bilkent University, Ankara, Turkey.

**1993-95** Teaching Assistant, Department of Electrical and Computer Engineering, Rutgers University, Piscataway, NJ.

**1995-98** Graduate Research Assistant, Department of Electrical and Computer Engineering, Rutgers University, Piscataway, NJ.

**1996** S. Ulukus and R. D. Yates, "Power Control Using Stochastic Measurements," Proceedings of *34th Allerton Conference on Communications, Control and Computing*, Monticello, ILL, October 1996.

**1997** S. Ulukus and R. D. Yates, "Adaptive Power Control with SIR Maximizing Multiuser Detectors," Proceedings of *Conference on Information Sciences and Systems*, Baltimore, MD, March 1997.

**1997** S. Ulukus and R. D. Yates, "Adaptive Power Control with MMSE Multiuser Detectors," Proceedings of *IEEE International Conference on Communications*, Montreal, Quebec, Canada, June 1997.

**1997** S. Ulukus and R. D. Yates, "A Blind Adaptive Decorrelating Detector for CDMA Systems," Proceedings of *IEEE Global Telecommunications Conference*, Phoenix, AZ, November 1997.

**1998** S. Ulukus and R. D. Yates, "Optimum Multiuser Detection is Tractable for Synchronous CDMA Systems Using M-sequences," *IEEE Communications Letters*, April 1998.

**1998** S. Ulukus and R. D. Yates, "Stochastic Power Control for Cellular Radio Systems," *IEEE Transactions on Communications*, June 1998.

**1998** S. Ulukus and G. P. Pollini, "Handover Delay in Cellular Wireless Systems," Proceedings of *IEEE International Conference on Communications*, Atlanta, GA, June 1998.

**1998** S. Ulukus and R. D. Yates, "Adaptive Power Control and MMSE Interference Suppression," *ACM Journal of Wireless Networks*, to appear.

**1998** S. Ulukus and R. D. Yates, "A Blind Adaptive Decorrelating Detector for CDMA Systems," *IEEE Journal on Selected Areas in Communications*, to appear.

**1998** Ph.D. in Electrical and Computer Engineering, Rutgers University, Piscataway, NJ.