EFFICIENT ACCESS AND INTERFERENCE MANAGEMENT FOR CDMA WIRELESS SYSTEMS

BY AYLIN YENER

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

May, 2000

ABSTRACT OF THE DISSERTATION

Efficient Access and Interference Management for CDMA Wireless Systems

by Aylin Yener

Dissertation Director: Professor Roy D. Yates

Code Division Multiple Access (CDMA) has emerged as a promising access scheme to meet the ever growing capacity demand of wireless services. This thesis investigates several issues related to the characterization and enhancement of the capacity of circuit-switched and packet-switched CDMA systems.

In circuit-switched CDMA systems, there exists continuous communication between a user and the base station. The capacity is limited by the mutual interference between the users and can be enhanced by efficient interference management. This thesis investigates interference management techniques that combine receiver based temporal and spatial processing and transmit power control.

A nonlinear programming approach is used to develop approximate solutions for the optimum multiuser bit detection problem. This approach is helpful in understanding the convergence properties of various iterative multiuser detectors. A new detector which approximates the optimum detector is derived.

Next, this thesis investigates the problem of combined multiuser detection and receiver beamforming for CDMA systems with base station antenna arrays. In this case, it is possible to further increase system capacity by employing both temporal (multiuser detection) and spatial (beamforming) filtering. Filter structures are derived and

analyzed and the performance of the proposed schemes are compared.

Power control is an interference management technique that balances the received powers of the users so that no one user creates excessive interference to another. This thesis investigates the improvement possible by adding transmit power control to temporal-spatial filtering. Iterative power control algorithms are derived and analyzed where transmit powers and filters of the users are updated to achieve minimum total transmit power of all users while guaranteeing each user its required quality of service.

When advanced receivers, such as the temporal-spatial filters discussed in this thesis, are not available at the base station, cell sectorization is an effective interference management method to reduce the interference each user experiences. This thesis studies simple spatial capacity enhancement techniques and gives the solution of the joint transmit power control and cell sectorization problem where sectors are designed in response to the spatial density of users.

This thesis provides an introduction for the characterization of the user capacity of connectionless packet-switched CDMA systems. In such systems, users keep no connection to the base station unless they are actively sending a packet. Thus, users and their timings have to be acquired each time they try to send a packet. No signature sequence assignment is assumed. This thesis investigates the multiuser access capacity of such systems and concludes that the connectionless system considered is limited by the efficiency of the timing acquisition process.

Acknowledgements

Being at the end of this long journey called the graduate school, I have a lot to thank for and a lot of people to express gratitude to.

First, I am truly indebted to Prof. Roy Yates for his support and guidance throughout the duration of this thesis. It is not an overstatement to say that, without him, this thesis would not be possible. I am proud to be educated by such an innovative thinker who has an outstanding research discipline and the deepest insight to interesting research problems. I am especially impressed by his willingness and generosity in sharing all these gifts and his wisdom with me and all his students. The last three years have been a great educational experience for me. It has been a genuine pleasure to have Roy as my advisor and teacher.

I would like to express my thanks to my PhD committee members, Prof. Narayan Mandayam and Prof. Christopher Rose. I would especially like to thank Prof. Vincent Poor of Princeton University for being the external examiner in the committee. I am also indebted to Prof. David Goodman of Polytechnic University for his invaluable comments in the earlier stages of this thesis as well as introducing me to the area of Wireless Communications and to WINLAB. Being a part of WINLAB truly changed my life for the better.

I would like to express my heartfelt gratitude to Dr. Sennur Ulukus. Her invaluable contributions to this thesis are greatly appreciated. My research collaboration with her has been most enjoyable and intellectually stimulating. I am proud to have Sennur not only as a most valuable research colleague, but also my best friend for the past six years and for the next sixty years to come.

My special thanks are to my colleague and friend, Cem U. Saraydar, for his research contributions to this thesis, as well as for all conversations we shared over the years at WINLAB and elsewhere, technical or otherwise. I would also like thank all my colleagues (past and present) at WINLAB for all their friendship and always enjoyable discussions. Thanks to WINLAB graduates Mohammad Saquib, Ashwin Sampath, Mathew Thomas, Fang-Chen Cheng, Shalinee Kishore, and WINLAB students Ivana Maric, Ana Iacono, Cristina Comaniciu, Dragan Samardzija, Rajnish Sinha, Zhuyu Lei, to name a few. Thanks to Jianghong Luo and Lang Lin for all the enjoyable conversations during our rides to Princeton all this semester. Many thanks to Ivan Seskar of WINLAB, the best system manager on the face of this earth, for putting up with me all these years.

I am thankful to my family, and all my friends here at Rutgers and elsewhere around the world, for their continuous support. Finally, I thank my teachers, past and present, for without their guidance, I would not be here today.

Dedication

To all loved ones, past, present and future...

Table of Contents

List of Figures

Chapter 1

Introduction

The demand for high capacity flexible wireless services is ever-growing. Code Division Multiple Access (CDMA) is an access scheme that shows promise in meeting these demands. CDMA has been implemented in second-generation wireless systems that use the IS-95 standard [2]. Furthermore, CDMA based standards such as W-CDMA [6, 15, 16, 45] and cdma2000 [1] are to be implemented in third-generation wireless systems. Among the many advantages of code division multiple access scheme are the graceful degradation of system performance with increasing number of users, the ability to operate in asynchronous environments, and resistance to intentional jamming and multipath [51, 60, 80].

In contrast to TDMA and FDMA systems that disjointly partition the resources between users, in a CDMA system, users share the time and frequency bands. In a CDMA system, each user transmits continuously and modulates each of its bits with a signature sequence whose bandwidth is several times that of the original message. The result is a collection of users transmitting all the time over the entire available bandwidth. The capacity of this time and frequency shared channel, i.e., the CDMA system, can be characterized in terms of the maximum number of users that can be *guaranteed* reliable communication by this system, termed the *user capacity*, or the maximum total information theoretic capacity, termed the *sum capacity* [79].

This thesis investigates user capacity of different types of CDMA systems. Specifically, we classify CDMA systems as **circuit-switched** and **packet-switched** and investigate the limitations on and improvements to the system capacity for each case. For each case we will only investigate the uplink, i.e., the mobile-to-base link.

1.1 Circuit-Switched CDMA Systems

We first consider circuit-switched, or connection oriented CDMA systems. Circuit switched CDMA represents the conventional CDMA concept: users share the time and frequency resources, but each user is uniquely identified through its assigned signature sequence with which it modulates its transmitted bits. Continuous transmissions or very long messages are assumed for all users. Each user keeps a connection to its assigned base station using its unique signature leading to a connection mode that requires dedicated resources, hence the term circuit-switched.

Although in theory, users could transmit using signature sequences that are orthogonal to each other, in practice, the asynchronous nature of CDMA transmissions renders the implementation of orthogonal codes virtually impossible. Thus, users transmissions naturally interfere with each other. The performance of a circuit-switched CDMA system can be limited by the *interference* that users create for each other during information transfer. As the interference a user experiences gets worse, it becomes more difficult for the user to maintain reliable communication. In particular, a strong interferer –most probably near the base station– may singlehandedly destroy the communication link between a more distant weak user and the base station, (see Figure 1.1). This effect is widely known in the literature as the *near-far effect* [66].

Several techniques have been investigated up to date to enhance the capacity of a CS-CDMA system. These techniques could be transmitter and/or receiver based, and all aim to alleviate the near-far problem by controlling, suppressing or avoiding the interference between the users. Transmitter side techniques aim to reduce or avoid the interference users experience: users can control their transmit power levels [86, 95], choose their signature sequences to avoid interference [68, 79] or form transmit beams in certain directions if multiple antenna elements are available at the handset [55]. Receiver side techniques on the other hand, aim to suppress the interference using advanced receivers: multiuser detection can be used to exploit the inherent structure of the CDMA received signal due to the usage of signature sequences [78] and receiver beamforming or cell sectorization can be used to better receive the desired user's signal

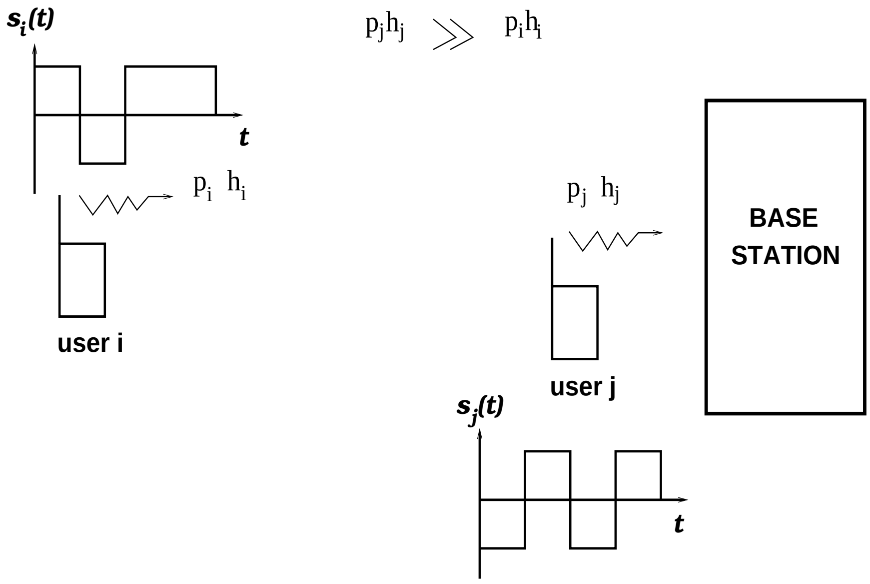

Figure 1.1: Illustration of near-far effect in CDMA: p_k and h_k transmit power and uplink gain of user k. The received power of user j, $p_j h_j$, is much higher than the received power of user i, $p_i h_i$ which causes user i to experience unacceptable interference levels, i.e., it can not communicate with the base station

if antenna arrays or directional antennas are available at the base station. All these techniques can be termed under the general heading of *Interference Management*.

In this thesis, we will investigate receiver processing techniques that aim to suppress the interference a user experiences, in most cases integrated with transmit power control which aims to reduce the interference each user creates.

We first look into the *multiuser detection* problem. In very general terms, multiuser detection deals with designing advanced receivers that process the CDMA received signal in time domain so that the interference a user experiences is suppressed [78]. Henceforth, we will use the terms *multiuser detection* and *temporal filtering* synonymously. The need for designing advanced temporal filters for multiuser CDMA systems was first shown in [74].

Consider the simplest multiuser CDMA system, a single cell system with N active users that are all synchronous and employing BPSK, for clarity of exposition. The received signal at the base station is given as

$$r(t) = \sum_{k=1}^{N} \sqrt{p_k h_k} b_k s_k(t) + n(t) \tag{1.1}$$

where p_k, h_k, b_k and $s_k(t)$ are the transmitter power, uplink gain, information bit and signature waveform of user k and $n(t)$ is the additive white Gaussian noise (AWGN) process with power spectral density σ^2. The multiuser detection problem is to determine the bits transmitted by each user $(b_i, k = 1, \cdots, N)$ given the received signal $r(t)$. It was shown in [76] that if the received signal is passed through a bank of filters each of which is matched to the signature of a particular user, the resulting outputs

$$y_k = \int_0^{T_b} r(t) s_k(t) dt \qquad k = 1, \cdots, N \tag{1.2}$$

are sufficient statistics for this detection problem. The kth matched filter output is given by

$$y_k = \sqrt{p_k h_k} b_k + \sum_{j \neq k} \sqrt{p_j h_j} \Gamma_{kj} b_j + n_k \tag{1.3}$$

where

$$\Gamma_{kj} = \int_0^{T_b} s_k(t) s_j(t) dt \tag{1.4}$$

is the cross correlation between the signatures of user k and j and

$$n_k = \int_0^{T_b} s_k(t) n(t) dt \tag{1.5}$$

is a Gaussian random variable with zero mean and variance σ^2.

In (1.3), only the first term contains the bit b_k of user k, the second and third terms are the multiple access interference and noise, respectively, and represent the total disturbance. Strictly speaking, unless all Γ_{kj} are identically 0, i.e., all interfering signatures are orthogonal to the kth signature, this disturbance is neither Gaussian nor white. Thus, the single user detector $\hat{b}_k = sgn(y_k)$, which is the minimum probability of error detector in a single user AWGN channel, is suboptimum in terms of the minimizing the bit error probability in a multiuser interference channel [75].

Reference [77] derived the joint maximum likelihood multiuser detector, i.e., *the optimum multiuser detector*, for all users. It was shown in [76] that the optimum multiuser detection problem is an NP-hard optimization problem. This implies there exist systems where, to detect the bits of all N active users, one needs to evaluate the resulting likelihood function for all possible bit sequences that may have been transmitted. For BPSK, this results in a computational complexity of 2^N. Thus, for a reasonable number of users, solving the optimum multiuser detection problem is not tractable for arbitrary non-orthogonal signature sequences.

This observation resulted in the development of a number of less complex *multiuser detectors* that nevertheless achieve near-optimum performance [78]. These suboptimum receivers have been motivated by several criteria other than the minimum bit error probability, see for example [35, 36]. In chapter 2, we construct a nonlinear programming based approach to the multiuser detection problem, where we derive several approximation algorithms to the original multiuser detection problem. This framework derives well known suboptimum multiuser detectors and a newly proposed detector as solutions to relaxations of the optimum multiuser detection problem. Using this approach provides us with an understanding of the behavior of different suboptimum detectors.

Multiuser detection is an important receiver processing technique to suppress the interference that users experience and consequently, it increases the capacity of a CDMA system. However, it is not the only possible receiver based method for interference management. When antenna arrays are employed at the base station, spatial processing may be used to suppress the interference. In fact, exploiting spatial diversity through the use of array combining to increase the system capacity is a familiar notion in both narrowband and wideband wireless systems [73]. For narrowband systems, it was shown that multiple antennas can be used to null out interferers and achieve diversity gain [84]. For CDMA systems, employing antenna arrays at the base station, and combining the outputs of multiple antenna array elements to make bit decisions for each user was first proposed in [42]. The array combiner for each user is matched to its *spatial signature*, i.e., the vector of the user's gains to each antenna array element. Array combiners, or *spatial filters*, that result in better performance in a multiuser environment by exploiting the correlations between users' spatial signatures are also available [50].

In chapter 3, we investigate receiver processing techniques when both multiuser detection and beamforming are employed at the base station. Linear processing is assumed in both the temporal and the spatial domains and the temporal-spatial filters are denoted by two-dimensional matrix filters. Within this framework, we formulate several possible filter structures and evaluate their performance.

An effective interference management method at the transmitter side is *power control*. Power control refers to assigning users with transmitter power levels so as to minimize the interference users create to each other while achieving a certain quality of service which is defined in terms of the Signal to Interference ratio (SIR) [86]. The power control problem was first identified as an eigenvalue problem for non-negative matrices and the solution was found by a matrix inversion, i.e., in a centralized and non-iterative fashion [24, 95]. This was followed by the development of iterative and distributed algorithms that require only local measurements [21,86,94]. The traditional iterative power control approach, henceforth called *conventional power control*, assumes that a single antenna and matched filter receivers are being used at the base stations. Each user employs an SIR based power update where the user's power is multiplied by the ratio of its target SIR to its current SIR, i.e., for user i, the update is

$$p_i(n+1) = \frac{\gamma_i^*}{\gamma_i(n)} p_i(n) \tag{1.6}$$

where $p_i(n)$ and $\gamma_i(n)$ are the power and SIR of user i at iteration n, and γ_i^* is the SIR *target* of user i. The simple intuition behind this iteration is that if the current SIR $\gamma_i(n)$ of user i is less than the target SIR γ_i^* then the power of that user is increased; otherwise, it is decreased. More recently, power control algorithms when multiuser detectors are present at the base station and power control algorithms when receiver beamforming is employed were studied in [67] and [56] respectively. It was shown that the use of temporal filtering [67] or the use of spatial filtering [56] in conjunction with power control increased the system capacity compared to conventional power control.

In chapter 4, we consider interference management through the combination of spatial *and* temporal combining and power control for CDMA systems. Specifically, we develop multistage iterative power control algorithms in which each user chooses its transmit power, linear receiver filter and beamformer so as to expend as little power as possible while achieving its required quality of service. Because the performance of the system is optimized over all three variables, power levels as well as the temporal and spatial receiver filters, it is conceivable that the system capacity will be superior to that of optimizing only one or two of the variables [56, 67, 86]. Indeed, the resulting algorithms yield solutions for which the total transmitted power of the users in the system is reduced significantly compared to systems employing conventional power control [86], joint power control and multiuser detection [67] and joint power control and beamforming [56]. Thus, the system can accommodate larger number of users and/or higher quality of service requirements, both of which translate into a system capacity increase.

The combined use of temporal and spatial filtering significantly enhances the system capacity. However, as the resulting receiver structures may be computationally complex compared to matched filters, their implementation may not always be practical. In such cases, one might ask how to reduce the interference at the receiver side via simpler techniques. One plausible answer is to sectorize the cells using directional antennas. While this approach still utilizes the spatial domain to introduce orthogonalization to the system, it is fundamentally different than receiver beamforming [50,73]. Beamforming combines the received signals from multiple antennas in a unique way for each user to suppress the interference that user sees. Sectorization merely employs directional antennas and each user's signal is received at only one of these antennas. Since only a subset of the users are received at each antenna, the interference that each user sees is less compared to a single antenna system. No spatial combiner is used. While this is a more *constrained* scheme than beamforming, the simplicity of the resulting receivers is appealing and sectorization can be quite beneficial especially for static systems, i.e., if the users do not have to handoff from sector to sector frequently. Sectorization combined with transmit power control thus shows promise in being a simple interference management technique for CDMA systems.

In chapter 5, we investigate the best way of sectorizing a cell so that the total power of all the users is minimized while each user gets its required SIR. It is assumed that the number of sectors and terminal locations are given. The design of the sectors are *adaptive* to user density in the particular cell. It is shown that power savings are achieved by the best sector assignment over the *uniform* sectorization which divides the cell into equal width sectors and power assignments that are infeasible with uniform sectorization can be handled with nonuniform sectorization. Thus, more users can be supported in the cell.

1.2 Packet-Switched CDMA Systems

For circuit-switched CDMA systems, interference management can be employed once a user is incorporated into the system, i.e., once its parameters are acquired. A signature sequence is assigned to the user and the connection between the system and the user is kept even when the user does not transmit information because of the need for synchronization information. The circuit-switched CDMA model is convenient for voice communications [2] and even for voice-data systems with long data sessions [54]. However, future CDMA systems need to provide more diverse types of applications. In particular, packet based services are being proposed for the third generation systems [6, 16, 45]. For these services to be successfully rendered, systems must be flexible in terms of resource partitioning. Thus, assigned signature sequences and keeping a continuous connection even with a lower synchronization rate become luxuries and should be avoided. In chapter 6 of this thesis, we turn our attention to CDMA systems with no dedicated resources. The *connectionless* architecture of these systems prompted the term *Packet-Switched CDMA*.

In a packet-switched CDMA system, multiple users share the same resources: the same frequency and time bands and the same set of signature sequences. Thus, for users to be able to communicate with the system, the base station has to determine the existence of active users and their parameters the most important of which is the timing delay of each user. We formulate the *Multiuser Access Detection* problem to jointly determine the existence of the active users and to estimate their timing. We find that the optimum solution is of high complexity and investigate suboptimum solutions.

The users in such a random access based system can collide with each other if users' delays are not sufficiently separated. The capacity of such systems, i.e., the number of users that can be served simultaneously, and the factors that affect the capacity are investigated in this thesis.

1.3 Thesis Road Map

Chapter 2 describes the nonlinear programming approach to multiuser detection. Chapter 3 investigates joint temporal-spatial receiver structures. Chapter 4 describes integrating optimum power control into a system employing joint temporal-spatial receivers. Chapter 5 solves the joint power control and cell sectorization problem. Chapter 6 deals with the multiaccess capacity of packet-switched CDMA systems. Chapter 7 presents the conclusions of this thesis along with suggestions for future work.

For each chapter, we give detailed introductions and literature reviews about the specific problem at the beginning of the chapter as necessary. Chapters 2 and 3 deal with a single cell CDMA system, consistent with the receiver processing literature. Chapter 4 gives results for a multicell CDMA system, consistent with the power control literature. Chapter 5 solves the optimum sectorization and power control problem for a single cell, and the results are also extended to sectorizing a *desired* cell in a multicell environment. Finally, chapter 6 investigates the multiaccess capacity of a packet-switched CDMA system with a single base station. We present the corresponding system model following the introduction in each chapter in an attempt to make the chapters more independently readable.

Chapter 2

CDMA Multiuser Detection: A Nonlinear Programming Approach

2.1 Introduction

In this chapter, we will investigate the problem of detecting the bits of multiple users transmitting CDMA signals.

Maximum likelihood detection of multiple users' bits has been shown to be an NP hard problem [76]. Following this observation, many suboptimum receivers with reasonable complexity yet near-optimum performance have been proposed [18, 35, 36, 71]. Among the most popular linear detectors are the decorrelator [35] and the MMSE receiver [36]. The decorrelator suppresses the multiple access interference totally while enhancing the Gaussian noise and is the optimum detector if the received powers of the users are unknown at the receiver [78]. The MMSE receiver [36] gives the minimum mean squared error between the filter output and the transmitted bit, and also maximizes the output signal to interference ratio. Both detectors are optimum when no noise is present. Decision directed nonlinear detectors that generally result in lower bit error rates have been proposed in [18, 71]. The decorrelating decision feedback detector [18] cancels the interference from the users that have already been decoded, and suppresses interference from users that have not yet been decoded, to obtain the decision statistic for the current user. The multistage detector [72] uses the previous stage bit estimates of all interferers to reconstruct and cancel the interference to the current user in the current stage. The interference canceller in [49] subtracts the interference estimate of the previously decoded users from the matched filter output for the current user and uses this decision statistic to decode the bit of the current user.

Our aim in this chapter is to approach the optimum multiuser detection problem from a nonlinear programming point of view. The original optimum multiuser detection problem (OMUD) is a $0-1$ quadratic program for which there exists no efficient algorithm. The general approach in the presence of such hardship is to approximate the solution by working on an *easier* problem that can be solved efficiently. The easier problem to be solved is a *relaxation* of the original problem. The relaxed solution is then mapped to the solution set of the original problem, ideally arriving at a near optimum solution. For the multiuser detection problem, a relaxation often corresponds to a near optimum multiuser detector.

Using a nonlinear programming approach, we observe that some popular suboptimum detectors are relaxed solutions, i.e., approximations to the optimum detection problem. This approach helps us understand previously unidentified convergence properties of some known iterative nonlinear detectors. Furthermore, a new relaxation method is proposed that results in a simple iterative detector whose performance is then evaluated.

2.2 OMUD and its Relaxations

We consider a synchronous CDMA system employing BPSK modulation. Consistent with the existing multiuser detection literature [78], we assume the system has a single cell. The received signal at the base station for such a system is given by (1.1). We restate it here for convenience.

$$r(t) = \sum_{i=1}^{N} \sqrt{q_i} b_i s_i(t) + n(t) \tag{2.1}$$

In (2.1), N is the number of users, $q_i = p_i h_i$ and b_i are received power and the transmitted bit (± 1 equiprobably) and $s_i(t)$ is the unit energy signature, i.e. $\int_0^T s_i(t)s_i(t)dt = 1$, of user i, and $n(t)$ is the additive white Gaussian noise (AWGN) process with power spectral density σ^2. The matched filter output of the ith user is given by

$$y_i = \int_0^T r(t)s_i(t)dt \tag{2.2}$$

The received signal vector at the output of the matched filters, $\mathbf{y} = [y_1 \cdots y_N]^\top$, is a sufficient statistic for the multiuser detection problem and is given by

$$\mathbf{y} = \mathbf{\Gamma}\mathbf{\Lambda}\mathbf{b} + \mathbf{n} \tag{2.3}$$

where $\mathbf{\Gamma}$ is the nonnegative definite cross correlation matrix with

$$\Gamma_{ij} = \int_0^T s_i(t)s_j(t)dt$$

, $\mathbf{\Lambda}$ is a diagonal matrix containing the users' received amplitudes $\Lambda_{ii} = \sqrt{q_i}$, $\mathbf{b}$ is the vector containing the information bits of the users and $\mathbf{n}$ is a zero mean Gaussian random vector with auto covariance matrix $E[\mathbf{n}\mathbf{n}^\top] = \sigma^2\mathbf{\Gamma}$.

The aim of multiuser detection is to recover the information bits, $\mathbf{b}$. The solution of the optimum multiuser detection problem (OMUD) [35, 75, 78] employs the maximum likelihood estimate of $\mathbf{b}$ given $\mathbf{y}$. The optimum estimate for the bit vector is $\mathbf{b}^*$, if $\mathbf{b}^*$ minimizes the cost function which is quadratic in $\mathbf{b}$. Specifically,

$$\mathbf{b}^* = \arg \min_{\mathbf{b}\in\{-1,1\}^N} \mathbf{b}^\top\mathbf{R}\mathbf{b} - 2\mathbf{b}^\top\mathbf{\Lambda}\mathbf{y} \tag{2.4}$$

where $\mathbf{R} = \mathbf{\Lambda}\mathbf{\Gamma}\mathbf{\Lambda}$ with $R_{ij} = \sqrt{q_i}\sqrt{q_j}\Gamma_{ij}$. Note that, 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. However, when this is not the case, the computational complexity of the problem is much worse. Although it has been shown recently that certain special $\mathbf{R}$ structures allow construction of polynomial time algorithms to find the optimum solution [58, 69], the problem for general correlation matrices remains NP hard and one can find the optimum $\mathbf{b}$ only by exhaustive search of 2^N candidate vectors [76].

It is interesting to note that this model can also be applied to detection on a multiple bit window in an asynchronous CDMA system with multipath fading [19]. If the arrival times of the L propagation paths are known, it is sufficient to observe the outputs of the $N \times L$ matched filters for all paths of all users over the length of the window. Recently, it has also been shown that when antenna arrays are employed at the receiver, the sufficient statistics are obtained by passing the received signal through beamformers directed to each path of each user and finally combining them with a bank of N RAKE combiners [81]. For each of these models, with careful definition of cross correlations between the signatures of all users over the multiple bit windows, one can construct the corresponding $\mathbf{R}$ matrix and the optimum detection problem can be shown to be of the form of (2.4); see, for example, equation (19) in [81]. Thus, the methodology and the resulting algorithms we use in this work remain valid for the most general model, although we will work with the synchronous model for simplicity.

In addition, we will concentrate on cases where the signatures of the users are independent and $\mathbf{\Gamma}$ and hence $\mathbf{R}$ are positive definite. In this case, the objective function in (2.4) is strictly convex in $\mathbf{b}$ and has a well defined unique minimizer over a convex set. Thus, we can find solutions by relaxing the constraint set –which in the original problem contains only the corners of the unit hypercube– such that the resulting "relaxed" constraint set is convex. Figure 2.1 shows the different relaxed constraint sets for the two-user case. Three basic relaxations that we will consider in the sequel are: relaxation of the constraint set to the whole unit hypercube (region I), relaxation of the constraint set to the sphere which covers the unit hypercube (regions I+II), and relaxation to whole space (regions I+II+III). Note that we require the constraint set for each relaxation to contain the feasible set of the original problem. The solution can then be mapped to the feasible set of the original problem by taking the sign of each component of the relaxed solution vector (since bits are equiprobably ± 1).

2.3 Decorrelator

We first consider the simplest relaxation, where the feasible set is relaxed to contain the N dimensional space $\mathcal{R}^N$.

$$\min_{\mathbf{b}\in\mathcal{R}^N} \mathbf{b}^\top\mathbf{R}\mathbf{b} - 2\mathbf{b}^\top\mathbf{\Lambda}\mathbf{y} \tag{2.5}$$

This relaxation effectively removes the constraints and converts the discrete optimization problem into a continuous one. Since the cost function is convex in its variable, this problem has a unique minimum:

$$\hat{\mathbf{b}} = \mathbf{R}^{-1}\mathbf{\Lambda}\mathbf{y} = \mathbf{b} + \mathbf{\Lambda}^{-1}\mathbf{\Gamma}^{-1}\mathbf{n} \tag{2.6}$$

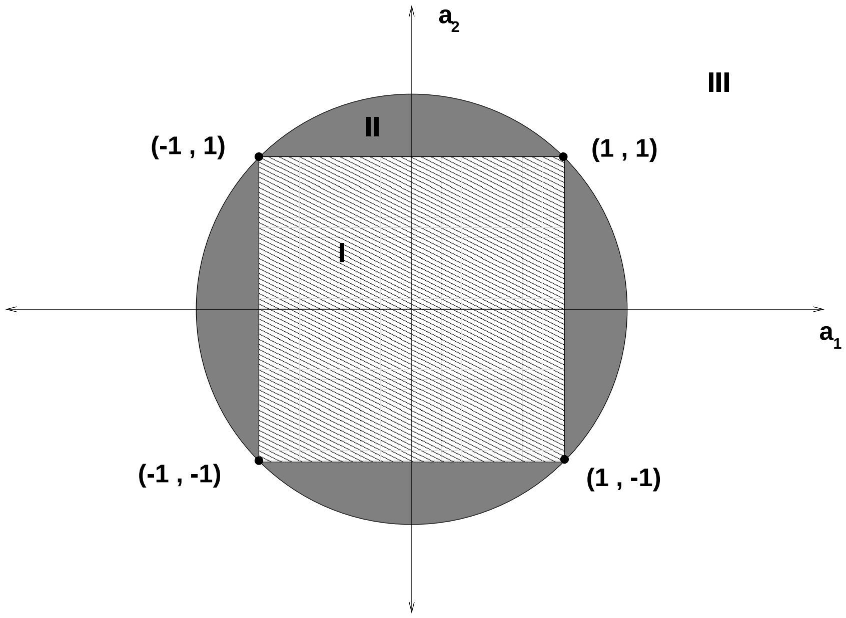

Figure 2.1: Relaxed constraint sets for the two user system which yield the following detectors: I - Soft Interference Canceller (Section 2.4), I+II - Generalized MMSE (Section 2.5), I+II+III - Decorrelator (Section 2.3)

Taking the sign of the solution vector yields the well known *decorrelating detector* [35].

2.4 Soft Interference Cancellation

The constraint set of the optimum multiuser detection problem (2.4) consists of the corner points of the unit hypercube. An effective approximation method is to relax the constraint set to cover to whole hypercube and use nonlinear programming algorithms to find the solution of the new convex programming problem [28]. The relaxed problem is:

$$\mathbf{b}^* = \arg \min_{\mathbf{b} \in [-1,1]^N} \mathbf{b}^\top \mathbf{R} \mathbf{b} - 2\mathbf{b}^\top \mathbf{\Lambda} \mathbf{y} \tag{2.7}$$

Both the cost function and the constraint set in (2.7) are convex. Thus, (2.7) has a unique minimum. However, the optimum point does not have a closed form and one should use iterative methods to get to the solution. One class of iterative methods that can be used are the constrained gradient methods. Further, the simplicity of the constraint set, i.e. the fact that it has a cartesian product form, enables us to define special iterative projection algorithms [13]. In particular, the following two algorithms, the *nonlinear Gauss-Seidel* and the *nonlinear Jacobi* algorithms respectively, converge to the minimum of (2.7) under certain conditions. Let

$$g(b_1, ..., b_N) = \mathbf{b}^\top \mathbf{R}\mathbf{b} - 2\mathbf{b}^\top \mathbf{\Lambda}\mathbf{y}$$

denote the function to be minimized. In the Gauss-Seidel iteration $b_i(t+1)$ is found by

$$b_i(t+1) = \arg \min_{x \in [-1,1]} g(b_1(t+1), \cdots, b_{i-1}(t+1), x, b_{i+1}(t), \cdots, b_N(t)) \tag{2.8}$$

while in the Jacobi iteration $b_i(t+1)$ is found by

$$b_i(t+1) = \arg \min_{x \in [-1,1]} g(b_1(t), \cdots, b_{i-1}(t), x, b_{i+1}(t), \cdots, b_N(t)) \tag{2.9}$$

respectively, where t is the stage (iteration) index. Both algorithms optimize one variable at a time to get to the optimum point of (2.7); however (2.8) uses the current stage estimates of some of the users while (2.9) allows a parallel implementation. Through a straightforward derivation, it can be shown that the above iterations yield the following two-step algorithms. For each user i, the first step for the Gauss-Seidel iteration is,

$$\hat{x}(t+1) = \frac{1}{\sqrt{q_i}} \left(y_i - \sum_{j=1}^{i-1} \sqrt{q_j} \Gamma_{ji} b_j(t+1) - \sum_{j=i+1}^{N} \sqrt{q_j} \Gamma_{ji} b_j(t) \right) \tag{2.10}$$

and the first step for the Jacobi iteration is,

$$\hat{x}(t+1) = \frac{1}{\sqrt{q_i}} \left(y_i - \sum_{j=1, j \neq i}^{N} \sqrt{q_j} \Gamma_{ji} b_j(t) \right) \tag{2.11}$$

The second step for both algorithms is

$$b_i(t+1) = \begin{cases} -1, & \hat{x}(t+1) < -1 \\ \hat{x}(t+1), & -1 \leq \hat{x}(t+1) \leq 1 \\ 1, & \hat{x}(t+1) > 1 \end{cases} \tag{2.12}$$

At each stage, to get the estimate of each user's bit, both receivers use soft estimates of the bits to reconstruct the interference and subtract this estimate from the user's matched filter output, scale the result by the amplitude of the user and project onto $[-1, 1]$. The difference between the two is that while the Gauss-Seidel algorithm uses the available current stage estimates of the users, i.e. feedback from a group of users whose bit estimates are already computed, the Jacobi algorithm uses only bit estimates from the previous stage. In multiuser detection terms, the nonlinear Gauss-Seidel algorithm corresponds to the *successive soft interference canceller*, whereas the nonlinear Jacobi algorithm corresponds to the *parallel soft interference canceller*.

In the following, we establish the convergence of these two algorithms. For that, we use results from [13] and omit the proofs here. First, we state the following theorem for the convergence of the nonlinear Gauss-Seidel algorithm.

Theorem 1 (Proposition 3.9 of Section 3.3 in [13]) *Suppose that $g : \mathcal{R}^N \to \mathcal{R}$ is continuously differentiable and convex on the set X where X is a Cartesian product of sets X_i and each X_i is a closed convex subset of $\mathcal{R}^{n_i}$. Furthermore, suppose that for each i, g is a strictly convex function of x_i, when the values of the other components of $\mathbf{x}$ are held constant. Let $\mathbf{x}(t)$ be the sequence generated by the nonlinear Gauss-Seidel algorithm, assumed to be well defined. Then every limit point of $\mathbf{x}(t)$ minimizes g over X.*

In our case, $g(\mathbf{b}) = \mathbf{b}^\top \mathbf{R}\mathbf{b} - 2\mathbf{b}^\top \mathbf{\Lambda}\mathbf{y}$, and for a positive definite $\mathbf{\Gamma}$, the function $g(\mathbf{b})$ is convex in $\mathbf{b}$ and is strictly convex in each variable b_i when the values of the other components of $\mathbf{b}$ are held constant. The convex set $X = [-1, 1]^N$ is in cartesian product form. The convexity of $g(\mathbf{b})$ ensures the uniqueness of the convergence point which is the global minimum. The following corollary states the convergence condition for the successive soft interference canceller.

Corollary 1 *For a CDMA system with linearly independent signature sequences, the nonlinear Gauss-Seidel algorithm (the successive soft interference canceller) always converges to the minimizer of (2.7).*

Theorem 1 shows that the nature of the problem at hand gives us the convergence of the Gauss-Seidel algorithm without any further restrictions. Unfortunately, this is not the case for the Jacobi algorithm. Convergence of the parallel soft interference canceller can be guaranteed under certain contraction assumptions as indicated by the following theorem. We first state the theorem and then derive the conditions for it to hold.

Theorem 2 (Proposition 3.10 of Section 3.3 in [13]) *Let $g : \mathcal{R}^N \to \mathcal{R}$ be continuously differentiable, let γ be a positive scalar, and suppose that the mapping $T : X \to \mathcal{R}^N$, defined by $T(\mathbf{x}) = \mathbf{x} - \gamma \nabla g(\mathbf{x})$, is a contraction with respect to the block maximum norm $\|\mathbf{x}\|_\infty^w = \max_i \|x_i\|_i / w_i$, where each $\|\cdot\|_i$ is the Euclidean norm on $\mathcal{R}^{n_i}$ and each w_i is a positive scalar. Then, there exists a unique vector $\mathbf{x}^*$ which minimizes g over X. Furthermore, the sequence $\mathbf{x}(t)$ generated by either of the Gauss-Seidel and the Jacobi algorithms converges to $\mathbf{x}^*$ geometrically.*

For our problem $g(\mathbf{x}) = \mathbf{x}^\top \mathbf{R}\mathbf{x} - 2\mathbf{x}^\top \boldsymbol{\Lambda}\mathbf{y}$ and $T(\mathbf{x})$ is given by

$$T(\mathbf{x}) = \mathbf{x} - \gamma(2\mathbf{R}\mathbf{x} - 2\boldsymbol{\Lambda}\mathbf{y}) = (\mathbf{I} - 2\gamma\mathbf{R})\mathbf{x} + 2\gamma\boldsymbol{\Lambda}\mathbf{y} \tag{2.13}$$

In order to guarantee the convergence of the Jacobi algorithm (parallel soft interference canceller), we need $T(\mathbf{x})$ to be a contraction mapping with respect to a block maximum norm. For any matrix $\mathbf{A}$, the block maximum norm $\|\mathbf{A}\|_\infty^w$ is defined [13] as

$$\|\mathbf{A}\|_\infty^w = \max_i \frac{1}{w_i} \sum_j |A_{ij}| w_j$$

A mapping $T(\mathbf{x})$ is called a contraction mapping with respect to a block maximum norm $\|\cdot\|_\infty^w$, iff, for some $\alpha \in [0, 1)$, $T(\mathbf{x})$ satisfies

$$\|T(\mathbf{x}) - T(\mathbf{x}')\|_\infty^w < \alpha \|\mathbf{x} - \mathbf{x}'\|_\infty^w \tag{2.14}$$

for all $\mathbf{x}$ and $\mathbf{x}'$. For the mapping $T(\mathbf{x})$ in (2.13), this condition is equivalent to

$$\|\mathbf{I} - 2\gamma\mathbf{R}\|_\infty^w < 1 \tag{2.15}$$

If (2.15) holds, then the nonlinear Jacobi algorithm converges by Theorem 2. However, for a given CDMA system, it may be difficult to verify whether (2.15) holds. Our aim is to derive an easier condition to check for the convergence of the nonlinear Jacobi algorithm. It is desirable for this condition to be independent of the received powers of the users, γ and a particular block maximum norm $\|\cdot\|_\infty^w$. The following Lemma states the condition for the convergence of the nonlinear Jacobi algorithm (parallel soft interference canceller) in terms of only the cross correlation matrix. Using $\rho(\cdot)$ to denote the maximum eigenvalue of a matrix, we have

Lemma 1 *For a CDMA system with linearly independent signature sequences, the non-linear Jacobi algorithm (parallel soft interference canceller) converges to the minimizer of (2.7), if*

$$\rho(\bar{\mathbf{\Gamma}}) < 1 \tag{2.16}$$

where $\bar{\mathbf{\Gamma}}$ *is constructed from* $\mathbf{\Gamma}$ *as* $\bar{\Gamma}_{ij} = |\Gamma_{ij}|$ *if* $i \neq j$*, and* $\bar{\Gamma}_{ii} = 0$*.*

Proof: Lemma 1 Using the definition of the block maximum norm, it can be shown that for a small enough γ, (2.15) is equivalent to

$$\max_i \left\{ (1 - 2\gamma R_{ii}) + 2\gamma \sum_{j \neq i} |R_{ij}| \frac{w_j}{w_i} \right\} < 1 \tag{2.17}$$

which is satisfied iff

$$-R_{ii} + \sum_{j \neq i} |R_{ij}| \frac{w_j}{w_i} < 0 \qquad \text{for all } i \tag{2.18}$$

Substituting $R_{ij} = \sqrt{q_i}\sqrt{q_j}\Gamma_{ij}$ in (2.18) and noting $\Gamma_{ii} = 1$, we arrive at

$$\sum_{j \neq i} |\Gamma_{ij}| \frac{\sqrt{q_j}}{\sqrt{q_i}} \frac{w_j}{w_i} < 1 \qquad \text{for all } i \tag{2.19}$$

Let $\bar{w}_j = w_j \sqrt{q_j}$. Using the definition of $\bar{\mathbf{\Gamma}}$, (2.19) is equivalent to

$$\|\bar{\mathbf{\Gamma}}\|_\infty^{\bar{w}} < 1 \tag{2.20}$$

Finally, Corollary 6.2 in Section 2.6 of [13] states that (2.20) is satisfied iff $\rho(\bar{\mathbf{\Gamma}}) < 1$, completing the proof. □

To summarize, for a given cross correlation matrix, one can check if (2.16) holds by constructing $\bar{\mathbf{\Gamma}}$ and finding its maximum eigenvalue. If it does, then by Theorem 2, the parallel soft interference canceller is guaranteed to converge to the minimizer of (2.7).

In the proof of Lemma 1, we arrived at (2.16) from the equivalent condition (2.20). An interesting observation is that choosing $\bar{w}_i = 1$ for all i, (2.20) reduces to a diagonal dominance condition, $\sum_{j \neq i} |\Gamma_{ij}| < \Gamma_{ii}$, for all i. This is clearly much easier to verify; however it is a more lenient condition for convergence since a particular set of $\bar{\mathbf{w}}$ values are used. It is possible for the parallel soft interference canceller to converge for a system with its cross correlation matrix satisfying (2.16) even if is not diagonally dominant (see Section 2.6).

It is also interesting to note that both the diagonal dominance condition mentioned above and the maximum eigenvalue condition in (2.16) are satisfied for a system where users' signatures are shifted versions of a basic m-sequence. In this case, $\Gamma_{ij} = -1/G$, $i \neq j$, and $\rho(\bar{\mathbf{\Gamma}}) = (N-1)/G$, and $\rho(\bar{\mathbf{\Gamma}}) < 1$ as long as $N \leq G$ which by definition is the case. Diagonal dominance condition is also equivalent to $(N-1)/G < 1$ for m-sequences. Thus, if these sequences are used, both Jacobi and Gauss-Seidel algorithms, i.e., parallel and successive soft interference cancellers, converge to the minimizer of (2.7).

In general, it takes more than one iteration for either algorithm to converge and thus the resulting receivers are *multi-stage* receivers. Multi-stage receivers are familiar in multiuser detection. In [71], hard decision bit estimates are used to reconstruct and subtract the interference for each user. The receiver is implemented in a parallel fashion as in (2.11) and is not convergent. Reference [44] proposes a class of receivers based on the SAGE algorithm, one of which is the successive multistage receiver (2.10) and argues that the SAGE based hard decision multistage receiver is convergent even when its parallel counterpart is not. The soft decision versions of these multistage receivers, i.e. (2.10) and (2.11), are proposed in [44] and [96] where they are called receivers with linear clippers. By representing these receivers in the form of iterative nonlinear programming algorithms, we have shown that *both* of these soft decision receivers, i.e. the parallel and the successive soft multi-stage interference cancellers, if they converge, converge to the same point which is the minimizer of (2.7).

We have observed that the Gauss-Seidel algorithm has a more lenient requirement in order to converge, i.e. linear independence of the signature sequences is sufficient, while the condition that guarantees the Jacobi algorithm to converge is more strict and is given in (2.16). Typically, Gauss-Seidel type iterations have faster convergence since they use the newest estimates. On the other hand, Jacobi type iterations can be executed in a completely parallel fashion since they do not require feedback from the current stage estimate of any user. Note that, if Theorem 2 is valid, any combination of the two algorithms also converges to the minimum of (2.7), i.e. some users can use the successive soft multi-stage receivers and others can use the parallel soft multi-stage receivers.

It is worthwhile to note that, one can implement the decorrelator given by (2.6) iteratively. Gauss-Seidel and Jacobi algorithms that converge to (2.6) can be found to be the algorithms derived in this section without the second stage $[-1, 1]$ clippers. It is also possible to derive Gauss-Seidel and Jacobi iterations that converge to the MMSE detector [36] which estimates the bits by taking the sign of $\bar{\mathbf{b}} = (\mathbf{\Gamma} + \sigma^2 \mathbf{\Lambda}^{-2})^{-1}\mathbf{y}$. It can be shown that the resulting algorithms differ from (2.10) and (2.11) only in the scaling factor. Specifically, we have

$$\hat{x}(t+1) = \frac{\sqrt{q_i}}{q_i + \sigma^2}\left(y_i - \sum_{j=1}^{i-1}\sqrt{q_j}\Gamma_{ji}b_j(t+1) - \sum_{j=i+1}^{N}\sqrt{q_j}\Gamma_{ji}b_j(t)\right) \tag{2.21}$$

$$\hat{x}(t+1) = \frac{\sqrt{q_i}}{q_i + \sigma^2}\left(y_i - \sum_{j=1, j\neq i}^{N}\sqrt{q_j}\Gamma_{ji}b_j(t)\right) \tag{2.22}$$

for the Gauss-Seidel and the Jacobi iterations respectively. Further, these iterations can be followed by the $[-1, 1]$ clippers. The resulting receivers converge to the minimizer of the following quadratic program which is similar to (2.7) where the matrix $\mathbf{R}$ is replaced by $\mathbf{R} + \sigma^2\mathbf{I}$.

$$\min_{\mathbf{b}\in[-1,1]^N} \mathbf{b}^\top(\mathbf{R} + \sigma^2\mathbf{I})\mathbf{b} - 2\mathbf{b}^\top\mathbf{\Lambda}\mathbf{y} \tag{2.23}$$

Note that this minimization problem when defined on the corners of the unit hypercube, i.e. $\mathbf{b} \in \{-1, 1\}^N$, is exactly the same as the original OMUD problem in (2.4), since on the corners of the unit hypercube the additional term $\sigma^2\mathbf{b}^\top\mathbf{b}$ in the cost function is independent of the choice of $\mathbf{b}$. An equivalent of the convergence condition (2.16) can be easily derived for the MMSE type receivers with clippers using the same steps we followed in this section. The algorithms without the clippers, i.e. the algorithms that converge to the decorrelating or the MMSE detectors, correspond to linear Gauss-Seidel and Jacobi iterations. For these iterations, Theorem 2 still gives the sufficient condition for convergence. However, less stringent conditions which are necessary and sufficient, are available for this case. For the decorrelator, for example, using Theorem 3.3 of Section 4.3 in [93], one can show that the necessary and sufficient condition for convergence of the Jacobi algorithm without the clippers is

$$\rho(\tilde{\mathbf{\Gamma}}) < 1 \tag{2.24}$$

where $\tilde{\mathbf{\Gamma}}$ is the matrix consisting of the off-diagonal elements of $\mathbf{\Gamma}$, i.e. $\tilde{\Gamma}_{ij} = \Gamma_{ij}$, if $i \neq j$, and $\tilde{\Gamma}_{ii} = 0$. For the MMSE detector, the necessary and sufficient condition is found by replacing $\mathbf{\Gamma}$ with $\mathbf{\Gamma} + \sigma^2 \mathbf{\Lambda}^{-2}$. To see why (2.24) is less stringent than (2.16), first observe that $\bar{\mathbf{\Gamma}} = |\tilde{\mathbf{\Gamma}}|$, and by definition of the $\|\cdot\|_\infty^w$ norm of a matrix, $\|\tilde{\mathbf{\Gamma}}\|_\infty^w = \|\bar{\mathbf{\Gamma}}\|_\infty^w$. We can then use Proposition 6.2(e) of Section 2.6 in [13] which states that $\rho(\mathbf{A}) \leq \|\mathbf{A}\|_\infty^w$ for any matrix $\mathbf{A}$. Thus, the following relationship holds

$$\rho(\tilde{\mathbf{\Gamma}}) \leq \|\bar{\mathbf{\Gamma}}\|_\infty^w \tag{2.25}$$

This means that if (2.20) (or equivalently (2.16)) holds, then (2.24) is guaranteed to hold.

Finally, we should emphasize that the implementations discussed here are not the unique way of solving for the minimizer of (2.7). There are other nonlinear programming methods that yield iterative algorithms whose bit error rate performance matches that of the soft interference cancellers.

2.5 Generalized MMSE Detector

The constraint on each $b_i \in \{-1, 1\}$ is equivalent to $b_i^2 = 1$ which implies $\mathbf{b}^\top \mathbf{b} = N$ at any feasible point for OMUD. Relaxing this set to $\mathbf{b}^\top \mathbf{b} \leq N$, i.e. to the smallest sphere that contains the corners of the unit hypercube, results in the following optimization problem:

$$\min_{\mathbf{b}^\top \mathbf{b} \leq N} \mathbf{b}^\top \mathbf{R} \mathbf{b} - 2\mathbf{b}^\top \mathbf{\Lambda} \mathbf{y} \tag{2.26}$$

Since (2.26) minimizes a convex function over a convex set, it has a unique minimum and iterative algorithms such as gradient descent can be employed to find this minimum [43]. Further, the *convex duality theorem* [43, Theorem 14.6] ensures that no duality gap exists and one can solve for the dual problem instead. Since (2.26) has a single constraint, there is only one dual variable. Thus, a simpler iterative algorithm can be found by solving the dual problem as outlined below.

The Lagrangian dual function can be expressed as

$$\mathcal{L}(\mathbf{b}, \lambda) = \mathbf{b}^\top \mathbf{R}\mathbf{b} - 2\mathbf{b}^\top \mathbf{\Lambda}\mathbf{y} + \lambda(\mathbf{b}^\top \mathbf{b} - N) \tag{2.27}$$

which is to be maximized over $\mathbf{b}$ and $\lambda \geq 0$. Solving for $\mathbf{b}$ in terms of λ and substituting back we arrive at:

$$\max_{\lambda \geq 0} \quad -\mathbf{y}^\top \mathbf{\Lambda}(\mathbf{R} + \lambda \mathbf{I})^{-1} \mathbf{\Lambda}\mathbf{y} - \lambda N \tag{2.28}$$

which is a one-dimensional optimization problem and can be solved with a variety of iterative algorithms [43]. A simple unconstrained gradient descent algorithm is guaranteed to converge for a small enough step size μ which can then be projected onto the positive axis. The algorithm is

$$\bar{\lambda}(t+1) = \bar{\lambda}(t) + \mu(\mathbf{y}^\top \mathbf{\Lambda}(\mathbf{R} + \bar{\lambda}(t)\mathbf{I})^{-2} \mathbf{\Lambda}\mathbf{y} - N) \tag{2.29}$$

which converges to $\bar{\lambda}$. The maximizer of (2.28) is given by

$$\lambda^* = \max(0, \bar{\lambda}) \tag{2.30}$$

Then, the unique minimizer of (2.26) can be found to be

$$\mathbf{b}^* = (\mathbf{R} + \lambda^* \mathbf{I})^{-1} \mathbf{\Lambda}\mathbf{y} = \mathbf{\Lambda}^{-1}(\mathbf{\Gamma} + \lambda^* \mathbf{\Lambda}^{-2})^{-1} \mathbf{y} \tag{2.31}$$

The form of this solution whose sign is the estimate of the bit vector is also familiar because of its similarity to the *MMSE detector* [36]. Thus, we term the relaxation (2.26) the *generalized MMSE* (GMMSE) solution. When $\lambda^* = \sigma^2$, (2.31) reduces to the MMSE detector. Note that, since λ^* is a function of $\mathbf{y}$, the GMMSE solution results in a nonlinear multiuser detector in contrast to the MMSE detector. On the other hand, the knowledge of the noise power value (σ^2) is not necessary for the GMMSE detector whereas the MMSE detector requires this knowledge if training or blind adaptation is not desired [29, 36].

The GMMSE detector is also an iterative detector since λ^* has to be found iteratively. However, since the iterations are in one dimension, they are expected to converge quickly compared to multidimensional algorithms.

2.6 Numerical Results

Since the probability of bit error expressions are not analytically tractable for arbitrary number of users and iterations, we have simulated the bit error performance of the detectors investigated in this work. All iterative detectors (multistage soft cancellers and the GMMSE) are evaluated at their convergence points. The first system simulated is a $N = 7$ user system with processing gain $G = 7$ that uses m-sequences, i.e. $\Gamma_{ij} = -1/7\ j \neq i$. Figure 2.2 shows the probability of bit error for one user when that user has 0 dB SNR and all the interferers have a common SNR that is varied. The soft interference cancellers ((2.10), (2.11), (2.12)) have almost invariable performance versus interference strength. We note that the performance of the GMMSE detector is similar to that of the linear MMSE detector. In particular, we observe that the GMMSE detector has the same trend of approaching the decorrelator performance as the MMSE detector when the interference dominates the noise. We have also simulated an $N = 4$ user system with processing gain $G = 7$ that uses Gold sequences with cross correlation matrix:

$$\mathbf{\Gamma} = \frac{1}{7}\begin{bmatrix} 7 & -1 & 3 & 3 \\ -1 & 7 & 3 & -1 \\ 3 & 3 & 7 & -1 \\ 3 & -1 & -1 & 7 \end{bmatrix} \tag{2.32}$$

and observed similar results (Figure 2.3). Notice that the diagonal dominance condition is not satisfied. However, the parallel soft interference canceller is still guaranteed to converge because $\rho(\bar{\mathbf{\Gamma}}) = 0.8747 < 1$.

2.7 Chapter Summary

In this chapter, we have shown that many popular suboptimum detectors are devices that attempt to approximate the solution of the joint minimum bit error rate detector (OMUD). Although it is analytically hard to characterize exactly how closely they approximate the OMUD cost function, we have observed that they achieve near-optimum cost values. Consequently, the near-optimum bit error rate performances of these detectors are not surprising. The nonlinear programming approach helped us identify the convergence conditions of multistage soft interference cancellers. We have also proposed and devised a simple iterative nonlinear detector with similar performance to the MMSE detector. It can be used in scenarios where adaptive or blind adaptive detection is not suitable, for instance when the channel is changing rapidly, and the ambient noise power is unknown.

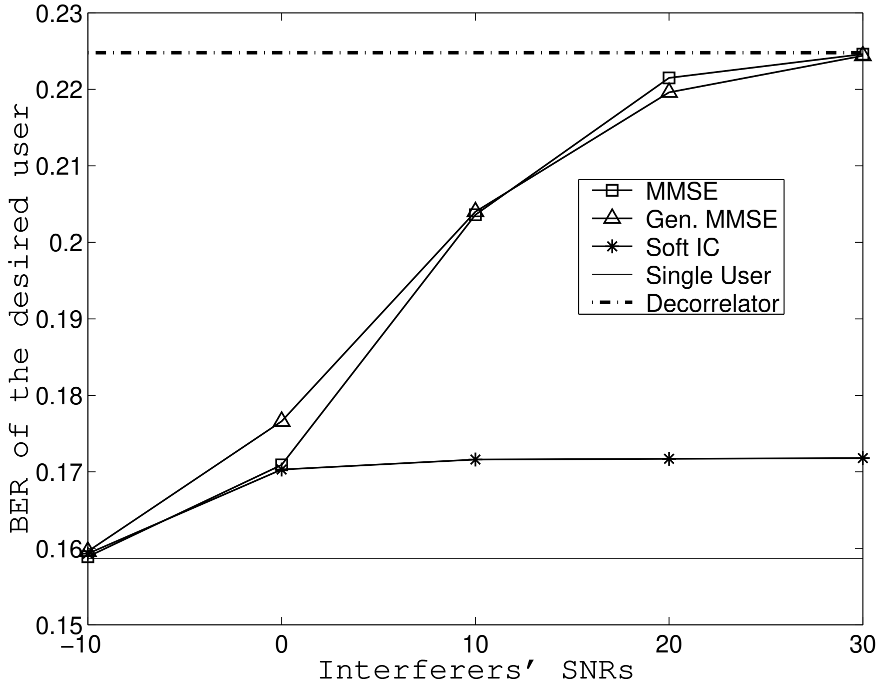

Figure 2.2: Comparison of error probabilities of near-optimum multiuser detectors: Near-far scenario (Desired user at 0 dB SNR). $G = 7$, $N = 7$, M-sequences.

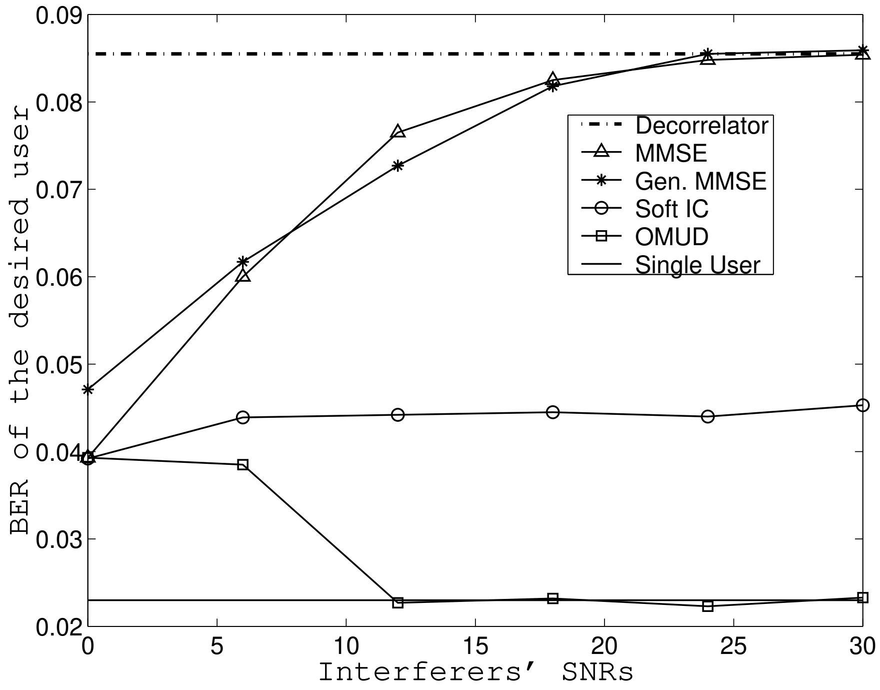

Figure 2.3: Comparison of error probabilities of near-optimum multiuser detectors: Near-far scenario (Desired user at 6 dB SNR). $G = 7$, $N = 4$, Gold sequences.

Chapter 3

Multiuser Detection and Beamforming for CDMA: Filter Structures

3.1 Introduction

In this chapter, we will concentrate on the two commonly used receiver processing based interference management methods: multiuser detection and receiver beamforming. Both methods aim at suppressing or cancelling the interference using receiver signal processing; multiuser detection, as we have seen in chapter 2, exploits the temporal structure of the interference, whereas receiver beamforming exploits the spatial structure of the interference for interference management.

Increasing the capacity of CDMA systems by employing antenna arrays at the base station has been proposed in [42], where the outputs of the multiple antenna array elements are combined to make bit decisions for the user. Matched filter receivers are assumed in the temporal domain for each user and the array observations are combined via a filter that is matched to the array response of the user, i.e., single user processing is employed in both domains.

Space-time processing for CDMA which traditionally refers to receiver beamforming (space processing) and multipath combining (time processing) has also been proposed as a capacity enhancement technique and studied extensively, see [50] for a survey. The received signals from different paths and antennas are combined to better decode the desired user's bits. However, the inherent structure of the multiple access interference is not exploited, i.e., no multiuser detection is employed [10,31,85]. A recent paper [81] addresses the derivation of the sufficient statistics and the minimum bit error rate receiver and some linear suboptimum multiuser detectors when an antenna array is present at the receiver for a multipath channel.

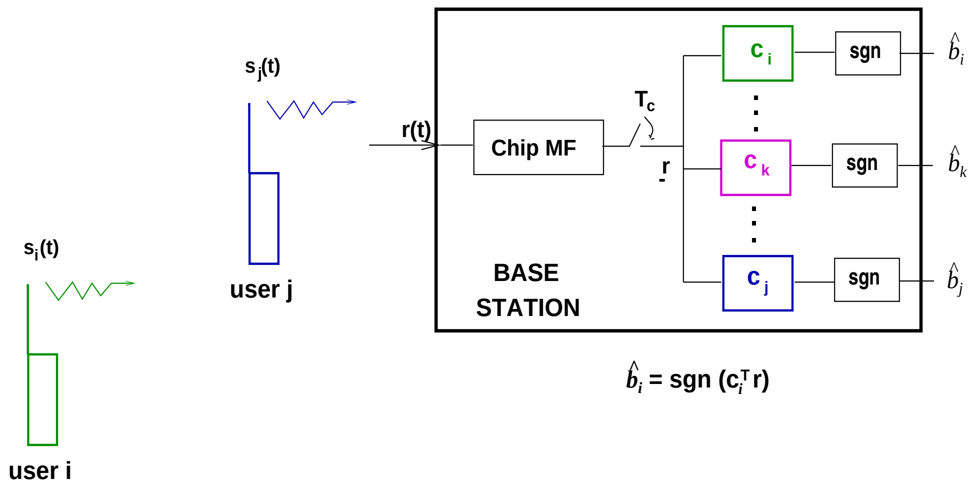

Figure 3.1: Linear multiuser detection employs intelligent temporal filters

In this chapter, we will investigate the possible receiver filter structures when both multiuser detection and beamforming are employed to further increase the uplink capacity of a CDMA system. The intuition behind this is to combine the two methods which work on similar principles in different domains and investigate the advantages of designing *temporal-spatial* filters. Linear processing is assumed in both the temporal and the spatial domains. When linear multiuser detection is employed at a single antenna base station receiver, a different linear, temporal vector filter, i.e., a *chip combiner*, is used to obtain the decision statistic for each user, after appropriately sampling the received signal (Figure 3.1). When linear receiver beamforming is employed at a base station receiver, a different linear spatial vector filter, i.e., an *antenna combiner* is used to obtain the decision statistic for each user, at the output of the array (Figure 3.2). When both are employed, the resulting temporal-spatial filters are denoted by two-dimensional matrix filters (Figure 3.3). Within this framework, we formulate and examine the performance of several filter structures all of which are based on minimum mean squared error criteria, but differ in how the MMSE problems are defined in the temporal and spatial domains, i.e., jointly or in cascade.

We give the jointly optimal temporal and spatial filter that minimizes the mean squared error between the information bit and filter output of a desired user. Since this joint MMSE filter may have high computational complexity, less complex filters which

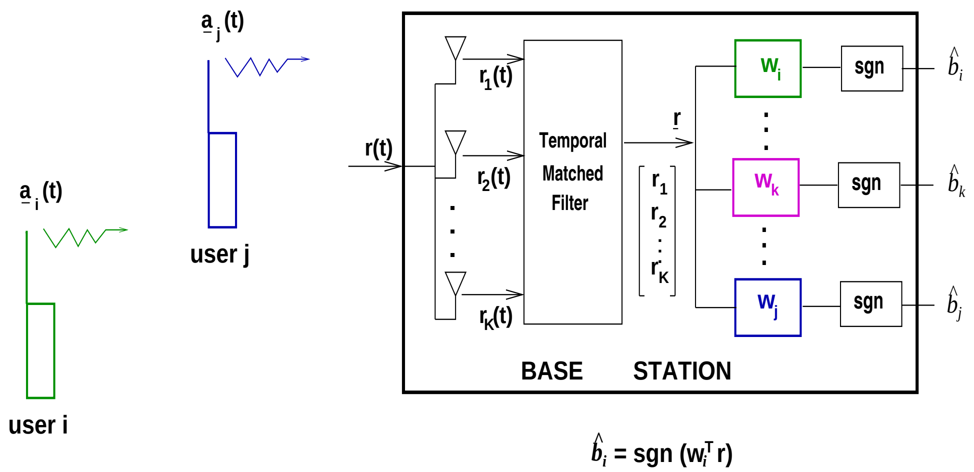

Figure 3.2: Linear receiver beamforming employs intelligent filters in spatial domain

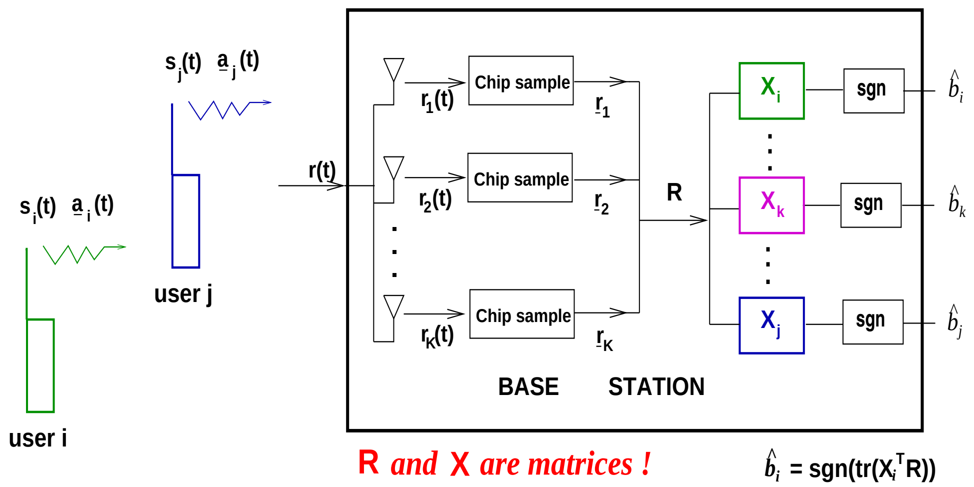

Figure 3.3: Combined linear multiuser detection and receiver beamforming employs intelligent filters in both temporal and spatial domains

nevertheless provide efficient interference suppression are of interest. To serve this purpose, we propose constrained optimum filters are found by forcing the joint receiver filter matrix to be of rank 1 and finding the optimum filter in this constrained space. We also derive cascaded filters where mean squared error optimization is carried over in temporal and spatial domains independently in tandem, in both spatial-temporal and temporal-spatial orders. While the cascaded spatial-temporal filter is similar to the idea of cascade optimum-space/optimum-time combiner proposed in [10], the temporal combiner in our case is a chip combiner that exploits the temporal structure of the interference composed of the temporal signatures of the interferers, as opposed to a multipath combiner in [10] which is a single user temporal processor. We consider all above filter structures and then give analytical performance comparisons among them. While the joint domain MMSE filter is clearly the signal-to-interference Ratio (SIR) maximizing temporal and spatial processor over all matrix filters, an interesting observation is that the constrained optimum temporal spatial processor outperforms all combined single user/multiple user approaches and the cascaded approaches either uniformly or asymptotically. Numerical results supporting the performance analysis is also given.

3.2 System Model

We consider a single cell, N user, DS-CDMA system with processing gain G. For clarity of exposition, we assume a synchronous system. At the base station, an antenna array of K elements is employed. Following references [42, 56, 64], over one bit period, the received signal at the output of the antenna array at the base station is:

$$\mathbf{r}(t) = \sum_{j=1}^{N} \sqrt{p_j h_j} b_j s_j(t) \mathbf{a}_j + \mathbf{n}(t) \tag{3.1}$$

where p_j,b_j and $s_j(t)$ are the transmit power, bit and the signature of user j, respectively. h_j and $\mathbf{a}_j$ are the uplink gain and the array response vector, i.e., the spatial signature, of user j . We assume that both the temporal and spatial signatures of the users are of unit energy, i.e., $\int_0^T s_j(t)^2 dt = \mathbf{a}_j^H \mathbf{a}_j = 1$ where T is the bit duration. We further assume that the temporal signatures are of the following form

$$s_j(t) = \sum_{l=1}^{G} s_j^{(l)} \psi(t-(l-1)T_c) \quad \text{with} \quad s_j^{(l)} = \pm 1/\sqrt{G} \tag{3.2}$$

where $\psi(t)$ is the chip waveform and T_c is the chip duration. Chip matched filtering and sampling the received signal $\mathbf{r}(t)$ at the lth chip interval, we obtain a K-dimensional observation vector $\mathbf{r}_l$. The vector $\mathbf{r}_l$ represents the chip matched filtered samples at the lth chip interval across the K antenna array elements and is given by

$$\mathbf{r}_l = \sum_{j=1}^{N} \sqrt{p_j h_j} b_j \mathbf{a}_j s_j^{(l)} + \mathbf{n}_l \tag{3.3}$$

Over one bit period $T = GT_c$, we collect a set of G K-dimensional vectors $\{\mathbf{r}_l, 1 \leq l \leq G\}$ which we can arrange in a $G \times K$ matrix $\mathbf{R} = [\mathbf{r}_1, \mathbf{r}_2, \cdots, \mathbf{r}_G]^\top$:

$$\mathbf{R} = \sum_{j=1}^{N} \sqrt{p_j h_j} b_j \mathbf{s}_j \mathbf{a}_j^\top + \mathbf{N} \tag{3.4}$$

These KG observation samples can also be seen as a collection of K G-dimensional vectors, i.e., $\mathbf{R} = [\mathbf{z}_1, \mathbf{z}_2, \cdots, \mathbf{z}_K]$, where $\mathbf{z}_k$ denotes the observation vector consisting of G chip matched filtered samples at the output of the kth antenna element and is expressed as

$$\mathbf{z}_k = \sum_{j=1}^{N} \sqrt{p_j h_j} b_j \mathbf{s}_j a_j^{(k)} + \mathbf{n}_k \tag{3.5}$$

In (3.4), $\mathbf{N}$ is the matrix that represents the spatially and temporally white noise, i.e., $E[N_{kl}^* N_{mn}] = \sigma^2 \delta_{km} \delta_{ln}$, where $(\cdot)^*$ denotes the conjugate of a complex number. We label user i as the *desired user* and the other users as *interferers.* The rest of this chapter deals with designing temporal-spatial filters for user i.

3.3 Filter Structures

The detection of the information bit of the desired user is done by taking the sign of the real part of the decision statistic which is found by combining the entries of the observation matrix $\mathbf{R}$ by using a matrix filter $\mathbf{X}_i$. Thus, the decision statistic, y_i, is the output of a two dimensional linear filter $\mathbf{X}_i$:

$$y_i = \sum_{j=1}^{G} \sum_{l=1}^{K} [X_i]_{jl}^* R_{jl} = \operatorname{tr}(\mathbf{X}_i^H \mathbf{R}) \tag{3.6}$$

where $\mathrm{tr}(\cdot)$ and $(\cdot)^H$ are the trace and the hermitian transpose operations on a matrix, respectively. The representation of the temporal-spatial filter as a matrix is useful in expressing any filtering scheme that is linear in the temporal and spatial domains.

In what follows, we investigate the possible filter structures. The filter structures in Sections 3.3.1, 3.3.2, and 3.3.3 use single user processing in at least one of the temporal and spatial domains and are well-known. Next we derive two temporal-spatial filters, the filter structures in Sections 3.3.4 and 3.3.5, which use *cascaded* MMSE optimizations in spatial and temporal domains. The structure of these receivers is to combine either the received chip samples at the output of each array element in the MMSE sense followed by a spatial filter that combines the resulting vector in the MMSE sense (Section 3.3.4); or to combine all array outputs for each chip sample in the MMSE sense followed by the temporal MMSE combiner for the resulting vector (Section 3.3.5).

Next, in Sections 3.3.6 and 3.3.7, we give the joint temporal-spatial MMSE filter structures. The difference between the two joint MMSE structures is the fact that while the joint optimum temporal-spatial MMSE in Section 3.3.6 is the best filter in terms of minimizing the mean squared error over all possible matrix filters, the constrained optimum temporal-spatial filter in Section 3.3.7 is the MMSE filter when the filter space is constrained to contain matrix filters of rank 1 only. The physical interpretation of this mathematical constraint on the matrix receiver filter is that it results in a *separable filter* with a single temporal and a single spatial combiner.

The difference between the joint MMSE structures in Sections 3.3.6 and 3.3.7 and the cascaded MMSE structures in Sections 3.3.4 and 3.3.5 lies in the fact that the cascaded structures use temporal and spatial filters that are optimized independently in each domain while the joint structures are found by optimization in both domains simultaneously.

3.3.1 Single User Temporal-Spatial Detector

This is a single user based approach for both the spatial and the temporal domains and has been proposed in [42]. The decision statistic in this case is $y_i = \mathbf{s}_i^\top \mathbf{R} \mathbf{a}_i^* = \mathrm{tr}(\mathbf{a}_i^* \mathbf{s}_i^\top \mathbf{R})$

leading to

$$\mathbf{X}_i^{\text{MF-MF}} = \mathbf{s}_i \mathbf{a}_i^\top \tag{3.7}$$

3.3.2 Temporal MMSE Filter-Spatial Matched Filter

This approach uses multiuser processing in temporal domain [36] and single user processing in spatial domain. The decision statistic in this case is $y_i = \mathbf{c}_i^\top \mathbf{R} \mathbf{a}_i^* = \text{tr}(\mathbf{a}_i^* \mathbf{c}_i^\top \mathbf{R})$ leading to

$$\mathbf{X}_i^{\text{MMSE-MF}} = \mathbf{c}_i \mathbf{a}_i^\top \tag{3.8}$$

where

$$\mathbf{c}_i = \sqrt{p_i h_i} \left(\sum_{j=1}^{N} p_j h_j |\mathbf{a}_i^H \mathbf{a}_j|^2 \mathbf{s}_j \mathbf{s}_j^\top + \sigma^2 \mathbf{I} \right)^{-1} \mathbf{s}_i \tag{3.9}$$

3.3.3 Temporal Matched Filter-Spatial MMSE Filter

This approach uses single user processing in the temporal domain combined with multi user processing in spatial domain [41]. The decision statistic in this case is $y_i = \mathbf{s}_i^\top \mathbf{R} \mathbf{w}_i^* = \text{tr}(\mathbf{w}_i^* \mathbf{s}_i^\top \mathbf{R})$ leading to

$$\mathbf{X}_i^{\text{MF-MMSE}} = \mathbf{s}_i \mathbf{w}_i^\top \tag{3.10}$$

where

$$\mathbf{w}_i = \sqrt{p_i h_i} \left(\sum_{j=1}^{N} p_j h_j (\mathbf{s}_i^\top \mathbf{s}_j)^2 \mathbf{a}_j \mathbf{a}_j^H + \sigma^2 \mathbf{I} \right)^{-1} \mathbf{a}_i \tag{3.11}$$

3.3.4 Cascaded Temporal-Spatial MMSE

Assume that at the output of each antenna array we are allowed to design a temporal filter, i.e., a linear chip combiner. Recall that the output of the kth element of the antenna array corresponds to the kth column, $\mathbf{z}_k$, of the received signal matrix $\mathbf{R}$ and is given (3.5).

We would like to design K temporal filters $\mathbf{c}_k, k = 1, \ldots, K$, such that the resulting statistic

$$\tilde{y}_k = \mathbf{c}_k^H \mathbf{z}_k = \sum_{j=1}^{N} \sqrt{p_j h_j} b_j \mathbf{c}_k^H \mathbf{s}_j a_j^{(k)} + \mathbf{c}_k^H \mathbf{n}_k \tag{3.12}$$

has minimum mean squared difference from the desired bit, b_i. The solution can be found as [36]:

$$\mathbf{c}_k = \arg\min_{\mathbf{c}} E[|\mathbf{c}^H \mathbf{z}_k - b_i|^2] = \sqrt{p_i h_i} a_i^{(k)} \left(\sum_{j=1}^{N} |a_j^{(k)}|^2 p_j h_j \mathbf{s}_j \mathbf{s}_j^\top + \sigma^2 \mathbf{I} \right)^{-1} \mathbf{s}_i \tag{3.13}$$

Note that what make $\mathbf{c}_k$, the temporal MMSE filter at the output of the kth antenna, different from $\mathbf{c}_l$, the temporal MMSE filter at the output of the lth antenna, are the different spatial gains users have for different antennas. Defining the modified gain at the output of the kth antenna for user j as $(\mathbf{c}_k^H \mathbf{s}_j) a_j^{(k)}$, from (3.12) we have

$$\tilde{y}_k = \sum_{j=1}^{N} \sqrt{p_j h_j} b_j \tilde{a}_j^{(k)} + \tilde{n}_k, \qquad k = 1, ..., K \tag{3.14}$$

with

$$\tilde{a}_j^{(k)} = \left(\sqrt{p_i h_i} \left(a_i^{(k)} \right)^* \mathbf{s}_i^\top \left(\sum_{j'=1}^{N} |a_{j'}^{(k)}|^2 p_j h_{j'} \mathbf{s}_{j'} \mathbf{s}_{j'}^\top + \sigma^2 \mathbf{I} \right)^{-1} \mathbf{s}_j \right) a_j^{(k)} \tag{3.15}$$

We can then combine $\{\tilde{y}_k\}$ in the MMSE sense and similar to (3.13) we can express the second stage of the cascaded filter as

$$\mathbf{w} = \arg\min_{\mathbf{w}} E[|\mathbf{w}^H \tilde{\mathbf{y}} - b_i|^2] = \sqrt{p_i h_i} \left(\sum_{j=1}^{N} p_j h_j \tilde{\mathbf{a}}_j \tilde{\mathbf{a}}_j^H + \mathbf{\Lambda} \right)^{-1} \tilde{\mathbf{a}}_i \tag{3.16}$$

where $\mathbf{\Lambda} = \mathrm{diag}\{\sigma^2 (\mathbf{c}_k^H \mathbf{c}_k)\}$ is the covariance matrix of $\tilde{\mathbf{n}}$. The final bit decision is done by taking the sign of the real part of $y_i = \mathbf{w}^H \tilde{\mathbf{y}}$.

Note that to construct the overall receiver we need to invert K $G \times G$ matrices and one $K \times K$ matrix. To see how the overall cascaded filter can be expressed as matrix filter $\mathbf{X}_i^{\text{TS-CMMSE}}$, observe that

$$\tilde{y}_k = \mathbf{c}_k^H \mathbf{z}_k = \mathbf{c}_k^H \mathbf{R} \mathbf{e}_k = \mathrm{tr}(\mathbf{e}_k \mathbf{c}_k^H \mathbf{R}) \tag{3.17}$$

and

$$y_i = \sum_{k=1}^{K} w_k^* \tilde{y}_k = \sum_{k=1}^{K} w_k^* \mathrm{tr}(\mathbf{e}_k \mathbf{c}_k^H \mathbf{R}) = \mathrm{tr}\left(\sum_{k=1}^{K} w_k^* \mathbf{e}_k \mathbf{c}_k^H \mathbf{R}\right) \tag{3.18}$$

where $\mathbf{e}_k$ is a K vector which has 1 in its kth entry and all zeros elsewhere. Comparing (3.18) with (3.6), we find that

$$\mathbf{X}_i^{\text{TS-CMMSE}} = \sum_{k=1}^{K} w_k \mathbf{c}_k \mathbf{e}_k^\top \tag{3.19}$$

which can be of rank up to K.

3.3.5 Cascaded Spatial-Temporal MMSE

Alternatively, one can think of first combining all antenna array elements for each chip sample, followed by a temporal combiner. In this case, first, each row of the received matrix, $\mathbf{r}_l$ from (3.3), is combined via a spatial filter $\mathbf{w}_l, l = 1, \ldots, G$.

We would like to find the filter $\mathbf{w}_l$ to combine the elements of $\mathbf{r}_l$ in the MMSE sense. The resulting filter can be expressed as

$$\mathbf{w}_l = \arg\min_{\mathbf{w}} E[|\mathbf{w}^H \mathbf{r}_l - b_i|^2] = \sqrt{p_i h_i} s_i^{(l)} \left(\sum_{j=1}^{N} (s_j^{(l)})^2 p_j h_j \mathbf{a}_j \mathbf{a}_j^H + \sigma^2 \mathbf{I}\right)^{-1} \mathbf{a}_i \tag{3.20}$$

Recall from (3.2) that $(s_j^{(l)})^2 = 1/G$, for all l. Thus, defining

$$\hat{\mathbf{w}} = \sqrt{p_i h_i} \left(\sum_{j=1}^{N} \frac{1}{G} p_j h_j \mathbf{a}_j \mathbf{a}_j^H + \sigma^2 \mathbf{I}\right)^{-1} \mathbf{a}_i \tag{3.21}$$

we arrive at

$$\mathbf{w}_l = s_i^{(l)} \hat{\mathbf{w}} \qquad l = 1, \ldots, G \tag{3.22}$$

Thus, for each chip sample sequence $\mathbf{w}_l$ is equal to $\pm\frac{1}{\sqrt{G}}\hat{\mathbf{w}}$ depending on the current chip value of the desired user. At the output of the lth combiner, the resulting statistic can be expressed as

$$\hat{y}_l = \mathbf{w}_l^H \mathbf{r}_l = \sum_{j=1}^{N} \sqrt{p_j h_j} b_j \mathbf{w}_l^H \mathbf{a}_j s_j^{(l)} + \mathbf{w}_l^H \mathbf{n}_l \tag{3.23}$$

$$= \sum_{j=1}^{N} \sqrt{p_j h_j} b_j \hat{\mathbf{w}}^H \mathbf{a}_j s_i^{(l)} s_j^{(l)} + \hat{n}_l \tag{3.24}$$

Defining

$$\hat{s}_j^{(l)} = \hat{\mathbf{w}}^H \mathbf{a}_j s_i^{(l)} s_j^{(l)} \tag{3.25}$$

we have

$$\hat{y}_l = \sum_{j=1}^{N} \sqrt{p_j h_j} b_j \hat{s}_j^{(l)} + \hat{n}_l, \qquad l = 1, \dots, G \tag{3.26}$$

It remains to find the MMSE combiner for $\hat{\mathbf{y}}$. We can express this second stage of the cascaded filter as

$$\mathbf{c} = \arg\min_{\mathbf{c}} E[|\mathbf{c}^H \hat{\mathbf{y}} - b_i|^2] = \sqrt{p_i h_i} \left(\sum_{j=1}^{N} p_j h_j \hat{\mathbf{s}}_j \hat{\mathbf{s}}_j^H + \frac{1}{G} \sigma^2 (\hat{\mathbf{w}}^H \hat{\mathbf{w}}) \mathbf{I} \right)^{-1} \hat{\mathbf{s}}_i \tag{3.27}$$

Notice that, as in the case of Section 3.3.4, the noise covariance matrix is given by $\mathrm{diag}\{\sigma^2(\mathbf{w}_l^T \mathbf{w}_l)\}$. Since $\mathbf{w}_l^H \mathbf{w}_l = (1/G)(\hat{\mathbf{w}}^H \hat{\mathbf{w}})$ for all l, the noise covariance matrix reduces to $(\sigma^2/G)(\hat{\mathbf{w}}^H \hat{\mathbf{w}})\mathbf{I}$. The final bit decision is done by taking the sign of the real part of $y_i = \mathbf{c}^H \hat{\mathbf{y}}$.

Note that to construct the overall receiver we need to invert one $K \times K$ matrix to calculate $\hat{\mathbf{w}}$ and one $G \times G$ matrix to calculate $\mathbf{c}$. To see how the overall cascaded filter can be expressed as a matrix filter $\mathbf{X}_i^{\text{ST-CMMSE}}$, observe that

$$\hat{y}_l = \mathbf{e}_l^\top \mathbf{R} \mathbf{w}_l^* = s_i^{(l)} \mathrm{tr}(\hat{\mathbf{w}}^* \mathbf{e}_l^\top \mathbf{R}) \tag{3.28}$$

and

$$y_i = \mathbf{c}^H \hat{\mathbf{y}} = \sum_{l=1}^{G} c_l^* \hat{y}_l = \sum_{l=1}^{G} c_l^* s_i^{(l)} \mathrm{tr}(\hat{\mathbf{w}}^* \mathbf{e}_l^\top \mathbf{R}) = \mathrm{tr}(\hat{\mathbf{w}}^* (\sum_{l=1}^{G} c_l^* s_i^{(l)} \mathbf{e}_l^\top) \mathbf{R}) \tag{3.29}$$

Then, comparing (3.29) with (3.6), we find that

$$\mathbf{X}_i^{\text{ST-CMMSE}} = \left(\sum_{l=1}^{G} c_l s_i^{(l)} \mathbf{e}_l \right) \hat{\mathbf{w}}^\top \tag{3.30}$$

3.3.6 Optimum Temporal-Spatial MMSE Filter

Let us now consider the jointly optimum matrix filter in temporal and spatial domains which minimizes the minimum mean squared error between y_i and b_i. That is, we want to find a matrix filter $\bar{\mathbf{X}}_i$ such that

$$\mathbf{X}_i^{\text{O-MMSE}} = \arg\min_{\mathbf{X}} E\left[\left| \mathrm{tr}(\mathbf{X}^H \mathbf{R}) - b_i \right|^2 \right] \tag{3.31}$$

The optimization problem (3.31) can be converted to an optimization problem with vector variables for easier manipulation [64]. Let $\mathbf{r}$ be the long vector obtained by stacking the columns of the received signal matrix, $\mathbf{R}$. The MMSE problem then can be reformulated as follows:

$$\bar{\mathbf{x}}_i = \arg\min_{\mathbf{x}} E\left[\left|\mathbf{x}^H \mathbf{r}_i - b_i\right|^2\right] \tag{3.32}$$

Let us define $\mathbf{q}_j$ as the combined temporal-spatial signature of user j. It is constructed by stacking columns of $\mathbf{s}_j \mathbf{a}_j^\top$ as a long vector of size KG. Then, the solution to the optimization problem (3.32) is given as [41]:

$$\begin{aligned} \bar{\mathbf{x}}_i &= k_i \left(\sum_{j \neq i} p_j h_j \mathbf{q}_j \mathbf{q}_j^H + \sigma^2 \mathbf{I} \right)^{-1} \mathbf{q}_i && (3.33) \\ &= l_i \left(\sum_{j=1}^{N} p_j h_j \mathbf{q}_j \mathbf{q}_j^H + \sigma^2 \mathbf{I} \right)^{-1} \mathbf{q}_i && (3.34) \end{aligned}$$

where (3.34) follows from (3.33) using the matrix inversion lemma, which states for an invertible matrix $\mathbf{M}$ and vectors $\mathbf{u}$ and $\mathbf{v}$

$$\left(\mathbf{M} + \mathbf{u}\mathbf{v}^\top\right)^{-1} = \mathbf{M}^{-1} - \frac{\mathbf{M}^{-1}\mathbf{u}\mathbf{v}^\top\mathbf{M}^{-1}}{1 + \mathbf{v}^\top\mathbf{M}^{-1}\mathbf{u}} \tag{3.35}$$

The constants k_i and l_i in (3.33) and (3.34) are given as

$$k_i = \frac{\sqrt{p_i h_i}}{1 + p_i h_i \mathbf{q}_i^H (\sum_{j \neq i} \mathbf{q}_j \mathbf{q}_j^H + \sigma^2 \mathbf{I})^{-1} \mathbf{q}_i} \qquad l_i = \sqrt{p_i h_i} \tag{3.36}$$

Note that the matrix $\sum_{j \neq i} \mathbf{q}_j \mathbf{q}_j^H + \sigma^2 \mathbf{I}$ is necessarily positive definite (and thus has an inverse) for all $\sigma^2 > 0$. $\bar{\mathbf{X}}_i$ then can be constructed by taking every G element of $\bar{\mathbf{x}}_i$ and putting as a column to $\bar{\mathbf{X}}_i$. Note also that, it is possible to use adaptive or blind adaptive approaches to find $\bar{\mathbf{x}}_i$ [29, 36, 64].

3.3.7 Constrained Optimum Temporal-Spatial Filtering

The joint MMSE filter requires a possibly large matrix ($KG \times KG$) to be inverted which can be computationally costly, or the corresponding adaptive implementation may be slow. Thus, one might want to consider less complex joint MMSE filters that nevertheless present capacity improvements for the system.

To this end, we consider a constrained class of rank 1 matrix filters, i.e., $\mathbf{X}_i \in \mathcal{L}$ where $\mathcal{L}$ is the space of rank 1 matrices in $\mathcal{C}^{G\times K}$. Note that all $\mathbf{X}_i \in \mathcal{L}$ can be expressed as $\mathbf{X}_i = \mathbf{c}_i\mathbf{w}_i^\top$. Physically, the scheme is to combine the chip matched filter outputs using a linear filter at the output of each of the antenna (or equivalently linearly combining the all antenna array observations for each chip) followed by a linear combination of the resulting statistics. The decision statistic to decode the bit for user i then becomes

$$y_i = \mathrm{tr}(\mathbf{w}_i^*\mathbf{c}_i^H\mathbf{R}) = \mathbf{c}_i^H\mathbf{R}\mathbf{w}_i^* = \sum_{j=1}^{N}\sqrt{p_jh_j}b_j(\mathbf{c}_i^H\mathbf{s}_j)(\mathbf{w}_i^H\mathbf{a}_j) + \mathbf{c}_i^H\mathbf{N}\mathbf{w}_i^* \tag{3.37}$$

We consider the joint optimal filter pair in the minimum mean squared error sense. In this case, the optimization problem (3.31) becomes

$$\begin{aligned}[\bar{\mathbf{c}}_i, \bar{\mathbf{w}}_i] &= \arg\min_{\mathbf{c}_i,\mathbf{w}_i} E\left[\left|\mathrm{tr}(\mathbf{w}_i^*\mathbf{c}_i^H\mathbf{R}) - b_i\right|^2\right] \\ &= \arg\min_{\mathbf{c}_i,\mathbf{w}_i} E\left[\left|\mathbf{c}_i^H\mathbf{R}\mathbf{w}_i^* - b_i\right|^2\right] \end{aligned}\tag{3.38}$$

The resulting $[\bar{\mathbf{c}}_i, \bar{\mathbf{w}}_i]$ pair yields the matrix filter

$$\mathbf{X}_i^{\text{CO-MMSE}} = \bar{\mathbf{c}}_i\bar{\mathbf{w}}_i^\top \tag{3.39}$$

This matrix filter is suboptimal for the optimization problem given in (3.31) since it is found in a constrained $\mathbf{X}_i$ space. The MSE function in (3.38) can be expressed as

$$\mathrm{MSE} = \sum_{j=1}^{N} p_jh_j|\mathbf{c}^H\mathbf{s}_j|^2|\mathbf{w}^H\mathbf{a}_j|^2 + \sigma^2(\mathbf{c}^H\mathbf{c})(\mathbf{w}^H\mathbf{w}) - 2\sqrt{p_ih_i}\Re\{(\mathbf{c}^H\mathbf{s}_i)(\mathbf{w}^H\mathbf{a}_i)\} + 1 \tag{3.40}$$

where $\Re\{\cdot\}$ denotes the real part of a complex number. It can be shown that, although (3.40) is convex in $\mathbf{c}_i$ for fixed $\mathbf{w}_i$ and convex in $\mathbf{w}_i$ for fixed $\mathbf{c}_i$, it is not jointly convex in both vector variables and the minimizer of MSE does not have a closed form expression. In this case, it is first necessary to ensure that the function indeed has a minimum. Fortunately, Weierstrass' Theorem [11] ensures that there exists a minimum if the function is continuous and coercive, i.e. $f(x) \to \infty$ when $\|x\| \to \infty$, as is the case for the MSE function given by (3.40). Unfortunately, this minimum is not attained by a unique pair due to nonconvexity. To see this, simply observe that any two pairs $[\mathbf{c}_i, \mathbf{w}_i]$

and $[\mathbf{c}_i', \mathbf{w}_i']$ will produce the same MSE if $\mathbf{c}' = \beta\mathbf{c}$ and $\mathbf{w}' = (1/\beta)\mathbf{w}$, where β is any nonzero scalar. Since this argument is true for all MSE values including the minimum MSE, we are guaranteed to have multiple global minima.

Due to the possible multimodality of the MSE function, standard iterative optimization algorithms cannot guarantee convergence to a global minima. We devise here an iterative algorithm based on block coordinate descent or nonlinear Gauss-Seidel method [11], also known as alternating minimization [20, 46], and investigate its convergence properties.

Consider fixing the value of one of the filters, say $\tilde{\mathbf{w}}$. It is then possible to find the filter, $\hat{\mathbf{c}}$, that maximally decreases the MSE function in (3.40). The solution is analogous to the *MMSE* detector described in [36] where user j's received amplitude is modified such that it is $\sqrt{p_j h_{ij}}(\tilde{\mathbf{w}}^H \mathbf{a}_{ij})$. With some abuse of notation, we will call this filter $\hat{\mathbf{c}} = \text{MMSE}(\tilde{\mathbf{w}})$:

$$
\begin{aligned}
\hat{\mathbf{c}} &= \text{MMSE}(\tilde{\mathbf{w}}) \\
&= \sqrt{p_i h_i}(\tilde{\mathbf{w}}^H \mathbf{a}_i)\left(\sum_{j=1}^{N} p_j h_j |\tilde{\mathbf{w}}^H \mathbf{a}_j|^2 \mathbf{s}_j \mathbf{s}_j^H + \sigma^2(\tilde{\mathbf{w}}^H \tilde{\mathbf{w}})\mathbf{I}\right)^{-1} \mathbf{s}_i \qquad (3.41)
\end{aligned}
$$

The same argument can be made for the case where $\mathbf{c}_i$ is fixed to $\tilde{\mathbf{c}}$ and the spatial filter is found to maximally decrease the MSE, $\hat{\mathbf{w}} = \text{MMSE}(\tilde{\mathbf{c}})$:

$$
\begin{aligned}
\hat{\mathbf{w}} &= \text{MMSE}(\tilde{\mathbf{c}}) \\
&= \sqrt{p_i h_i}(\tilde{\mathbf{c}}^H \mathbf{s}_i)\left(\sum_{j=1}^{N} p_j h_j |\tilde{\mathbf{c}}^H \mathbf{s}_j|^2 \mathbf{a}_j \mathbf{a}_j^H + \sigma^2(\tilde{\mathbf{c}}^H \tilde{\mathbf{c}})\mathbf{I}\right)^{-1} \mathbf{a}_i \qquad (3.42)
\end{aligned}
$$

Now, consider the following algorithm. Starting with the filter pair $\mathbf{c}(0), \mathbf{w}(0)$ and keeping $\mathbf{w}(0)$ fixed, one can find $\mathbf{c}(1) = \text{MMSE}(\mathbf{w}(0))$. Then keeping $\mathbf{c}(1)$ fixed, one can find $\mathbf{w}(1) = \text{MMSE}(\mathbf{c}(1))$ that further decreases the MSE in (3.40). Iteration $n+1$ of this two step iterative algorithm for user i is given below.

$$
\begin{aligned}
\mathbf{c}_i(n+1) &= \text{MMSE}(\mathbf{w}_i(n)) &\qquad (3.43)\\
\mathbf{w}_i(n+1) &= \text{MMSE}(\mathbf{c}_i(n+1)) &\qquad (3.44)
\end{aligned}
$$

Note that the order in which $\mathbf{c}_i$ and $\mathbf{w}_i$ are updated could be reversed. That is, we could devise a similar algorithm where $\mathbf{w}_i$ is updated before $\mathbf{c}_i$. The resulting MSE sequence given by the algorithm (3.43) and (3.44) is decreasing since

$$\mathrm{MSE}(\mathbf{c}(n-1), \mathbf{w}(n-1)) \geq \mathrm{MSE}(\mathbf{c}(n), \mathbf{w}(n-1)) \geq \mathrm{MSE}(\mathbf{c}(n), \mathbf{w}(n)) \tag{3.45}$$

and is bounded from below by the minimum MSE value. Thus, the algorithm is convergent. However, since the function is possibly multimodal, care must be taken to avoid undesirable stopping points. In particular, one can observe that $\mathbf{c} = \mathbf{w} = \mathbf{0}$ is an undesirable fixed point of the algorithm. A moment's thought reveals however, that this point is reachable only from a point where the filter in either temporal or spatial domains is orthogonal to the desired user's signature in that domain and this situation can be avoided by judicious choice of starting points.

To see this, we observe that the linear transformation (3.41) produces $\hat{\mathbf{c}} = \mathbf{0}$ iff $\tilde{\mathbf{w}}^H \mathbf{a}_i = 0$. Similarly, the linear transformation (3.42) produces $\hat{\mathbf{w}} = \mathbf{0}$ iff $\tilde{\mathbf{c}}^H \mathbf{s}_i = 0$. Thus, $[\mathbf{c}, \mathbf{w}] = [\mathbf{0}, \mathbf{0}]$ is reachable only from one of the following set of points: $[\mathbf{c}, \mathbf{w}_\perp]$ with $\mathbf{w}_\perp^H \mathbf{a}_i = 0$ or $[\mathbf{c}_\perp, \mathbf{w}]$ with $\mathbf{c}_\perp^H \mathbf{s}_i = 0$.

Now, recall the linear transformations in (3.41) and (3.42). They are projections onto the corresponding -temporal or spatial- signature spaces. Thus, if we start the algorithm at a point $\mathbf{w}(0)$ that lies in the linear vector space spanned by the spatial signatures, we can never arrive at a point outside the corresponding space for $\mathbf{c}$ or $\mathbf{w}$. Thus, we can avoid the undesirable fixed point $[\mathbf{c}, \mathbf{w}] = [\mathbf{0}, \mathbf{0}]$. Matched filter to the desired user's spatial signature, or any linear combination of the spatial signatures of all users are safe starting points.

Other than the obvious undesirable fixed point $[\mathbf{c}, \mathbf{w}] = [\mathbf{0}, \mathbf{0}]$, we have not encountered any other points where the algorithm would get stuck. Experimentally, we have observed that the MMSE (global minimum value) is achieved starting from random points and the resulting $[\mathbf{c}, \mathbf{w}]$ vectors are scalar multiples of each other in the form of $[\mathbf{c}, \mathbf{w}]$ and $[\beta\mathbf{c}, \mathbf{w}/\beta]$ where β is a nonzero scalar.

3.4 Performance Comparison

An important performance comparison criterion is the probability of bit error. Unfortunately, for general system parameters, it is difficult to derive analytical results for the probability of bit error rendering this comparison intractable. A commonly used approximation to the probability of bit error is obtained when a Gaussian approximation is applied to the total interference. It was reported in [53] that this Gaussian approximation is particularly accurate when MMSE receivers are employed. In this case, the bit error probability becomes a one-to-one decreasing function of the SIR. In what follows, we will give and compare the performance of all filters given in Sections 3.3.1 through 3.3.7 in terms of their SIRs. We will also compare the performances of the filters in asymptotic regimes when appropriate.

3.4.1 Optimum and Constrained Optimum Filters

The optimum temporal-spatial filter, $\mathbf{X}_i^{\text{O-MMSE}}$, is superior to the constrained optimum temporal-spatial filter, $\mathbf{X}_i^{\text{CO-MMSE}}$ in terms of the achievable MSE. This is a mere consequence of the fact that $\mathbf{X}_i^{\text{CO-MMSE}}$ is constrained to be in the rank 1 matrix space, while $\mathbf{X}_i^{\text{O-MMSE}}$ could take any value in the $G \times K$ dimensional matrix space.

The following proposition states that temporal-spatial filters that minimize the MSE in the corresponding filter spaces also maximize the SIR. Thus, by the same token, we can claim that $\mathbf{X}_i^{\text{O-MMSE}}$ is superior to $\mathbf{X}_i^{\text{CO-MMSE}}$ in terms of the achievable SIR also.

Proposition 1 *The filters that solve (3.31) and (3.38) achieve the maximum SIR over all filters in $\mathcal{C}^{G\times K}$ and $\mathcal{L}$, respectively.*

The proof of this Proposition is given in Appendix 4.A and is a simple extension of the solution to [78, Problem 6.5] (as given in [65]) to include complex numbers and constrained optimization.

It is also instructive to observe the difference between the performances of the optimum and the constrained optimum filters in terms of their asymptotic efficiencies and near-far resistances. Asymptotic efficiency of user i, with energy E_i, and bit error rate of the user in the multiuser environment, as a function of the background noise power σ^2, $P_i(\sigma)$, is defined as [78]

$$\eta_i = \sup\left\{0 \le r \le 1; \lim_{\sigma\to 0} \frac{P_i(\sigma)}{Q\left(\frac{\sqrt{rE_i}}{\sigma}\right)} < +\infty\right\} \tag{3.46}$$

where $Q(x) = \int_x^\infty (1/\sqrt{2\pi}) \exp(-y^2/2) dy$. Then the near-far resistance is defined as [78]

$$\bar{\eta}_i = \inf_{E_j, j \neq i} \eta_i \tag{3.47}$$

It is well-known that the asymptotic efficiency of an MMSE receiver is equal to the asymptotic efficiency of a decorrelating (zero-forcing) receiver [78]. Therefore, in order to study the asymptotic efficiencies of the unconstrained and constrained MMSE receivers ($\mathbf{X}_i^{\text{O-MMSE}}$ and $\mathbf{X}_i^{\text{CO-MMSE}}$ respectively), we will study the asymptotic efficiencies of the unconstrained and constrained (to rank 1 matrix space) decorrelating receivers. The unconstrained temporal-spatial decorrelating receiver, $\mathbf{X}_i$, for the ith user, is defined to be the solution of the following optimization problem [35].

$$\begin{aligned} \max \quad & \operatorname{tr}(\mathbf{X}_i^H \mathbf{s}_i \mathbf{a}_i^\top) \\ \text{s.t.} \quad & \operatorname{tr}(\mathbf{X}_i^H \mathbf{s}_j \mathbf{a}_j^\top) = 0 \qquad \text{for all} \quad j \neq i \\ & \operatorname{tr}(\mathbf{X}_i^H \mathbf{X}_i) = 1 \end{aligned} \tag{3.48}$$

Using the long vector notation, the unconstrained decorrelating receiver can be expressed as

$$\mathbf{x}_i = \mathbf{Q} \left(\mathbf{Q}^H \mathbf{Q} \right)^{-1} \mathbf{e}_i \tag{3.49}$$

where $\mathbf{Q}$ is a $KG \times N$ matrix whose columns are $\mathbf{q}_i$, i.e., $\mathbf{Q} = [\mathbf{q}_1 \cdots \mathbf{q}_N]$, and $\mathbf{q}_i$ are obtained by stacking the columns of $\mathbf{s}_i \mathbf{a}_i^\top$ into a KG dimensional vector, and $\mathbf{e}_i$ is the ith N dimensional unit vector, i.e., all entries of $\mathbf{e}_i$ are zero except the ith entry which is equal to 1. From the construction of the $\mathbf{Q}$ matrix, we have $\mathbf{Q}^H \mathbf{Q} = (\mathbf{A}^H \mathbf{A}) \circ (\mathbf{S}^H \mathbf{S})$, where $\mathbf{A} = [\mathbf{a}_1 \cdots \mathbf{a}_N]$ and $\mathbf{S} = [\mathbf{s}_1 \cdots \mathbf{s}_N]$, and $\circ$ denotes the Hadamard (elementwise) product of matrices. The asymptotic efficiency of the ith user is given as [35, 78]

$$\eta_X^i = \frac{1}{[(\mathbf{Q}^H \mathbf{Q})^{-1}]_{ii}} = \frac{1}{\left[((\mathbf{A}^H \mathbf{A}) \circ (\mathbf{S}^H \mathbf{S}))^{-1} \right]_{ii}} \tag{3.50}$$

The constrained temporal-spatial decorrelating filter is given in a similar way, except the addition of $\mathbf{X}_i \in \mathcal{L}$ into the constraint set of (3.48). Denoting the constrained decorrelating filter as $\mathbf{X}_i = \mathbf{c}_i \mathbf{w}_i^\top$, we can write (3.48) as

$$\begin{aligned} \max \quad & (\mathbf{c}_i^H \mathbf{s}_i)(\mathbf{w}_i^H \mathbf{a}_i) \\ \text{s.t.} \quad & (\mathbf{c}_i^H \mathbf{s}_j)(\mathbf{w}_i^H \mathbf{a}_j) = 0 \qquad \text{for all} \quad j \neq i \\ & (\mathbf{c}_i^H \mathbf{c}_i)(\mathbf{w}_i^H \mathbf{w}_i) = 1 \end{aligned} \tag{3.51}$$

Note that in order to satisfy the first constraint in (3.51) we may decorrelate the ith user from an interfering jth user either in time (by choosing $\mathbf{c}_i^H \mathbf{s}_j = 0$), or in space (by choosing $\mathbf{w}_i^H \mathbf{a}_j = 0$). Decorrelating the ith user from an interfering jth user both in time and space results in more enhanced background noise, and equivalently reduced asymptotic efficiency. Therefore, we have to partition all interfering users (to user i) into two subsets: those that will be decorrelated temporally and those that will be decorrelated spatially. Let us consider an arbitrary partition of interfering users into two sets. Let I_s and I_t denote the indices of the users that will be decorrelated from in space and time, respectively. Let $\mathbf{A}_s$ be a matrix composed of spatial signatures of the users in the set I_s and the spatial signature of the ith user as its columns, and $\mathbf{S}_t$ be a matrix composed of temporal signatures of the users in the set I_t and the temporal signature of the ith user as its columns. Then it can be shown that the asymptotic efficiency of the ith user with the constrained temporal-spatial decorrelating filter is given as

$$\eta_{cw}^i = \max_{I_s, I_t} \frac{1}{[(\mathbf{A}_s^H \mathbf{A}_s)^{-1}]_{ii} \left[(\mathbf{S}_t^H \mathbf{S}_t)^{-1}\right]_{ii}} \tag{3.52}$$

where the maximization is defined over all possible partitions of the interfering users into two sets. Note that for both the unconstrained and the constrained filters, near-far resistances are equal to the corresponding asymptotic efficiencies, i.e., $\bar{\eta}_X^i = \eta_X^i$ and $\bar{\eta}_{cw}^i = \eta_{cw}^i$, for all i, since the asymptotic efficiencies do not depend on the energies of the users. At first sight, it may seem difficult to compare the quantities in (3.50) and (3.52). In order to do this comparison, we will use a slightly different definition of the asymptotic efficiency. For a normalized decorrelating detector $\mathbf{X}_i$, i.e., the solutions of (3.48) and (3.51), the asymptotic efficiency is given as [35, 48, 78]

$$\eta_X^i = \left(\mathrm{tr}(\mathbf{X}_i^H \mathbf{s}_i \mathbf{a}_i^\top) \right)^2 \tag{3.53}$$

Note that this is the square of the cost function of the maximization problems in (3.48) and (3.51). Since the feasible set of (3.51) is contained in the feasible set of (3.48) we can conclude that the cost function at the solution of (3.48) is larger than or equal to the cost function at the solution of (3.51). Thus,

$$\eta_X^i \geq \eta_{cw}^i \tag{3.54}$$

and the optimum temporal-spatial MMSE filter has greater near-far resistance than the constrained optimum temporal-spatial MMSE filter.

A related interesting issue is how many interfering users a given user can cope with, in the sense of having a non-zero asymptotic efficiency, using the constrained and the unconstrained decorrelating receivers. As long as a user has a non-zero asymptotic efficiency, that user can achieve its quality of service requirement by increasing its transmit power. With the constrained decorrelating receiver, a user can suppress up to $G-1$ users in time, and up to $K-1$ users in space, as long as the temporal signature sequences of the users to be suppressed in time and the spatial signature sequences of the users to be suppressed in space are linearly independent. Therefore, a user can suppress at most $K+G-2$ interferers by using a constrained temporal-spatial decorrelating receiver. Clearly, the number of of interferers a user can suppress by using an unconstrained temporal-spatial decorrelating receiver is higher, since the constrained decorrelating receiver is a special case of an unconstrained decorrelating receiver. Using the long vector formulation in (3.49), one would conclude that a desired user can suppress up to $KG-1$ interfering users since the dimensionality of the temporal-spatial signatures is KG. However, these "long" temporal-spatial signatures have a repetitive structure. In other words, the temporal-spatial signature of a given user is composed of the temporal signature of the same user concatenated K times after being multiplied by the K antenna gains. This repetitive structure may result in loss of dimensionality. However, in recent work [27], for a large system where $N \to \infty$ and $G \to \infty$ but $\alpha = N/G$ and K are fixed and finite, the dimensionality of the temporal-spatial signatures has been shown to be KG as long as the temporal signatures are random, and antenna gains are uncorrelated.

3.4.2 Constrained Optimum Filter and Cascaded Optimum Filters

From the arguments in Section 3.4.1, it is clear that the optimum MMSE receiver of Section 3.3.6 outperforms all other receiver structures mentioned in Sections 3.3.1 through 3.3.5, as well as the constrained optimum MMSE receiver in Section 3.3.7, in terms of both the MSE and the SIR. It only remains to compare the performance of the constrained optimum MMSE receiver of Section 3.3.7 with the receiver structures in Sections 3.3.1 through 3.3.5. First we observe from (3.7), (3.8), (3.10), (3.30) that the filters $\mathbf{X}_i^{\text{MF-MF}}$, $\mathbf{X}_i^{\text{MMSE-MF}}$, $\mathbf{X}_i^{\text{MF-MMSE}}$, $\mathbf{X}_i^{\text{ST-CMMSE}}$ are of rank 1. Given that the constrained optimum MMSE, $\mathbf{X}_i^{\text{CO-MMSE}}$, minimizes the MSE and therefore maximizes the SIR among all possible rank 1 matrix filters, we conclude that the constrained optimum MMSE receiver filter outperforms all of these suboptimum receiver filters. In fact, the iterative algorithm described by ((3.43),(3.44)) can be started at any of the temporal-spatial filter pairs that define $\mathbf{X}_i^{\text{MF-MF}}$, $\mathbf{X}_i^{\text{MMSE-MF}}$, $\mathbf{X}_i^{\text{MF-MMSE}}$ or $\mathbf{X}_i^{\text{ST-CMMSE}}$. Since each iteration increases the SIR and decreases the MSE monotonically, with each iteration, the performance of the resulting filter pair is better than the previous one and the convergence point temporal-spatial filter pair, $\mathbf{X}_i^{\text{CO-MMSE}}$, outperforms the starting point filter pair.

The cascaded temporal-spatial MMSE filter, $\mathbf{X}_i^{\text{TS-CMMSE}}$, in Section 3.3.4 can have rank up to K, just as the joint optimum MMSE filter of Section 3.3.6 which can have rank up to $\min\{K, G\}$. Thus, there could be cases under which $\mathbf{X}_i^{\text{TS-CMMSE}}$ performs better than the constrained optimum filter; see Section 4.5. However, the fact that $\mathbf{X}_i^{\text{TS-CMMSE}}$ has higher rank than $\mathbf{X}_i^{\text{CO-MMSE}}$ does not necessarily guarantee that it yields a lower MSE or a higher SIR than $\mathbf{X}_i^{\text{CO-MMSE}}$. This is also demonstrated in Section 4.5 and is intuitively pleasing since $\mathbf{X}_i^{\text{TS-CMMSE}}$, in contrast with $\mathbf{X}_i^{\text{CO-MMSE}}$, is not designed with the objective of minimizing the end-to-end MSE of the user (i.e., from the transmitted bit of the ith user to the final filter output of the ith user). Thus, neither $\mathbf{X}_i^{\text{CO-MMSE}}$ nor $\mathbf{X}_i^{\text{TS-CMMSE}}$ performs uniformly better than the other; depending on the system parameters (N, G, K, spatial and temporal signatures of the users, etc.) one may outperform the other.

It is possible to compare the two filters in the asymptotic case when the background noise σ^2 goes to zero, or equivalently the received powers of the interfering users go to infinity. Once again, the MMSE receiver reduces to a decorrelating receiver as the background noise power goes to zero or the received powers of the interfering users go to infinity [78]. The temporal decorrelating filter for user i, $\mathbf{d}_i$, is found by

$$\mathbf{d}_i = \mathbf{S}(\mathbf{S}^H\mathbf{S})^{-1}\mathbf{e}_i = \mathbf{S}\mathbf{\Gamma}^{-1}\mathbf{e}_i \tag{3.55}$$

for a chip matched filtered CDMA received signal where users' temporal signatures $\mathbf{s}_j$ constitute the columns of the signature matrix $\mathbf{S}$. Note that $\mathbf{\Gamma}$ is the cross correlation matrix of users' temporal signatures and is invertible if users' signatures are linearly independent. When this is not the case, the decorrelator is found by using the Moore-Penrose generalized inverse of the cross correlation matrix, $\mathbf{\Gamma}^+$, and remains independent of the received powers of the users [35].

Recall now that in the calculation of the cascaded temporal-spatial MMSE receiver filter $\mathbf{X}_i^{\text{TS-CMMSE}}$, first, K temporal MMSE receiver filters are found; see (3.13) and that these K MMSE receivers, $\mathbf{c}_k$s, are different due to the fact that the received powers of the users are different at each antenna array element. This is because the actual received powers of the users are multiplied with the square magnitudes of the antenna gains $|a_j^{(k)}|$ at the outputs of different antenna elements. Since in the interference limited regime where the ratio of the background noise power to received interfering powers goes to zero, the MMSE receivers go to decorrelators and since decorrelators are independent of the received powers of the users, all K temporal receiver filters become identical, i.e., $\mathbf{c}_k = \mathbf{c}$ for all k. Note that this is true even when the cross correlation matrix is not invertible and the Moore-Penrose generalized inverse is used. When the temporal filters at the output of all antenna array elements are the same, the receiver filter $\mathbf{X}_i^{\text{TS-CMMSE}}$ becomes a rank 1 filter. The MSE achieved by $\mathbf{X}_i^{\text{TS-CMMSE}}$, $\text{MSE}_i^{\text{TS-CMMSE}}$, is larger than that of the constrained optimum MMSE filter, $\text{MSE}_i^{\text{CO-MMSE}}$, simply because $\mathbf{X}_i^{\text{CO-MMSE}}$ is the filter that yields the minimum MSE among all rank 1 matrix filters. Thus,

$$\lim_{\sigma^2 \to 0} \text{MSE}_i^{\text{TS-CMMSE}} \geq \lim_{\sigma^2 \to 0} \text{MSE}_i^{\text{CO-MMSE}} \tag{3.56}$$

Equivalently, using (3.66) from Appendix 3.A, the SIRs achieved by these two filters in this interference limited regime is compared as

$$\lim_{\sigma^2 \to 0} \mathrm{SIR}_i^{\text{TS-CMMSE}} \leq \lim_{\sigma^2 \to 0} \mathrm{SIR}_i^{\text{CO-MMSE}} \tag{3.57}$$

Hence, the constrained optimum MMSE filter outperforms the cascaded temporal-spatial MMSE filter of Section 3.3.4 asymptotically.

3.5 Results

Numerical results showing the performance of the filters considered in this chapter are now given. We consider a single cell CDMA system, the base station of which employs a linear antenna array with half a wavelength spacing. Thus, the array response for user i at the kth array element is given by

$$a_{i,k} = \exp[j(k-1)\pi \sin(\theta_k) \tag{3.58}$$

where θ_k is the direction of arrival for user k. For each experiment we present here, the temporal signatures and users' positions which in turn determine the spatial signatures are created randomly, but then kept fixed for that experiment. The experiments are intended to demonstrate the behavior and performance of all the filters under different system loading conditions. For each experiment, we plot the output signal-to-interference ratio (SIR) for the desired user (in linear scale) versus the received signal-to-noise ratios (SNRs) of all interferers (in dB scale). The severe near-far conditions where the interferers' powers are very high are intended to show the asymptotic behavior of the filters, i.e., their performance in the interference limited environment where noise level becomes negligible compared to the interference power. The desired user's SNR is kept at 10dB.

Consider first a system with processing gain $G = 8$, $K = 2$ array elements and $N = 2$ users. Figure 3.4 shows the output SIR of the desired user. As expected, with only a single interferer present, all filters perform well even under very severe near-far conditions where the interferer's power is as much as 60dB above the desired user's. The only exception is the temporal-spatial matched filter which is well-known to be not near-far resistant. The more interesting observation about this system is better observed in Figure 3.5. Recall that we concluded in Section 3.4.2 that the constrained optimum MMSE filter does not necessarily outperform the cascaded temporal-spatial MMSE filter of Section 3.3.4, it only is as good or better than the cascaded temporal-spatial MMSE filter asymptotically. Indeed, in Figure 3.5, we see that the cascaded temporal-spatial MMSE outperforms the constrained optimum MMSE filter. When the system becomes interference limited, both detectors have identical performance. The constrained optimum MMSE filter outperforms all other filters, except for the optimum MMSE filter which is the SIR maximizer among all matrix filters.

For the rest of this section, the system considered has $K = 4$ antenna array elements and $G = 16$ processing gain. We will examine the performance of the filter structures for this system under different loading conditions.

We first consider $N = 8$ users. Figure 3.6 shows the output SIR of the desired user. The constrained optimum MMSE filter outperforms all filters, except for the optimum MMSE filter. Note that, for this system, user's temporal signature sequences are linearly independent and thus, when the system is interference limited, the filters that perform temporal MMSE first, i.e., the temporal MMSE filter-spatial matched filter of Section 3.3.3 and the cascaded temporal-spatial MMSE filter of Section 3.3.4 end up decorrelating all interferers in temporal domain. Specifically, the temporal MMSE filter at the output of each antenna becomes a decorrelator, $\mathbf{c}_k = \mathbf{c}$, for all $k = 1, ..., K$, for the the cascaded temporal-spatial MMSE filter. In this case, the output statistics of the first stage of the cascaded temporal-spatial MMSE filter are interference free, i.e., (3.14) can be expressed as

$$\tilde{\mathbf{y}} = \sqrt{p_i h_i} b_i (\mathbf{c}^H \mathbf{s}_i) \mathbf{a}_i + \tilde{\mathbf{n}} \tag{3.59}$$

where $\tilde{n}_k$ is the enhanced noise at the output of the kth antenna and the components of $\tilde{\mathbf{n}}$ are independent. Thus, the second stage spatial MMSE combiner $\mathbf{w}$ in (3.16) becomes the spatial matched filter, $\mathbf{a}_i$, which explains why the temporal MMSE-spatial matched filter and the cascaded temporal-spatial MMSE filter have identical performance asymptotically. It is also notable that constrained optimum MMSE detector becomes a temporal-spatial decorrelator and chooses to suppress some of the interferers in temporal domain and others in spatial domain such that it gets the best asymptotic SIR among such temporal-spatial decorrelators.

Next we consider $N = 16$ users. The output SIR of the desired user is plotted in Figure 3.7. There are still enough temporal dimensions for users to be decorrelated in the temporal domain, i.e., $N \leq G$, even if they can not all be suppressed in the spatial domain. As a result, all filters that employ MMSE combining in the temporal domain have non-zero SIRs asymptotically, i.e., $X^{\text{MMSE-MF}}$, $X^{\text{ST-CMMSE}}$, $X^{\text{TS-CMMSE}}$, $X^{\text{CO-MMSE}}$, and $X^{\text{O-MMSE}}$ are all near-far resistant. However, the constrained optimum MMSE filter, again by choosing the appropriate users to suppress in the spatial or temporal domains achieve higher SIR over all filters except the optimum MMSE filter.

The next system to be considered has $N = 18$ users and results are shown in Figure 3.8. Since the number of interferers are larger than both the processing gain and the number of array elements, all interferers cannot be suppressed in a single domain, thus the combined single user/multiple user filter structures, i.e., $X^{\text{MMSE-MF}}$ and $X^{\text{MF-MMSE}}$, are not near-far resistant. For this example, the cascaded structures, i.e., $X^{\text{ST-CMMSE}}$ and $X^{\text{TS-CMMSE}}$ are not near-far resistant either, since in the interference dominated regime, each stage tries to suppress all interferers independently in cascade. When an interference suppressor is designed by considering both domains jointly, as in the case of constrained optimum and optimum detectors, $X^{\text{CO-MMSE}}$, and $X^{\text{O-MMSE}}$, near-far resistance is achieved.

The last example we consider is a very highly loaded system with $N = 36$ users. The purpose of this experiment is to show the difference between the optimum MMSE filter and the constrained optimum MMSE filter. We observe from Figure 3.9 that although the constrained optimum MMSE filter, $X^{\text{CO-MMSE}}$, results in acceptable SIR values in near-far situations, e.g. an output SIR of 5 (7dB) when all interferers' powers are 10dB higher than the desired user, it is *not* near-far resistant. This is simply due to the fact that the constrained optimum MMSE filter can suppress up to $G - 1$ users in the temporal domain and $K - 1$ users in the spatial domain. Thus, for this example, when $N > 19$, the constrained optimum MMSE filter is not able to suppress all the interference and the output SIR it produces approaches 0 when the interferers' powers approach infinity. The optimum MMSE detector on the other hand can suppress up to $KG-1$ users and for this example is near-far resistant.

3.6 Chapter Summary

In this chapter, we have examined different possible filter structures for a CDMA system when both temporal and spatial filters are used at the receiver. It is shown that the joint domain approaches, the optimum MMSE filter and the constrained optimum MMSE filter outperform the approaches where the temporal and spatial receivers are designed independently in cascade. While the optimum MMSE is the filter that maximizes the output SIR of the desired user over all possible temporal-spatial filter structures, its complexity may render its usage impractical and other structures have to be considered. We have observed that the constrained optimum MMSE filter performs close to the optimum MMSE filter for all but extremely loaded systems and outperform all other structures either uniformly or asymptotically (in the decorrelating regime) and shows promise for implementability because of its low complexity.

3.A Appendix

Proof: Proposition 1 Let us assume a general matrix receiver filter $\mathbf{X}$. Let us represent the desired signal part of the received signal, i.e., the signal of user i, with $b_i\mathbf{S}_i$, and the multiaccess interference and AWGN part of the received signal with $\mathbf{Y}_i$, i.e. $\mathbf{R} = b_i\mathbf{S}_i + \mathbf{Y}_i$:

$$\begin{aligned} \mathbf{S}_i &= \sqrt{p_i h_i}\mathbf{s}_i\mathbf{a}_i^\top \\ \mathbf{Y}_i &= \sum_{j\neq i} \sqrt{p_j h_j} b_j \mathbf{s}_j \mathbf{a}_j^\top + \mathbf{N} \end{aligned} \tag{3.60}$$

The MSE and SIR with filter $\mathbf{X}$ are given by

$$\mathrm{MSE}(\mathbf{X}) = E\big[|\mathrm{tr}(\mathbf{X}^H\mathbf{R}) - b_i|^2\big] = |\mathrm{tr}(\mathbf{X}^H\mathbf{S}_i)|^2 + E\big[|\mathrm{tr}(\mathbf{X}^H\mathbf{Y}_i)|^2\big] - 2\Re\{\mathrm{tr}(\mathbf{X}^\top\mathbf{S}_i^*)\} + 1 \tag{3.61}$$

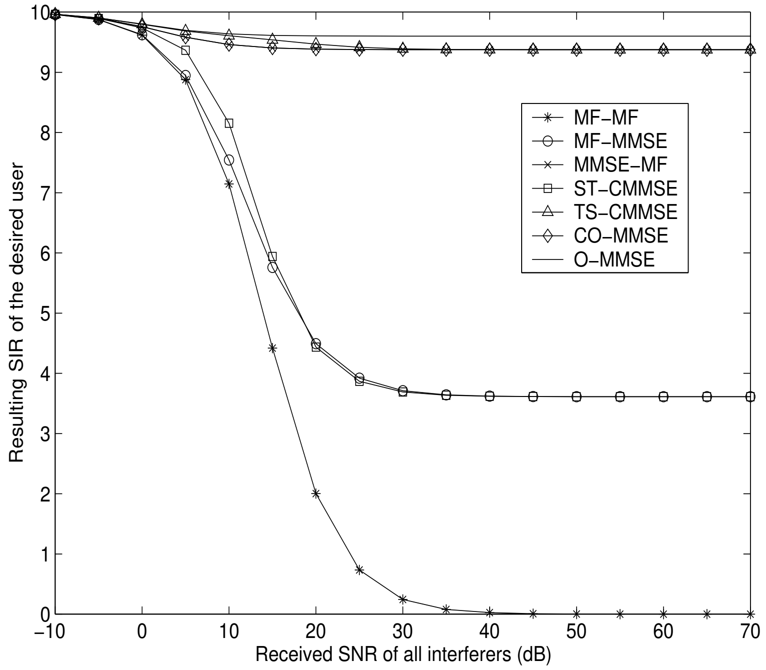

Figure 3.4: Resulting SIR of the desired user versus received SNR of all the interferers. $N = 2, K = 2, G = 8$.

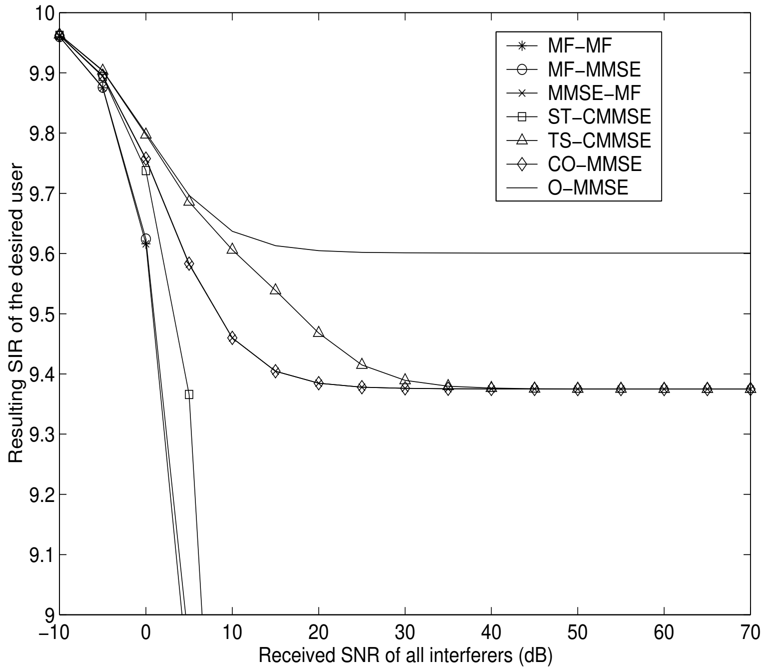

Figure 3.5: Resulting SIR of the desired user versus received SNR of all the interferers. $N = 2, K = 2, G = 8$, Figure 3.4 magnified.

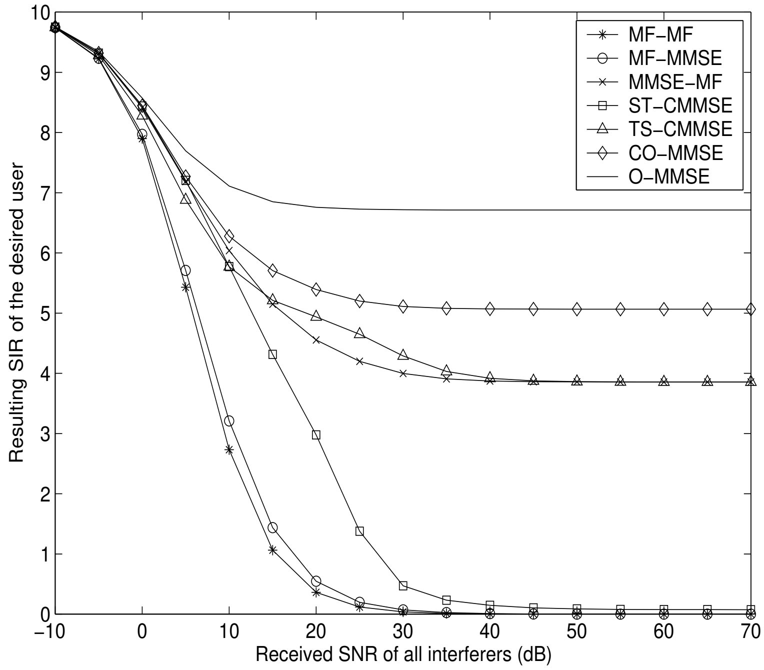

Figure 3.6: Resulting SIR of the desired user versus received SNR of all the interferers. $N = 8, K = 4, G = 16$.

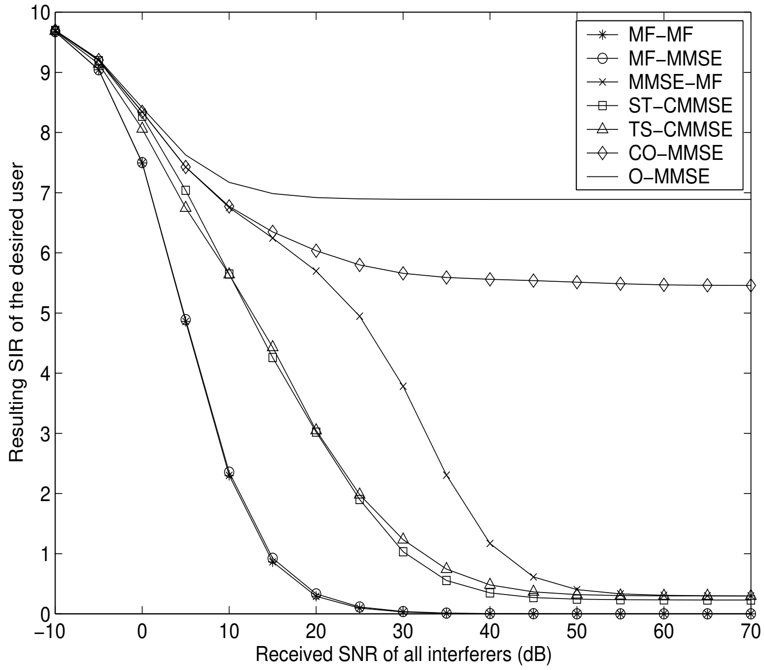

Figure 3.7: Resulting SIR of the desired user versus received SNR of all the interferers. $N = 16, K = 4, G = 16$.

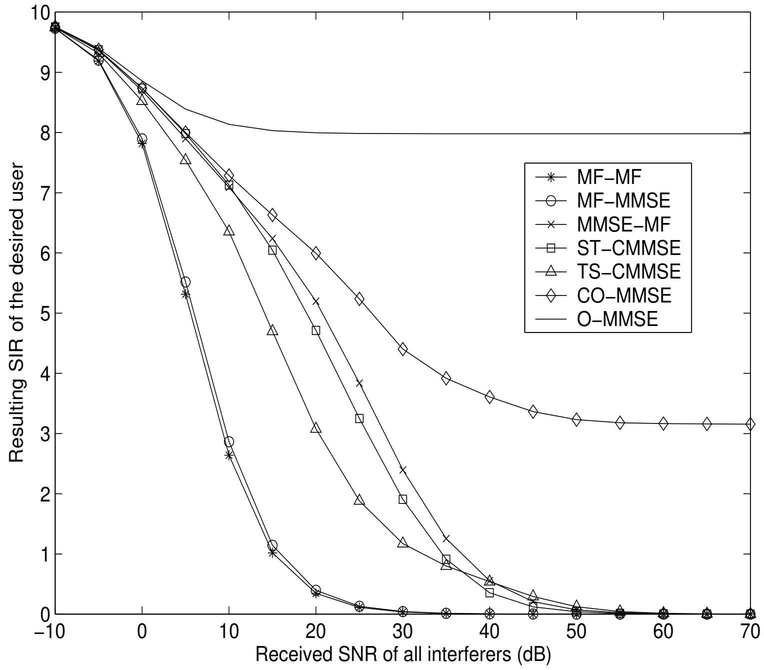

Figure 3.8: Resulting SIR of the desired user versus received SNR of all the interferers. $N = 18, K = 4, G = 16$.

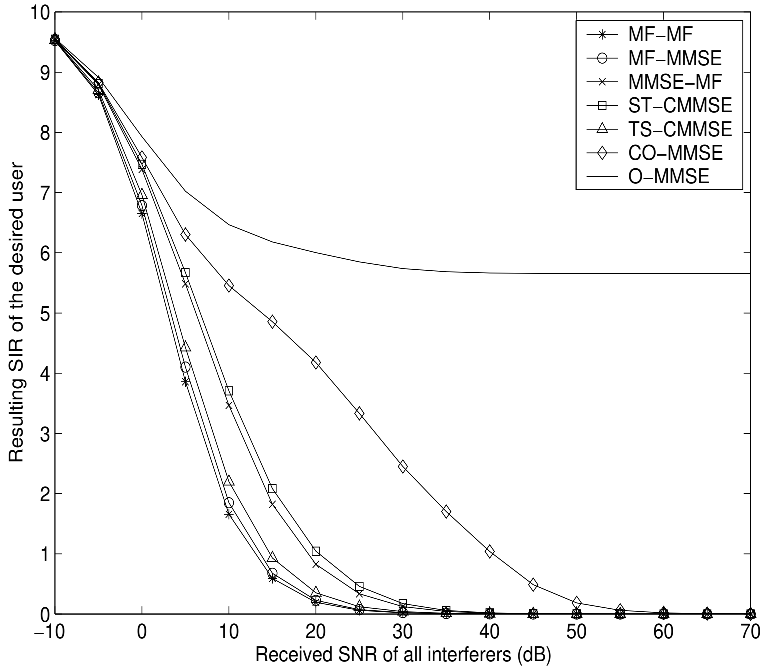

Figure 3.9: Resulting SIR of the desired user versus received SNR of all the interferers. $N = 36, K = 4, G = 16$.

and

$$\mathrm{SIR}(\mathbf{X}) = \frac{|\mathrm{tr}(\mathbf{X}^H \mathbf{S}_i)|^2}{E[|\mathrm{tr}(\mathbf{X}^H \mathbf{Y}_i)|^2]} \tag{3.62}$$

Now let us consider the MSE with a scaled version of the filter $\mathbf{X}$.

$$\mathrm{MSE}(\alpha\mathbf{X}) = |\alpha|^2 \left\{|\mathrm{tr}(\mathbf{X}^H \mathbf{S}_i)|^2 + E\left[|\mathrm{tr}(\mathbf{X}^H \mathbf{Y}_i)|^2\right]\right\} - 2\Re\{\mathrm{tr}(\alpha\mathbf{X}^T \mathbf{S}_i^*)\} + 1 \tag{3.63}$$

Setting the derivative of $\mathrm{MSE}(\alpha\mathbf{X})$ with respect to the real and the imaginary parts of α equal to zero, the complex scalar α that minimizes $\mathrm{MSE}(\alpha\mathbf{X})$ can be found as

$$\bar{\alpha} = \arg\min_{\alpha} \mathrm{MSE}(\alpha\mathbf{X}) = \frac{\mathrm{tr}(\mathbf{X}^H \mathbf{S}_i)}{|\mathrm{tr}(\mathbf{X}^H \mathbf{S}_i)|^2 + E[|\mathrm{tr}(\mathbf{X}^H \mathbf{Y}_i)|^2]} \tag{3.64}$$

Thus,

$$\min_{\alpha} \mathrm{MSE}(\alpha\mathbf{X}) = 1 - \frac{|\mathrm{tr}(\mathbf{X}^H \mathbf{S}_i)|^2}{|\mathrm{tr}(\mathbf{X}^H \mathbf{S}_i)|^2 + E[|\mathrm{tr}(\mathbf{X}^H \mathbf{Y}_i)|^2]} \tag{3.65}$$

Using (3.65) and (3.62) we can write

$$\frac{1}{\min_{\alpha} \mathrm{MSE}(\alpha\mathbf{X})} = 1 + \mathrm{SIR}(\mathbf{X}) \tag{3.66}$$

Equation (3.66) is true for any complex filter $\mathbf{X}$. In the unconstrained temporal-spatial filtering case, $\mathbf{X}$ can take any value in $\mathcal{C}^{G\times K}$, and the constrained case it is constrained to be in the rank 1 matrix space denoted by $\mathcal{L}$. In order to represent the constrained and unconstrained cases in a unified fashion, we will restrict $\mathbf{X}$ to be $\mathbf{X} \in \mathcal{S}$ with $\mathcal{S} = \mathcal{C}^{G\times K}$ for the unconstrained and $\mathcal{S} = \mathcal{L}$ in the constrained case. Maximizing both sides of (3.66) with respect to $\mathbf{X} \in \mathcal{S}$, i.e.,

$$\max_{\mathbf{X}\in\mathcal{S}} \frac{1}{\min_{\alpha} \mathrm{MSE}(\alpha\mathbf{X})} = 1 + \max_{\mathbf{X}\in\mathcal{S}} \mathrm{SIR}(\mathbf{X}) \tag{3.67}$$

is equivalent to

$$\frac{1}{\min\limits_{\mathbf{X}\in\mathcal{S}} \min\limits_{\alpha} \mathrm{MSE}(\alpha\mathbf{X})} = 1 + \max_{\mathbf{X}\in\mathcal{S}} \mathrm{SIR}(\mathbf{X}) \tag{3.68}$$

Combining continuous variables $\mathbf{X}$ and α into one variable and noting that, for both unconstrained and constrained cases, if $\mathbf{X} \in \mathcal{S}$, then $\alpha\mathbf{X} \in \mathcal{S}$ yields

$$\frac{1}{\min\limits_{\mathbf{X}\in\mathcal{S}} \mathrm{MSE}(\mathbf{X})} = 1 + \max_{\mathbf{X}\in\mathcal{S}} \mathrm{SIR}(\mathbf{X}) \tag{3.69}$$

where we also used the fact that $\text{SIR}(\mathbf{X})$ is insensitive to the scaling of its argument.

Therefore, (3.69) verifies that the filter $\mathbf{X} \in \mathcal{S}$ that minimizes the MSE is the one that maximizes the SIR. As shown above, this result is valid for both $\mathcal{S} = \mathcal{C}^{G \times K}$ and $\mathcal{S} = \mathcal{L}$. We conclude that the MMSE filters in both the unconstrained and the constrained spatial-temporal filter spaces maximize the SIR. $\square$

Chapter 4

Power Control, Multiuser Detection and Beamforming for CDMA

4.1 Introduction

In Chapters 2 and 3, we investigated the multiuser detection and combined multiuser detection and receiver beamforming for interference management. In this chapter, we will investigate the capacity gain that these receiver techniques can provide when combined together optimally with transmit power control, and the algorithms that can realize this gain.

Transmit power control is an efficient interference management technique. It is currently being used in second generation CDMA systems, i.e., IS-95, and is being proposed for third generation CDMA systems, i.e., W-CDMA and cdma2000. In very general terms, power control balances received powers of all users so that no user creates excessive interference to other users in the system. Power control has been studied extensively in the literature, starting with centralized and non-iterative solutions [24,95]. This is followed by the development of iterative and distributed algorithms that require only local measurements [21, 86, 94]. Traditional iterative power control approaches assume that only one antenna and matched filter receivers are being used at the base stations and each user employs an SIR based power update where the user's power is multiplied by the ratio of its target SIR to its current SIR as given by (1.6), see Figure 4.1.

It is notable that third generation CDMA systems anticipates the use of antenna arrays at the base station and also multiuser detection techniques [16]. Thus, it is meaningful to investigate the combined use of transmit power control with other interference management methods.

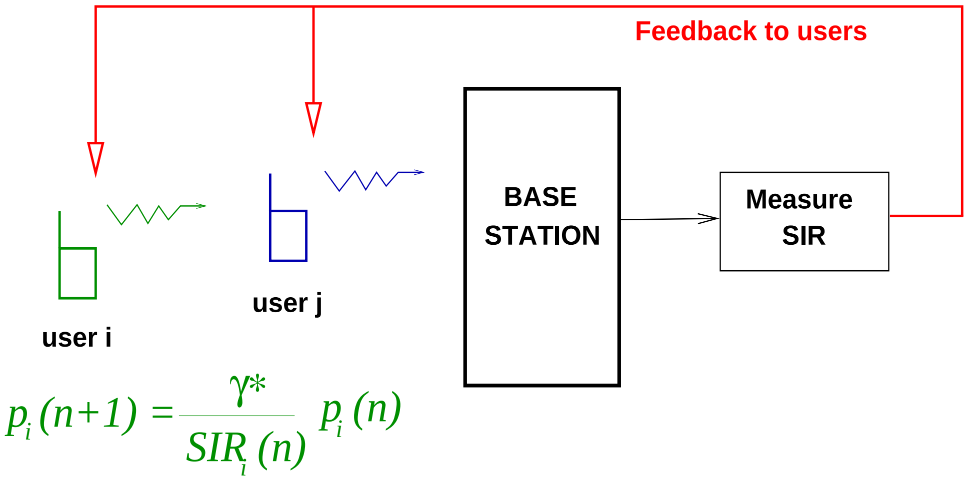

Figure 4.1: Conventional Power Control uses matched filters as receivers

Most of the receiver processing literature concentrated on developing signal processing algorithms without considering the issue of optimum transmit power control, assuming the need for power control can be alleviated by intelligent receiver design. More recently, combining power control and multiuser detection for CDMA has been studied in [32, 67]. In [67], the problem of finding the jointly optimum powers and linear receiver filters was studied. It was shown that a distributed and iterative power control algorithm where each user optimizes its linear receiver filter before each power control update converged to the point where all users expend minimum transmit power and use the corresponding MMSE linear filters. This work assumed a single antenna at each base station. A similar development arose in joint power control and beamforming for wireless networks in [56] where it was shown that a capacity increase is possible with power control if array observations are combined in the MMSE sense. For its applications to CDMA, this work assumed matched filters, i.e. no multiuser detection.

In this chapter, we combine the three basic interference management approaches, transmit power control, multiuser detection and beamforming to further increase the uplink capacity of a CDMA system. Linear processing is assumed in both the temporal and the spatial domains. As in chapter 3, we use matrix filters as receivers. The aim is to assign each user with just enough transmit power and find the best temporal-spatial filter to process the received signal such that each user achieves its target SIR. We

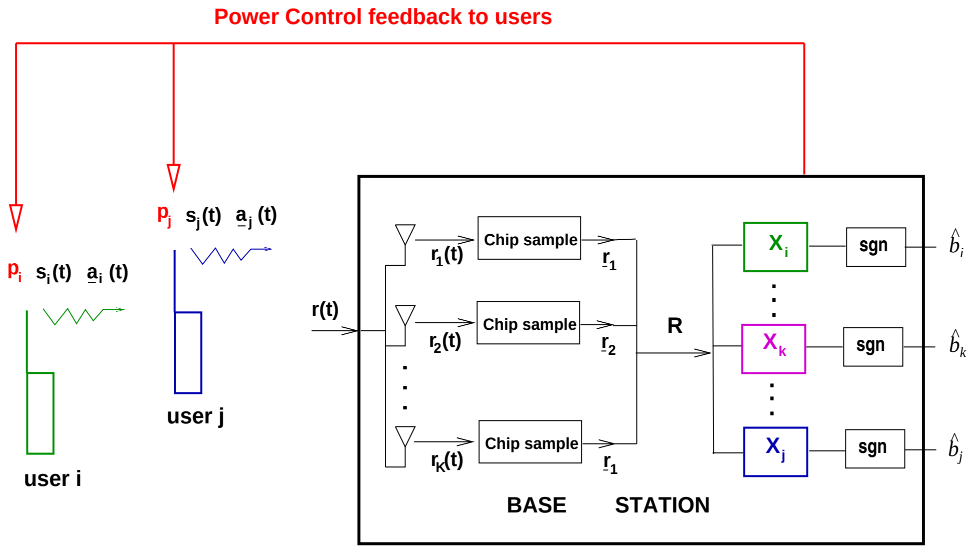

Figure 4.2: Power Control when temporal-spatial filters are used as receivers

assume two receiver structures: the first is a temporal-spatial matrix filter that has no constraints and the second is constrained to be of rank 1, i.e. a separable filter. We derive iterative power control algorithms that update the filters and the powers of all users that converge to the joint optimal powers and temporal-spatial filters, for both the unconstrained and the constrained receiver architectures (Figure 4.1). It is observed that combining the three approaches, i.e. power control and intelligent combining in both spatial and temporal domains, leads to significant savings in total transmit power and can increase capacity by supporting all users in some highly loaded systems that would otherwise be infeasible.

4.2 System Model

We consider a multicell DS-CDMA system with processing gain G. Number of active users over all the cells is N. At each base station an antenna array of K elements is employed. Once again, for clarity of exposition, we assume a synchronous system. Initially, we will assume that base station assignment has been done for all users. The base station selection will be incorporated into our interference management algorithms in Section 4.4. To incorporate the multicell model, we will use the base station index as necessary in our notation. Similar to (3.1), the received signal at the output of the antenna array at the assigned base station of user i is

$$\mathbf{r}_i(t) = \sum_{j=1}^{N} \sqrt{p_j h_{ij}} b_j s_j(t) \mathbf{a}_{ij} + \mathbf{n}_i(t) \tag{4.1}$$

where, as usual, p_j, b_j and $s_j(t)$ are the transmit power, bit and the signature of user j, respectively. The uplink gain of user j to the assigned base station of i is h_{ij} and $\mathbf{a}_{ij}$ is the K dimensional array response vector of user j (spatial signature) at the base station of i. Chip matched filtering the received signal and sampling at the chip rate, we have G observations at the output of each of the K antenna elements. Thus, the received matrix at the base station of user i is

$$\mathbf{R}_i = \sum_{j=1}^{N} \sqrt{p_j h_{ij}} b_j \mathbf{s}_j \mathbf{a}_{ij}^\top + \mathbf{N}_i \tag{4.2}$$

where k^{th} column of $\mathbf{R}_i$ represents the chip sampled outputs at the output of the k^{th} antenna array element. $\mathbf{N}_i$ is the matrix that represents the spatially and temporally white noise, i.e. $E[[N_i]_{kl}^*[N_i]_{mn}] = \sigma^2 \delta_{km} \delta_{ln}$.

4.3 Joint Power Control and Temporal-Spatial Filtering

Our aim, in this section, is to find optimal transmit powers, p_i, and matrix filters, $\mathbf{X}_i$, for both constrained and unconstrained cases, for $i = 1, \dots, N$, such that the total transmitter power is minimized while each user i satisfies its quality of service requirement, $\text{SIR}_i \geq \gamma_i^*$, where γ_i^*, called the *target* SIR, is the minimum acceptable level of SIR for user i. The SIR of user i at the output of the joint spatial and temporal filter which is employed at the assigned base station of user i, can be expressed as

$$\text{SIR}_i = \frac{p_i h_{ii} \left|\text{tr}(\mathbf{X}_i^H \mathbf{s}_i \mathbf{a}_{ii}^\top)\right|^2}{\sum_{j \neq i} p_j h_{ij} \left|\text{tr}(\mathbf{X}_i^H \mathbf{s}_j \mathbf{a}_{ij}^\top)\right|^2 + \sigma^2 \text{tr}(\mathbf{X}_i^H \mathbf{X}_i)} \tag{4.3}$$

We can then state the optimization problem as

$$\begin{aligned} \min \quad & \sum_{i=1}^{N} p_i \\ \text{s.t.} \quad & p_i \geq \frac{\gamma_i^*}{h_{ii}} \frac{\sum_{j \neq i} p_j h_{ij} |\text{tr}(\mathbf{X}_i^H \mathbf{s}_j \mathbf{a}_{ij}^\top)|^2 + \sigma^2 \text{tr}(\mathbf{X}_i^H \mathbf{X}_i)}{|\text{tr}(\mathbf{X}_i^H \mathbf{s}_i \mathbf{a}_{ii}^\top)|^2} \\ & p_i \geq 0, \quad \mathbf{X}_i \in \mathcal{S} \qquad i = 1, \dots, N \end{aligned} \tag{4.4}$$

where, following the definitions given in Section 3.3.7, $\mathcal{S} = \mathcal{C}^{G\times K}$ for unconstrained temporal-spatial filtering and $\mathcal{S} = \mathcal{L}$ for constrained temporal-spatial filtering. As in the case for the joint power control and temporal filtering [67], we can write (4.4) as

$$\begin{aligned}
\min_{\mathbf{p}} \quad & \sum_{i=1}^{N} p_i \\
\text{s.t.} \quad & p_i \geq \frac{\gamma_i^*}{h_{ii}} \min_{\mathbf{X}_i \in \mathcal{S}} \frac{\sum_{j\neq i} p_j h_{ij} |\mathrm{tr}(\mathbf{X}_i^H \mathbf{s}_j \mathbf{a}_{ij}^\top)|^2 + \sigma^2 \mathrm{tr}(\mathbf{X}_i^H \mathbf{X}_i)}{|\mathrm{tr}(\mathbf{X}_i^H \mathbf{s}_i \mathbf{a}_{ii}^\top)|^2} \\
& p_i \geq 0 \qquad i = 1, \ldots, N
\end{aligned} \tag{4.5}$$

The minimization over $\mathbf{X}_i$, on the right side of each of the power constraints above, is equivalent to maximizing SIR_i given by (4.3) for a fixed power p_i in the corresponding filter spaces. Recall that, as stated by Proposition 1, temporal-spatial filters that minimize the MSE in the corresponding filter spaces also maximize the SIR. Thus, the filter $\mathbf{X}_i$ that solves the minimization on the right side of p_i is the optimum or the constrained optimum temporal-spatial filters described in Sections 3.3.6 and 3.3.7.

Now, let us devise an iterative algorithm that converges to the optimum of (4.4). Iterative power control algorithms of the form

$$\mathbf{p}(n+1) = \mathbf{I}(\mathbf{p}(n)) \tag{4.6}$$

are analyzed for *standard interference functions*, $\mathbf{I}(\mathbf{p})$, in [86]. The definition of a standard interference function and the corresponding convergence result will be used throughout this chapter and are restated here for convenience.

Definition 1 $\mathbf{I}(\mathbf{p})$ *is a standard interference function if for all* $\mathbf{p} \geq \mathbf{0}$ *the following properties are satisfied.*

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

Theorem 3 *If there exists* $\mathbf{p}' \geq \mathbf{I}(\mathbf{p}')$, *then for any initial power vector* $\mathbf{p}(0)$, *the sequence* $\mathbf{p}(n) = \mathbf{I}(\mathbf{p}(n-1))$ *converges to a unique fixed point* $\bar{\mathbf{p}}$ *such that* $\bar{\mathbf{p}} \leq \mathbf{p}'$ *for any* $\mathbf{p}' \geq \mathbf{I}(\mathbf{p}')$.

The condition that there exists $\mathbf{p}' \geq \mathbf{I}(\mathbf{p}')$ is simply a requirement that a feasible power vector exists. The fixed point $\bar{\mathbf{p}}$ is a minimum power solution in that $\bar{\mathbf{p}} \leq \mathbf{p}'$ for any feasible power vector $\mathbf{p}'$.

We define the interference function $\mathbf{I}(\mathbf{p})$, which is valid for both unconstrained and constrained cases by a proper selection of $\mathcal{S}$, as

$$I_i(\mathbf{p}) = \frac{\gamma_i^*}{h_{ii}} \min_{\mathbf{X}_i \in \mathcal{S}} \frac{\sum_{j \neq i} p_j h_{ij} |\mathrm{tr}(\mathbf{X}_i^H \mathbf{s}_j \mathbf{a}_{ij}^\top)|^2 + \sigma^2 \mathrm{tr}(\mathbf{X}_i^H \mathbf{X}_i)}{|\mathrm{tr}(\mathbf{X}_i^H \mathbf{s}_i \mathbf{a}_{ii}^\top)|^2} \tag{4.7}$$

Note that the interference function for the unconstrained temporal-spatial filtering, $\bar{\mathbf{I}}(\mathbf{p})$, can be obtained from (4.7) by choosing $\mathcal{S} = \mathcal{C}^{G \times K}$ as

$$\begin{aligned} \bar{I}_i(\mathbf{p}) &= \frac{\gamma_i^*}{h_{ii}} \min_{\mathbf{X}_i \in \mathcal{C}^{G \times K}} \frac{\sum_{j \neq i} p_j h_{ij} |\mathrm{tr}(\mathbf{X}_i^H \mathbf{s}_j \mathbf{a}_{ij}^\top)|^2 + \sigma^2 \mathrm{tr}(\mathbf{X}_i^H \mathbf{X}_i)}{|\mathrm{tr}(\mathbf{X}_i^H \mathbf{s}_i \mathbf{a}_{ii}^\top)|^2} \\ &= \frac{\gamma_i^*}{h_{ii}} \min_{\mathbf{x}_i \in \mathcal{C}^{GK}} \frac{\sum_{j \neq i} p_j h_{ij} |\mathbf{x}_i^H \mathbf{q}_{ij}|^2 + \sigma^2 \mathbf{x}_i^H \mathbf{x}_i}{|\mathbf{x}_i^H \mathbf{q}_{ii}|^2} \end{aligned} \tag{4.8}$$

where the long vector notation as introduced in (3.3.6) is used in a multicell setting, i.e. $\mathbf{q}_{ij}$ is the combined temporal-spatial signature of user j at the base station of user i. The interference function for the constrained temporal-spatial filtering case, $\tilde{\mathbf{I}}(\mathbf{p})$, can be obtained from (4.7), by choosing $\mathcal{S} = \mathcal{L}$ as

$$\begin{aligned} \tilde{I}_i(\mathbf{p}) &= \frac{\gamma_i^*}{h_{ii}} \min_{\mathbf{X}_i \in \mathcal{L}} \frac{\sum_{j \neq i} p_j h_{ij} |\mathrm{tr}(\mathbf{X}_i^H \mathbf{s}_j \mathbf{a}_{ij}^\top)|^2 + \sigma^2 \mathrm{tr}(\mathbf{X}_i^H \mathbf{X}_i)}{|\mathrm{tr}(\mathbf{X}_i^H \mathbf{s}_i \mathbf{a}_{ii}^\top)|^2} \\ &= \frac{\gamma_i^*}{h_{ii}} \min_{\mathbf{c}_i, \mathbf{w}_i} \frac{\sum_{j \neq i} p_j h_{ij} |\mathbf{c}_i^H \mathbf{s}_j|^2 |\mathbf{w}_i^H \mathbf{a}_{ij}|^2 + \sigma^2 |\mathbf{c}_i|^2 |\mathbf{w}_i|^2}{|\mathbf{c}_i^H \mathbf{s}_i|^2 |\mathbf{w}_i^H \mathbf{a}_{ii}|^2} \end{aligned} \tag{4.9}$$

We make the following observation.

Proposition 2 $\mathbf{I}(\mathbf{p})$ *in (4.7) is a standard interference function.*

The proof of this Proposition is given in Appendix 4.A. Proposition 2 implies that the power control iteration of the form of (4.6) converges to the optimum power vector.

The resulting power control algorithm for both unconstrained and constrained temporal-spatial filtering cases are two-step iterative algorithms. In both cases, in the first step, the filter is found by solving the minimization problem in (4.7) and in the second step, the power of the user is updated using (4.6). Below we will state the resulting power control algorithms for unconstrained and constrained cases separately.

The implementation of the two step iterative power control algorithm in the case of unconstrained temporal-spatial filtering for user i at iteration $n+1$ is given by

Algorithm A1:

$$\hat{\mathbf{x}}_i = k_i(n)(\textstyle\sum_{j\neq i} p_j(n) h_{ij} \mathbf{q}_{ij}\mathbf{q}_{ij}^H + \sigma^2 \mathbf{I})^{-1}\mathbf{q}_{ii} \tag{4.10}$$

$$p_i(n+1) = \frac{\gamma_i^* \left(\sum_{j\neq i} p_j(n) h_{ij} |\hat{\mathbf{x}}_i^H \mathbf{q}_{ij}|^2 + \sigma^2 \hat{\mathbf{x}}_i^H \hat{\mathbf{x}}_i\right)}{h_{ii}|\hat{\mathbf{x}}_i^H \mathbf{q}_{ii}|^2} \tag{4.11}$$

where $k_i(n)$ is the scaling factor. Notice that any positive scalar multiple of $\hat{\mathbf{x}}_i$ yields the same power update in (4.11), so, calculation of $k_i(n)$ is actually not needed.

The implementation of the two step iterative power control algorithm in the case of constrained temporal-spatial filtering for user i at iteration $n+1$ is given by

Algorithm A2:

$$[\hat{\mathbf{c}}_i, \hat{\mathbf{w}}_i] = \arg\min_{\mathbf{c},\mathbf{w}} E[(\mathbf{c}^H \mathbf{R}_i(\mathbf{p}(n))\mathbf{w}^*) - b_i)^2] \tag{4.12}$$

$$p_i(n+1) = \frac{\gamma_i^* \left(\sum_{j\neq i} p_j(n) h_{ij} |\hat{\mathbf{c}}_i^H \mathbf{s}_j|^2 |\hat{\mathbf{w}}_i^H \mathbf{a}_{ij}|^2 + \sigma^2 (\hat{\mathbf{c}}_i^H \hat{\mathbf{c}}_i)(\hat{\mathbf{w}}_i^H \hat{\mathbf{w}}_i)\right)}{h_{ii}|\hat{\mathbf{c}}_i^H \mathbf{s}_i|^2 |\hat{\mathbf{w}}_i^H \mathbf{a}_{ij}|^2} \tag{4.13}$$

If the SIR targets are feasible, then starting from any initial power vector and filter coefficients, the algorithms in (4.11) and (4.13) converge to the minimum power fixed point with best possible temporal-spatial filters in the corresponding filter space.

Note that to implement (4.12) we need to use the iterative algorithm given by (3.43) and (3.44). The filters need to converge to the exact optimum $\hat{\mathbf{c}}_i, \hat{\mathbf{w}}_i$ pair for fixed powers before each power update for the power control algorithm to be standard. So, theoretically, many filter updates in the form of (3.43) and (3.44) have to be done before the power of the user is updated. In practice, a finite number, L, of iterations are performed which is in a way an approximation for the standard power control algorithm. Our observations is that a small number of iterations is sufficient. In all experiments, we used $L = 5$ iterations. Curiously, we have also observed the convergence of the following algorithm with $L = 1$ of each filter update per power iteration, and that the convergence point of the following algorithm and the algorithm A2 are the same (see Section 4.5):

Algorithm A3:

$$\mathbf{c}_i(n+1) = \text{MMSE}(\mathbf{w}_i(n)) \tag{4.14}$$

$$\mathbf{w}_i(n+1) = \text{MMSE}(\mathbf{c}_i(n+1)) \tag{4.15}$$

$$p_i(n+1) = \frac{\gamma_i^*}{h_{ii}} \frac{\sum_{j\neq i} p_j h_{ij} |\mathbf{c}_i^H \mathbf{s}_j|^2 |\mathbf{w}_i^H \mathbf{a}_{ij}|^2 + \sigma^2 |\mathbf{c}_i|^2 |\mathbf{w}_i|^2}{|\mathbf{c}_i^H \mathbf{s}_i|^2 |\mathbf{w}_i^H \mathbf{a}_{ii}|^2} \tag{4.16}$$

The intuition behind this algorithm is simple: before each power control update, each user chooses a *better*, not necessarily the best, filter pair. This is a mere consequence of the fact that in update (4.14), the SIR of user i is maximized by replacing the temporal filter with $\mathbf{c}_i(n+1)$, for the given power vector and $\mathbf{w}_i(n)$ and in update (4.16) the SIR is further increased by the maximization when the power vector and $\mathbf{c}_i(n+1)$ are fixed. Thus, $[\mathbf{c}_i(n+1), \mathbf{w}_i(n+1)]$ are a *better* filter pair than $[\mathbf{c}_i(n), \mathbf{w}_i(n)]$ for the power vector $\mathbf{p}(n)$. The simulation results about the performance and convergence of this algorithm are given in Section 4.5.

4.4 Simple Extensions

It was shown in [86] that standard power iterations in the presence of maximum power constraints are standard and thus convergent. So, it is possible to modify A1 and A2 to incorporate the maximum power constraints. In particular, the transmit power update steps (4.11), (4.13) can be modified respectively as follows

$$p_i(n+1) = \min\left\{p_{\max}, \bar{I}_i(\mathbf{p}(n))\right\} \tag{4.17}$$

$$p_i(n+1) = \min\left\{p_{\max}, \tilde{I}_i(\mathbf{p}(n))\right\} \tag{4.18}$$

where $\bar{I}_i(\cdot)$ and $\tilde{I}_i(\cdot)$ are defined in (4.8) and (4.9) respectively.

Another possible extension for the power control algorithms we proposed is to incorporate base station selection into the algorithms. Base station selection as a means of further interference suppression compared to fixed assignment combined with transmit power control has been addressed in [26, 87]. These works assumed a single antenna at each base station and conventional processing in time and found the best base station assignment that minimized the total power of all users. Similar to that case, by finding the optimum assignment of users to base stations, we can further decrease the total transmit power in the systems where temporal and spatial processing are used. Consider M base stations to which all users will be assigned. $\mathbf{b}$ is the N dimensional vector that denotes the assigned base station indices. In particular, $b_i = k$ means that user i is assigned to base k. In this case, the optimization problem is

$$\begin{aligned} \min_{\mathbf{p},\mathbf{b},\{\mathbf{X}_i\in\mathcal{S}\}} \quad & \sum_{i=1}^{N} p_i \\ \text{s.t.} \quad & p_i \geq \frac{\gamma_i^*}{h_{b_i i}} \frac{\sum_{j\neq i} p_j h_{b_i j} |\mathrm{tr}(\mathbf{X}_i^H \mathbf{s}_j \mathbf{a}_{b_i j}^\top)|^2 + \sigma^2 \mathrm{tr}(\mathbf{X}_i^H \mathbf{X}_i)}{|\mathrm{tr}(\mathbf{X}_i^H \mathbf{s}_i \mathbf{a}_{b_i i}^\top)|^2} \\ & p_i \geq 0, \quad b_i \in \{1,\cdots,M\}, \quad \mathbf{X}_i \in \mathcal{S}, \qquad i = 1,\ldots,N \end{aligned} \tag{4.19}$$

where, again, $\mathcal{S} = \mathcal{C}^{G\times K}$ for unconstrained temporal-spatial filtering and $\mathcal{S} = \mathcal{L}$ for constrained temporal-spatial filtering. We can once again move filter and base station optimization to the constraint set and define

$$T_i(\mathbf{p}) = \min_{\mathbf{b},\mathbf{X}_i\in\mathcal{S}} \frac{\gamma_i^*}{h_{b_i i}} \frac{\sum_{j\neq i} p_j h_{b_i j} |\mathrm{tr}(\mathbf{X}_i^H \mathbf{s}_j \mathbf{a}_{b_i j}^\top)|^2 + \sigma^2 \mathrm{tr}(\mathbf{X}_i^H \mathbf{X}_i)}{|\mathrm{tr}(\mathbf{X}_i^H \mathbf{s}_i \mathbf{a}_{b_i i}^\top)|^2} \tag{4.20}$$

Simply extending the proof we had for Proposition 2, we arrive at the following Proposition.

Proposition 3 $\mathbf{T}(\mathbf{p})$ *is a standard interference function.*

Thus, we can devise the algorithm where each user evaluates its SIR at each base station with the best corresponding temporal and spatial filters and then chooses the best base station. The transmit power of the user is then adjusted. This algorithm will converge to the best temporal and spatial filters in the intended filter space with the best base station assignment and transmit power for each user.

4.5 Simulation Results

We consider a 9-cell CDMA system on a 3×3 grid. We consider fixed base station assignment for simplicity. We assume a linear array of omni directional antennas equispaced at half a wavelength [81]. The positions of the users and their temporal signatures are generated at random, but then kept fixed for the particular experiment. The SIR target value is the same for all users and is set to $\gamma^* = 5$ (7dB). Results are generated to compare the following algorithms:

1. Conventional Power control (C-PC): Each base station has a single antenna and matched filter receivers are employed in the temporal domain [86, 94].

2. Power control and Multiuser Detection (MMSE-PC): Each base station has a single antenna and SIR maximizing (MMSE) receivers are employed in the temporal domain [67].

3. Power control and Beamforming (PC-BF-MF): An antenna array of K elements is employed at each base station. Array outputs are combined in MMSE sense. Matched filter receivers are employed in the temporal domain [56].

4. Power control with Constrained Optimum Temporal-Spatial filtering, CO-MMSE, (BF-MMSE-PC $c - wL$): Constrained temporal-spatial filtering is employed. We performed L=5 iterations of the algorithm given by (3.43), (3.44) before each power update and observed that the resulting filters $\mathbf{c}, \mathbf{w}$ converged to the global minimizer of the MSE function given by (3.40) in each case (**Algorithm A2**).

5. Power control with single step CO-MMSE (BF-MMSE-PC $c - w$): Constrained temporal-spatial filtering is employed, but only L=1 iteration of the algorithm given by (3.43), (3.44) is employed before each power update (**Algorithm A3**).

6. Power Control with Optimum Temporal-Spatial Filtering, O-MMSE, (J-MMSE): Joint unconstrained filtering in temporal and spatial domains is employed as given in (3.33) (**Algorithm A1**).

Figure 4.3 shows the comparison of total transmit power usage when there are $N = 12$ users in the system. An antenna array of $K = 2$ elements is used, and the processing gain is $G = 10$. For this small system, all power control algorithms are feasible, i.e. all users can achieve γ^*. However, we see that the joint spatial temporal algorithms (items 4, 5, 6 above) offer savings in total transmit power over the C-PC and the combined power control and MMSE filtering in one domain (items 2, 3 above). Compared to C-PC, the savings are as high as 7.2dB.

Next, we consider a highly loaded system with $N = 60$ users. The number of antenna array elements is $K = 2$ and processing gain is $G = 10$. In Figure 4.4, we see that only power control algorithms with joint processing in both domains are feasible (items 4, 5, 6 above). The system can support this many users only by utilizing the structure in both temporal and spatial domains in conjunction with power control. Figure 4.5 is a magnified version of Figure 4.4 which emphasizes the fact that the O-MMSE with power control offers more savings in total transmit power as compared to CO-MMSE with power control (see Section 3.3.7). This figure also emphasizes our observation about the convergence of the $L = 1$ algorithm implemented as in (4.14)-(4.16) to the optimal power vector with CO-MMSE.

Similar observations are obtained for a larger system with processing gain $G = 64$ and $K = 4$ antennas. Figure 4.6 shows the distribution of the $N = 300$ users in 9 cells for this experiment. The corresponding total transmit power curves are plotted in Figure 4.7. Once again, intelligent signal combining methods in both temporal and spatial domains used with optimal power control are superior to that of single domain combining with power control. C-PC and combined power control and beamforming are simply infeasible for this example. Figure 4.8 shows an even more crowded system with the same parameters and $N = 500$ users. For this example, only the temporal and spatial filtering with power control methods (items 4,5,6) are feasible, i.e. the system can support this many users only if all three interference management methods are combined as proposed in this chapter.

4.6 Chapter Summary

In this chapter, we have shown that when antenna arrays are employed at each base station, the system performance can be improved by jointly combining the array observations and the temporal observations and employing power control. The total transmit power expended by all users is less as compared to algorithms that do not utilize both temporal and spatial domains. In cases where other algorithms result in an infeasible system, power control with multiuser detection and beamforming can convert the system into a feasible one. Thus, it increases the system capacity by allowing the SIR

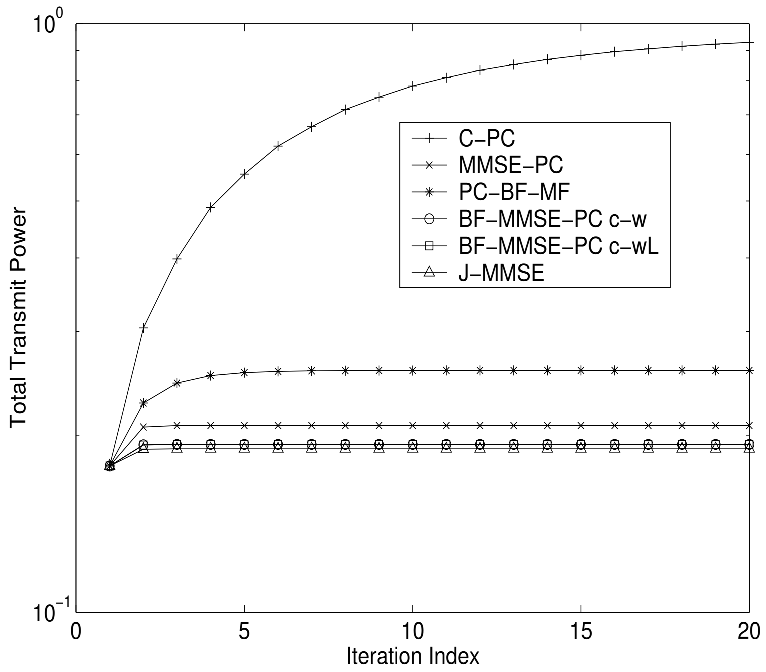

Figure 4.3: Total transmit power in the system versus power control iteration index. $N = 12, K = 2, G = 10$

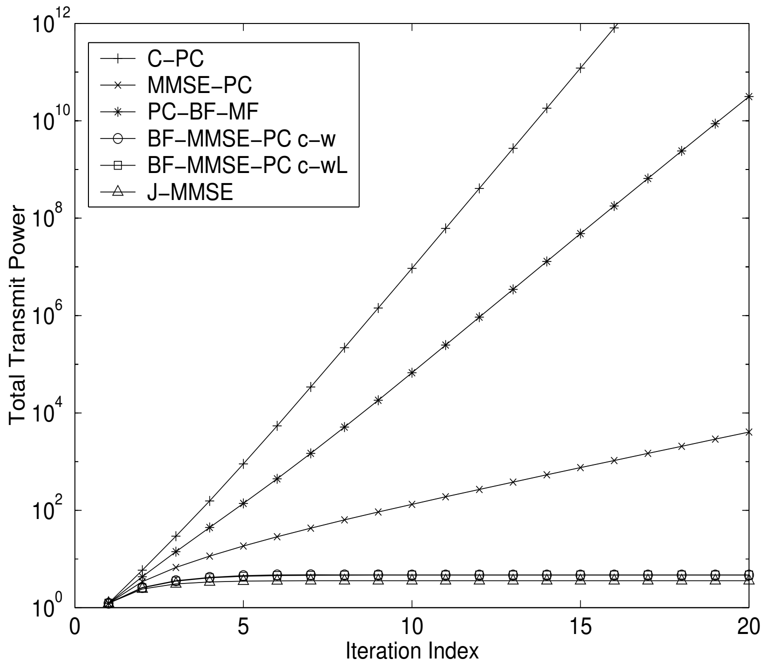

Figure 4.4: Total transmit power in the system versus power control iteration index. $N = 60, K = 2, G = 10$

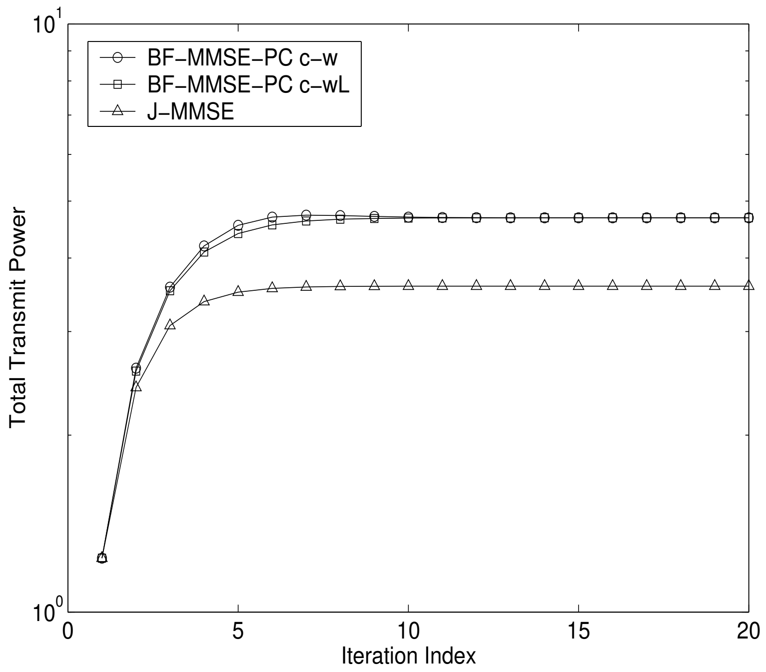

Figure 4.5: Total transmit power in the system versus power control iteration index. $N = 60, K = 2, G = 10$

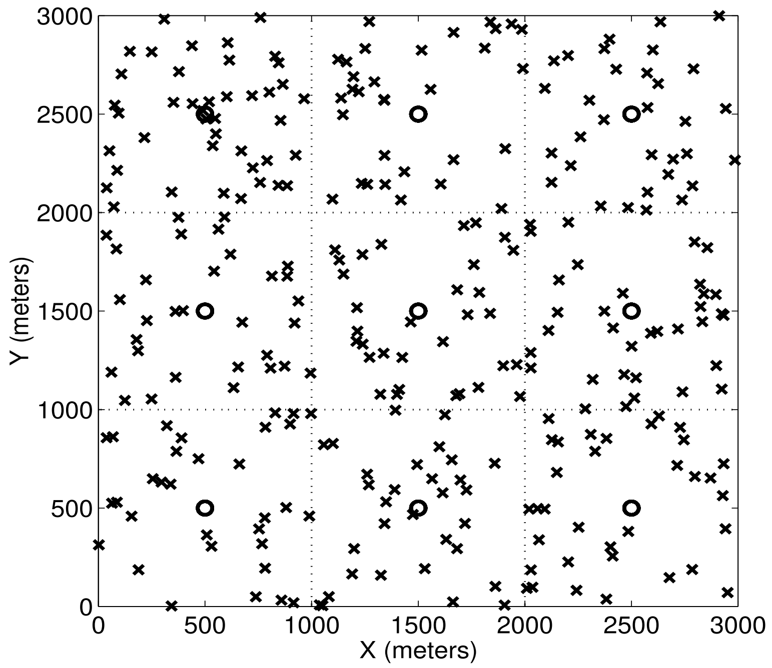

Figure 4.6: Distribution of $N = 300$ users on a 9-cell grid. x and o denote users and bases respectively.

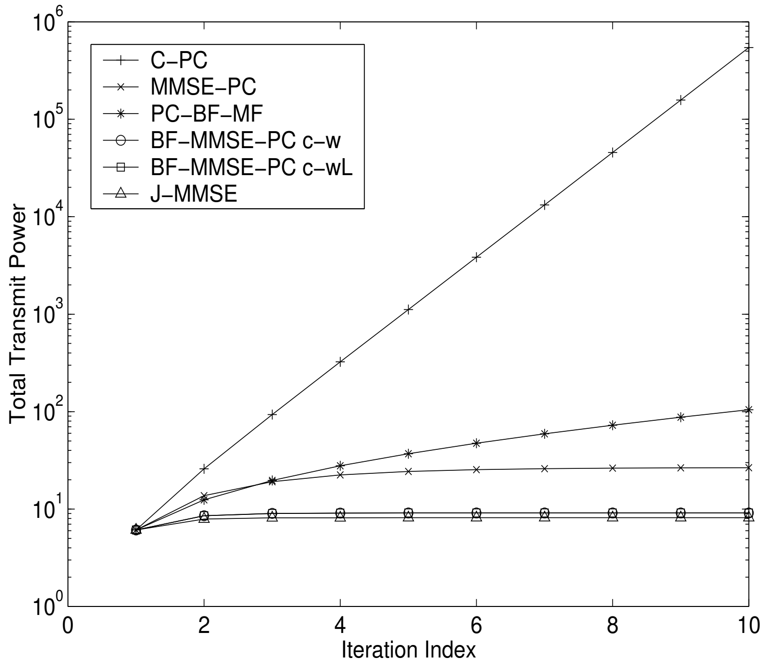

Figure 4.7: Total transmit power in the system versus power control iteration index. $N = 300, K = 4, G = 64$

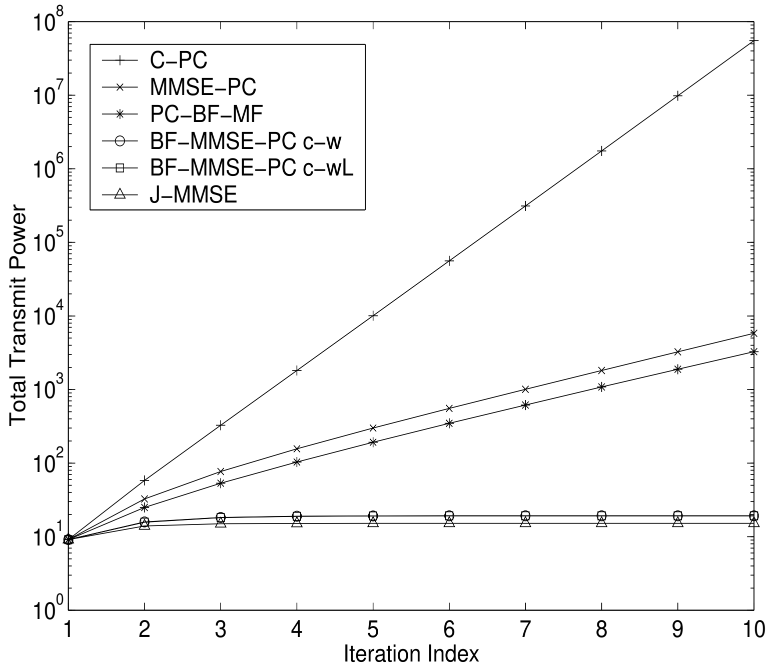

Figure 4.8: Total transmit power in the system versus power control iteration index. $N = 500, K = 4, G = 64$

targets of the users to be higher, or by increasing the number of users supportable at a fixed SIR target level.

One should note that the results we have presented assume knowledge of all users' parameters in the system, e.g. spreading codes, timing information, spatial signatures and link gains. Further, we have assumed perfect measurements of interference functions for the power control algorithms proposed. In practice, especially for out of cell interferers, all parameters may not be available to the system. In such cases, adaptive [36] or blind adaptive [29, 64] methods should be used to find the filters in power control iterations (4.10), (4.12), (4.14) and (4.15). The speed of these adaptive algorithms brings a natural constraint on the time frame in which the power updates can be done and consequently the effectiveness of power control. The same is valid for interference measurement speed and accuracy [70]. Thus, similar to [67] and [56], care must be exercised in choosing the practical filter update algorithms and interference measurement methods to ensure the maximal gain in employing the three interference management techniques, transmit power control, multiuser detection and receiver beamforming, jointly.

4.A Appendix

Proof: Proposition 2 We first define

$$J_i(\mathbf{p}, \mathbf{X}_i) = \frac{\gamma_i^*}{h_{ii}} \frac{\sum_{j \neq i} p_j h_{ij} |\mathrm{tr}(\mathbf{X}_i^H \mathbf{s}_j \mathbf{a}_{ij}^\top)|^2 + \sigma^2 \mathrm{tr}(\mathbf{X}_i^H \mathbf{X}_i)}{|\mathrm{tr}(\mathbf{X}_i^H \mathbf{s}_i \mathbf{a}_{ii}^\top)|^2} \tag{4.21}$$

Therefore,

$$I_i(\mathbf{p}) = \min_{\mathbf{X}_i \in \mathcal{S}} J_i(\mathbf{p}, \mathbf{X}_i) \tag{4.22}$$

To show that $\mathbf{I}(\mathbf{p})$ is standard, we need to check that the three conditions in Definition 1 are satisfied for (4.7). Similar to [67]:

- Positivity: For any fixed $\mathbf{X}_i \in \mathcal{S}$, we have $J_i(\mathbf{p}, \mathbf{X}_i) > 0$. Therefore, this is true for the minimizer filter also, i.e., $I_i(\mathbf{p}) = \min_{\mathbf{X}_i \in \mathcal{S}} J_i(\mathbf{p}, \mathbf{X}_i) > 0$

- Monotonicity: For any fixed $\mathbf{X}_i$, $\mathbf{p} \geq \mathbf{p}'$ implies $J_i(\mathbf{p}, \mathbf{X}_i) \geq J_i(\mathbf{p}', \mathbf{X}_i)$. If the minimum of $J_i(\mathbf{p}, \mathbf{X}_i)$ in $\mathcal{S}$ is achieved at $\mathbf{X}_i^*$, then,

$$\begin{aligned} I_i(\mathbf{p}) &= \min_{\mathbf{X}_i \in \mathcal{S}} J_i(\mathbf{p}, \mathbf{X}_i) && (4.23) \\ &= J_i(\mathbf{p}, \mathbf{X}_i^*) && (4.24) \\ &\geq J_i(\mathbf{p}', \mathbf{X}_i^*) && (4.25) \\ &\geq \min_{\mathbf{X}_i \in \mathcal{S}} J_i(\mathbf{p}', \mathbf{X}_i) = I_i(\mathbf{p}') && (4.26) \end{aligned}$$

- Scalability: For any fixed $\mathbf{X}_i$ and $\alpha > 1$ we have $\alpha J_i(\mathbf{p}, \mathbf{X}_i) > J_i(\alpha\mathbf{p}, \mathbf{X}_i)$. Again let $\mathbf{X}_i^*$ be the minimizer of $J_i(\mathbf{p}, \mathbf{X}_i)$ in $\mathcal{S}$. Then

$$\begin{aligned} \alpha I_i(\mathbf{p}) &= \min_{\mathbf{X}_i \in \mathcal{S}} \alpha J_i(\mathbf{p}, \mathbf{X}_i) && (4.27) \\ &= \alpha J_i(\mathbf{p}, \mathbf{X}_i^*) && (4.28) \\ &> J_i(\alpha\mathbf{p}, \mathbf{X}_i^*) && (4.29) \\ &\geq \min_{\mathbf{X}_i \in \mathcal{S}} J_i(\alpha\mathbf{p}, \mathbf{X}_i) = I_i(\alpha\mathbf{p}) && (4.30) \end{aligned}$$

$\square$

Chapter 5

Adaptive Cell Sectorization for CDMA

5.1 Introduction

Chapter 3 dealt with designing temporal-spatial filters geared towards detecting the bits of each user in a better way. Chapter 4 integrated transmit power control to a system employing these advanced receivers and investigated the performance gain. The systems considered in Chapters 3 and 4 require that a sophisticated receiver be designed for each user. There may be cases where such an architecture is not feasible to be implemented in the receiver due to stringent complexity constraints. In such cases, one might wonder if there are interference management methods that will enhance the system capacity while using the simplest receiver architecture possible. The motivation of this chapter is to investigate such a scheme.

A simple way of reducing the interference between users is to sectorize the cells using directional antennas. Sectorization simply employs directional antennas and each user's signal is received at only one of these antennas. Since only a subset of the users are received at each antenna, the interference that each user sees is less compared to a single antenna system. In contrast with receiver beamforming, no spatial combiner designed for each user. While this is a more *rigid* scheme than beamforming, the simplicity of the resulting receivers is appealing and sectorization can be quite beneficial especially for static systems, i.e. if the users do not have to handoff from sector to sector frequently.

It has been shown that sectorization increases the number of users admissible in a system, i.e. system capacity [14]. However, under highly non-uniform traffic loads, conventional sectorization which divides the cell into equal sectors, might fail to bring much capacity improvement. If information is available about the geographic distribution of the call traffic, it may be possible to design sectors such that maximum capacity gain can be achieved. Recently, fixed wireless systems, Wireless Local Loops (WLL), are posed as an alternative to conventional local wired telephone service [23]. Contrary to mobile macro-cellular systems, for fixed wireless systems (or systems with low mobility), users' whereabouts, and thus their uplink gains to the base station, do not change much with time and the information about the traffic load as a function of the position in the cell can be obtained at the base station. It is then possible to use this information to sectorize the cells such that the maximum coverage is achieved, i.e. the maximum number of subscribers get their required quality of service.

In this chapter, given the number of sectors and terminal locations, we study the best way of sectorizing the cell such that all subscribers in the cell meet their quality of service requirements while using as little power as possible. We call this combined power control and cell sectorization approach *adaptive cell sectorization* since the sector configurations are adaptive to the geographical load in the cell. We first consider a single cell system where we study the minimization of the total received power as well as the total transmit power. We show that in the received power optimization case, the best arrangement is such that the number of users in each sector is equalized (as much as possible) leading to a discrete waterfilling solution. The minimum total transmit power problem is formulated as a graph partitioning problem whose solution is found by solving a shortest path problem on a network of nodes constructed from the instance of the sectorization problem. We then formulate and solve the cell sectorization problem considering the effects of out-of-cell interference. We provide performance comparisons of solutions by both the received power and transmit power problems with the equal-sectors configuration implemented in current systems. We observe that power savings are achieved and power assignments that are infeasible with uniform sectorization can be handled with nonuniform sectorization. Thus, more users can be supported by the cell.

5.2 Model

We first consider a single cell DS-CDMA system with processing gain G and a total of N users. The multiple cell scenario is studied in Section 5.5. The user (call traffic)

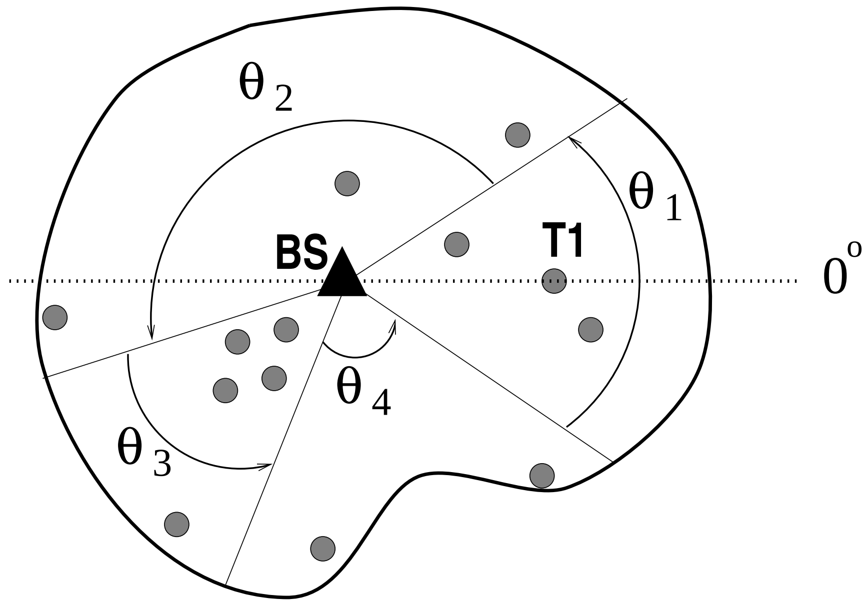

Figure 5.1: Sectorization of a cell.

density in the cell is a known function of the distance and angle (referenced to the origin) from the base station. In a WLL, this information is readily available through the addresses of the subscribers. Note that the user density can be highly non-uniform in which case partitioning the cell into equal width sectors as is done conventionally is not the best way of maximizing the number of users to be served by the cell. We assume the cell is to be partitioned into M sectors and the antennas have ideal directivity, i.e. there is no interference between the sectors.

Our aim is to find the best sectorization arrangement such that all users get their required quality of service while expending as little power as possible. Figure 5.1 shows an example of a sectorization configuration which is characterized by the reference point (0° angle) and the vector containing sector width angles, $\boldsymbol{\theta}$. In each sector, all users belonging to that sector interfere with each other. We assume that each user has a pseudo-random signature sequence and single user receivers, i.e., matched filters, are employed at the base station. We consider a synchronous system for simplicity. The quality of service measure is the signal to interference ratio (SIR). A user has reliable communication if its SIR at its assigned sector antenna is larger than a common target SIR, γ^*.

5.3 Received Power Optimization

We first consider the optimization problem of minimizing total received power at the base station by choosing the transmit power for each user and a sector partition. Although the focus of this study is the *transmit* power minimization problem, we first analyze the received power minimization problem for reasons that will become clear by the end of this section.

We consider the problem of optimally placing N terminals into M sectors such that the total power received at the base station is minimized while ensuring each user maintains the target SIR of γ^*. Let $s_i(\boldsymbol{\theta})$ denote the set of users that reside in the area spanned by sector i, and $\boldsymbol{\theta}$ denote the M-tuple vector that contains the sector angles. Since sectorization creates disjoint subsets of the set of terminals, each terminal belongs to only one of the sets $s_i(\boldsymbol{\theta})$ where $i = 1, \cdots, M$. We denote the sector index that user j belongs to by a_j. The power expended by terminal j is p_j which is received after being attenuated by its channel gain h_j. The SIR of terminal j at the output of its matched filter can be expressed as

$$\gamma_j = \frac{p_j h_j}{\sum_{k \neq j} \delta_{kj} p_k h_k / G + \sigma^2} \tag{5.1}$$

where σ^2 is due to AWGN and δ_{kj} is defined as,

$$\delta_{kj} = \begin{cases} 1 & \text{if } a_j = a_k \\ 0 & \text{otherwise} \end{cases} \tag{5.2}$$

Notice that each terminal experiences interference from the terminals in its own sector only. Although the δ_{kj} notation seems complicated, it will be convenient for the graph theoretic formulation in Section 5.4.2.

The received power optimization problem can be expressed as

$$\begin{aligned} \min_{\boldsymbol{\theta}, \mathbf{p}} \quad & \sum_{i=1}^{M} \sum_{j \in s_i(\boldsymbol{\theta})} p_j h_j \qquad\qquad \textbf{(RP)} \\ \text{s.t.} \quad & \gamma_j = \frac{p_j h_j}{\sum_{k \neq j} \delta_{kj} p_k h_k / G + \sigma^2} \geq \gamma^* \\ & j = 1, ..., N \\ & \mathbf{p} \geq \mathbf{0} \qquad \mathbf{1}^T \boldsymbol{\theta} = 2\pi \end{aligned} \tag{5.3}$$

where $\mathbf{0}$ and $\mathbf{1}$ denote the all zero and all 1 vectors, respectively. Note that the constraints represent each user's reliable connection requirement.

It is not hard to see that the SIR constraints ought to be satisfied with equality at the solution of (5.3). Note also that within each sector, minimum power is achieved by assigning equal received powers to all users belonging to the sector. Denote the received power in sector i, q_i^*, such that $q_i^* = p_j h_j$ for all users j in sector i. Given the above observations, we can express q_i^* as

$$q_i^* = \frac{\gamma^* \sigma^2}{1 - \frac{\gamma^* N_i(\boldsymbol{\theta})}{G}} \tag{5.4}$$

where $N_i(\boldsymbol{\theta}) = |s_i(\boldsymbol{\theta})|$ is the number of users in sector i, i.e. the cardinality of set $s_i(\boldsymbol{\theta})$. Given (5.4), we can rewrite (5.3) as:

$$\begin{aligned}
\min_{\boldsymbol{\theta}, \mathbf{q}^*} \quad & \sum_{i=1}^{M} N_i(\boldsymbol{\theta}) q_i^* \\
\text{s.t.} \quad & q_i^* = \frac{\gamma^* \sigma^2}{1 - \frac{\gamma^* N_i(\boldsymbol{\theta})}{G}} \qquad i = 1, ..., M \\
& \mathbf{q}^* \geq \mathbf{0} \qquad \sum_{i=1}^{M} N_i(\boldsymbol{\theta}) = N \qquad \mathbf{1}^T \boldsymbol{\theta} = 2\pi
\end{aligned} \tag{5.5}$$

where $\mathbf{q}^*$ is the M-vector that contains the received power values associated with each sector. Optimization problem (5.5) is equivalent to

$$\begin{aligned}
\min_{\boldsymbol{\theta}} \quad & G \sum_{i=1}^{M} \frac{\sigma^2 \frac{\gamma^* N_i(\boldsymbol{\theta})}{G}}{1 - \frac{\gamma^* N_i(\boldsymbol{\theta})}{G}} \\
& N_i(\boldsymbol{\theta}) < G/\gamma^* \qquad i = 1, ..., M \\
& \sum_{i=1}^{M} N_i(\boldsymbol{\theta}) = N \qquad \mathbf{1}^T \boldsymbol{\theta} = 2\pi
\end{aligned} \tag{5.6}$$

Substituting $N_M(\boldsymbol{\theta}) = N - \sum_{i=1}^{M-1} N_i(\boldsymbol{\theta})$ and taking derivatives with respect to N_j reveals that the optimum point has to satisfy

$$\frac{1}{1 - \frac{\gamma^*}{G} N_j(\boldsymbol{\theta})} = \frac{1}{1 - \frac{\gamma^*}{G}(N - \sum_{i=1}^{M-1} N_i(\boldsymbol{\theta}))} \tag{5.7}$$

for all $j = 1, \cdots, M$. Note that we have treated the number of users in each sector as continuous variables which is a relaxation. Since the space of integers is a subset of the continuous solution space, finding a set of integer values that satisfy (5.7) for all j guarantees that we achieve the solution of (5.6). If we assume $N = kM$ where k is an integer, then (5.7) implies that the sectors should be arranged, i.e. $\boldsymbol{\theta}$ should be determined, such that

$$N_j(\boldsymbol{\theta}) = N/M \qquad j = 1, ..., M \tag{5.8}$$

Thus, to minimize the total power received at the base station while all users get their minimum required quality of service, the cell should be sectorized such that each sector has the same number of active users. Clearly, any set of $\boldsymbol{\theta}$ for which each sector covers N/M users is optimum. So, starting from one reference point, the sectorization can be arranged such that each sector is terminated when a coverage of N/M users is reached.

Note that the generalization of the result to N/M not an integer is immediate in the sense that the *left over* users after $k = \lfloor N/M \rfloor$ are placed in each sector, should be distributed evenly to the sectors. Thus, as a result, some sectors have to serve $k + 1$ users while some serve k. The total interference is placed as equally as possible which is the discrete water-filling solution.

One would consider that designing sector boundaries such that each sector contains an equal number of terminals is an *intuitive* sectorization method. In this section, we have shown that indeed such a method has an analytical basis and emerges as a solution to the received power minimization problem. We will use this result to compare results from sectorization by transmit power minimization.

5.4 Transmit Power Optimization

5.4.1 Problem Statement

Using the notation introduced in Section 5.3, the transmit power optimization problem can be stated in a similar fashion as

$$\begin{aligned}
\min_{\boldsymbol{\theta},\mathbf{p}} \quad & \sum_{i=1}^{M} \sum_{j\in s_i(\boldsymbol{\theta})} p_j \qquad\qquad \textbf{(TP)} \\
\text{s.t.} \quad & \gamma_j = \frac{p_j h_j}{\sum_{k\neq j} \delta_{kj} p_k h_k / G + \sigma^2} \geq \gamma^* \\
& j = 1, ..., N \\
& \mathbf{p} \geq \mathbf{0} \qquad \mathbf{1}^T \boldsymbol{\theta} = 2\pi
\end{aligned} \tag{5.9}$$

The observations that led to (5.4), i.e. (i) SIR constraints are satisfied with equality at the optimum, and (ii) Within each sector, minimum power assignment is such that each user in that sector is received with the same power at the base station, also hold for (5.9). Thus, we can rewrite (5.9) as

$$\begin{aligned}
\min_{\boldsymbol{\theta},\mathbf{q}^*} \quad & \sum_{i=1}^{M} q_i^* \sum_{j\in s_i(\boldsymbol{\theta})} \frac{1}{h_j} \\
\text{s.t.} \quad & q_i^* = \frac{\gamma^* \sigma^2}{1 - \frac{\gamma^* N_i(\boldsymbol{\theta})}{G}} \qquad i = 1, ..., M \\
& \mathbf{q}^* \geq \mathbf{0} \qquad \sum_{i=1}^{M} N_i(\boldsymbol{\theta}) = N \qquad \mathbf{1}^T \boldsymbol{\theta} = 2\pi
\end{aligned} \tag{5.10}$$

which is equivalent to

$$\begin{aligned}
\min_{\boldsymbol{\theta},N_i} \quad & \sum_{i=1}^{M} \frac{\gamma^* \sigma^2}{1 - \frac{\gamma^* N_i(\boldsymbol{\theta})}{G}} \sum_{j\in s_i(\boldsymbol{\theta})} \frac{1}{h_j} \qquad\qquad \textbf{(TP)} \\
& 1 - \frac{\gamma^* N_i(\boldsymbol{\theta})}{G} \geq \mathbf{0} \qquad \sum_{i=1}^{M} N_i(\boldsymbol{\theta}) = N \qquad \mathbf{1}^T \boldsymbol{\theta} = 2\pi
\end{aligned} \tag{5.11}$$

We should note that the joint sector assignment and transmit power control problem is different from those that can be unified by the theory of standard interference functions described in [86]. In particular, although an iterative solution to the transmit power control problem under a specific sectorization arrangement is possible using the framework in [86], this is not the case where an optimization over the sectorization arrangements is added to the problem. In this case, the problem becomes similar to joint optimal power control and frequency/time slot assignment and is reminiscent of bin packing problems [22]. Fortunately, it is possible to formulate the joint sector assignment and transmit power control problem as a graph theoretic problem that is solvable in polynomial time. In particular, we observe that optimization problem (5.11) can be formulated as a graph partitioning (GP) problem. We also observe that because of the special topology of the problem at hand, this problem is polynomially solvable [61, 62]. We will first discuss how the physical wireless system entities can be represented in a graph setting. The equivalent graph theoretic problem is then defined and the solution is presented. We give an algorithm that solves the TP problem. We then discuss the constrained problem where sector width is upper and lower bounded and sector boundaries are not allowed between close terminals. Finally, sectorization in a multicell system is examined in Section 5.5.

5.4.2 Graph Theoretic Formulation of TP

TP does not have a closed form solution. However, through appropriate transformations, it can be formulated as a graph partitioning problem for which we present an algorithm that is guaranteed to find the optimal solution in polynomial time.

First, label all the terminals within a cell by assigning an angular and radial coordinate to each one of them. The radial position is the distance of the terminal to the base station. The angular position is the angular distance of the terminal to a reference terminal. Without loss of generality, the reference terminal can be selected arbitrarily and assigned 0° as its angular position. Each terminal within the cell is represented by a vertex (node) along a ring. Let the position of the vertex on the ring be determined by the angular position of the terminal regardless of its radial position. We assume that all angular positions are unique. The radius of the ring is independent of the radial position of the users. The segment of the ring that connects each pair of neighboring nodes is called an edge. Figure 5.2 demonstrates the relation between the physical location of terminals and the corresponding nodes along a ring.

The topology just described constitutes a ring $G = (V, E)$, with vertices $V = \{v_1, v_2, \cdots, v_N\}$ and edges $E = \{(v_1, v_2), \cdots, (v_{N-1}, v_N)\} \cup \{(v_N, v_1)\}$. Denote element (v_i, v_j) of the edge set by e_i. Let the weight of vertex v_i be denoted by w_i and the weight of edge e_i be denoted by c_i. Vertex i is assigned weight $w_i = \frac{1}{h_i}$ where h_i is the uplink gain of the ith terminal. The weight of edge i, c_i, is the angular distance between the two vertices connected by this edge.

It is worth noting that two or more randomly placed terminals sharing the same angular location is a probability zero event. Nevertheless, the model is able to handle in case such a degenerate situation actually occurs. If there exist two such terminals, the ring is modified by combining these terminals to form a single vertex and assign a weight that is the summation of the weights of the original nodes. This reduction can be done with no effect on the optimal sectorization outcome since it is physically not possible to place users sharing the same angular position into different sectors.

The sectorization problem for minimum transmitted power (TP) solves for an optimal set of sector angles that identify sector boundaries. The set of terminals in the cell are optimally grouped into disjoint subsets. Graph theoretic methods emerge as powerful tools to find the best partition.

In the next section, we will present general results in graph partitioning. These results apply to optimal partitioning on a string which is a graph where all vertices are located along a line. An immediate observation is that a ring *broken* at an edge results in a string as shown in Figure 5.3(a) and 5.3(b). We will use this fact in Section 5.4.4.

5.4.3 Optimal Partitioning and Shortest Path Algorithms

Given a graph $G = (V, E)$, let $\pi = \{S_1, \cdots, S_M\}$ be a *partition* of the set of vertices V into M subsets. π is considered feasible in G if for every $k \in \{1, \cdots, M\}$, the subgraph $G(S_k)$ induced by S_k is connected and the weight $W(S_k)$ of S_k as defined below is greater than or equal to zero. The set of all such feasible partitions is denoted by $\Pi(G, M)$. Note that in a feasible partition, $S_k \neq \emptyset$ for all k. The cost function $C : \Pi(G, M) \rightarrow \Re$ assigns a value to each feasible partition and the cost of a particular partition $\pi = \{S_1, \cdots, S_M\}$ is given by

$$C(\pi) = \sum_{i=1}^{M} W(S_i) \tag{5.12}$$

The graph partitioning problem (GP) is to find a feasible partition such that some given cost is minimized. Mathematically, it can be expressed as,

$$\min_{\pi=\{S_1,\cdots,S_M\}\in\Pi(G,M)} \qquad C(\pi) = \sum_{i=1}^{M} W(S_i) \qquad \textbf{(GP)} \tag{5.13}$$

We now show that the graph partitioning problem GP is polynomially solvable under certain conditions. We first observe that the cost function of GP given in (5.13) is separable. This fact follows from the definition of a separable function given below.

Definition 2 *A function of M variables $f(x_1, x_2, \cdots, x_M)$ is said to be separable if it can be expressed as a sum of M functions of a single variable; i.e. $f(x_1, x_2, \cdots, x_M) = \sum_{i=1}^{M} f_i(x_i)$.*

Although the problem of optimally partitioning an arbitrary graph with an arbitrary cost function is an NP-hard optimization problem, partitioning a string optimally with a separable cost function can be solved in polynomial time. We use the following optimal graph partitioning result presented in [62].

Theorem 4 *If the objective function is separable, the problem of optimally partitioning a string can be reduced to a shortest path problem with computational complexity $\mathcal{O}(N^2M)$.*

Proof: A full proof can be found in [62]. A sketch of the proof is as follows. The main idea is to establish a one-to-one correspondence between the cost of a feasible partition in a string and the cost of a path from the origin to destination in a special acyclic network obtained from the partitioning problem at hand. An *acyclic network* is a directed graph with a *length* associated with each edge and that has no paths that start and end at the same node.

Suppose H is a string with N vertices labeled $v_1, \cdots, v_N$ so that adjacent vertices have consecutive indices and let the weight of the vertices be denoted as $w_1, \cdots, w_N$.

An edge e_i exists between vertex v_i and v_{i+1} and has weight c_i. H is to be partitioned into M subsets such that a separable objective function is optimized. Suppose that in any feasible partition $\{S_1, \cdots, S_M\}$, adjacent subsets have consecutive indices. Given the string and the problem, construct an acyclic network Q as follows. Let $(0,0)$ and (M,N) be the origin and the destination nodes in Q, respectively. The nodes that lie between the origin-destination pair is given by the set

$$\{(k,i) : 1 \leq k \leq M-1; k \leq i \leq N-M+k\} \tag{5.14}$$

An edge is placed *from* node (k,i) *to* (h,j) if $h = k+1$ and $j > i$. Otherwise, there is no edge between (k,i) and (h,j). There is a one-to-one correspondence between the feasible partitions of H into M subsets and the paths from $(0,0)$ to (M,N) in Q. The path P corresponding to a given $\pi \in \Pi(H,M)$ is defined as follows: The node (k,i) belongs to P if and only if i is the last vertex of the subset S_k of π. Furthermore, when the weight of edge u in Q is assigned the cost of the corresponding subset in H, the cost of partition π, $C(\pi)$ in (5.13) coincides with the total length of path P in network Q. It follows that the optimum partition of H can be determined by finding a shortest path from $(0,0)$ to (M,N) in Q since each path corresponds to one and only one feasible partition.

□

Figure 5.3(b) and 5.3(c) show how the network Q is obtained from the GP problem with G, $N = 5$ nodes and $M = 3$ partitions. Notice that as a result of the construction of the network, no empty sets are allowed.

A shortest path from origin to destination can be found in $\mathcal{O}(N^2M)$ by using well known shortest path algorithms. Dijkstra's method that finds the shortest path from an origin node to the destination node in a network with positive edge lengths is particularly suitable for use in our case [7, 12, 33, 34].

5.4.4 Graph Theoretic Solution of TP

Let us return to the discussion of TP and its solution in the graph theoretic framework. First, we observe that the cost function of TP given in (5.11) is separable, i.e. the cost of a sectorization arrangement in TP can be expressed as a sum of the costs of the individual sectors. Let $S_i \subset V$ represent the set of users that is in the ith sector in a given partition. We define the cost of sector S_i by

$$W(S_i) = \frac{\gamma^* \sigma^2}{1 - \frac{\gamma^* |S_i|}{G}} \sum_{k: v_k \in S_i} w_k = \frac{\gamma^* \sigma^2}{1 - \frac{\gamma^* |S_i|}{G}} \sum_{k: v_k \in S_i} \frac{1}{h_k} \tag{5.15}$$

where $|S_i|$ is the number of vertices in S_i. Then, the total cost, which is the summation of $W(S_i)$ over all sectors, has the same structure as the cost function (5.13) for GP. In addition, each feasible partition π corresponds to a set of cell sectorization configurations $\{s_i(\boldsymbol{\theta})\}$, both of which have the same cost value. Therefore, it follows that a solution for GP is also a solution for TP. Notice that the graph theoretic formulation returns the optimal sectors in terms of a list of terminals that belong to each sector. However, it does not say anything about where the actual sector boundary is located. The choice of the actual sector boundary is flexible since it can lie anywhere between the closest two terminals that belong to different sectors. The location of the boundary between neighboring vertices might have practical significance when considering physical constraints as explained in Section 5.4.5, however it has no impact on the cost function and the optimization given in (5.13).

So far, we established that TP is actually equivalent to GP with a separable objective function. Recall from Theorem 4 that GP with a separable objective function can be reduced to a shortest path problem on a string. We observe that the graph corresponding to TP is a ring instead of the string that Theorem 4 applies for. However, as mentioned in Section 5.4.2, if the ring in TP is *broken* at an edge (by removing the edge), the resulting graph is a string. Thus, with some increase in computational complexity, we can solve TP by solving a number of shortest path problems. Recall that the computational complexity of solving for the best connected M partition is $\mathcal{O}(N^2 M)$. Repeating the same procedure for each edge in E, and choosing the best among these solutions, we end up with the optimal solution for TP which has complexity $\mathcal{O}(N^3 M)$. The sectors are then formed based on this optimal partition.

5.4.5 Physical Constraints

In this section, we discuss how to incorporate certain physical constraints into the original TP problem. These physical constraints can be accommodated easily in the graph theoretical formulation of TP and do not bring any additional computational complexity to the solution algorithm. They do however have practical significance from a system point of view which is why we address them here. The weight c_i for edge e_i was defined earlier in Section 5.4.2 as the angular distance between the two vertices connected by e_i. However, for the results presented up to here, the edge weights had no impact on the optimum solution of the problem. In what follows, physical constraints we consider do appear in the form of constraints on the edge weights and thus affect the optimum solution.

The three constraints we consider are (i) minimum allowable angular distance between two users that belong to two neighboring sectors, (ii) maximum allowable sector width and (iii) minimum allowable sector width. The first constraint is needed to ensure that two users that are physically very close to each other are not assigned to different sectors, since accomplishing such precise beamforming could be difficult. We modify the vertex set such that vertices that are "very close" to each other are combined to form a supernode. This ensures that two such nodes are not placed in different sets in a partition. The minimum angular distance between two users in neighboring sectors, $c_{\min}$, is a system design parameter. The edges with weight, $c_i < c_{\min}$ are deleted from the edge set and the two neighboring nodes are combined to form a supernode which is placed midway along the edge that is deleted. The weight of the supernode and the edges that connect it to the rest of the ring are adjusted appropriately.

Next, let us turn our attention to the next two constraints, maximum and minimum sector size. Suppose the maximum and minimum allowable sector width is denoted by $\theta_{\max}$ and $\theta_{\min}$, respectively. We require that $\theta_{\min} < \theta_i < \theta_{\max}$ for all sectors i. This constraint can be incorporated in the solution described in Section 5.4.3 by modifying the way edges are assigned in the network. An infinite weight is assigned to edges in the network that correspond to sectors that are wider or narrower than the specified limit. Thus, while searching for a shortest path, such sectors do not emerge as possible solutions. The rest of the procedure of finding the optimal partition remains the same.

Recall that without constraints, the optimal sector boundaries are flexible, i.e. as long as each sector contains the optimal set of users, the exact locations of sector boundaries are not important. However, with constraints present, the angular space between neighboring terminals becomes significant and hence boundaries become less flexible. It is also worth noting that feasibility issues arise with sector size constraints. It may be the case that there exists no sectorization configuration satisfying a particular size constraint. For example, it is not possible to find 3 sectors each less than $120°$.

5.5 Multicell Environment

In this section, optimal cell sectorization and transmit power assignment for a cell in a multicell system is considered. The scenario we consider is appropriate for enhancing the capacity of a highly loaded cell, a *hot spot*, in the network. It is assumed that the transmit power values for out-of-cell interferers are assigned and do not change significantly as a result of the sectorization of the hot spot cell. Suppose a cell with N users is to be divided into M sectors where L users interfere from outside the cell. Define the out-of-cell interference (OCI) profile, $I(\phi)$, as a discrete function that gives the value of the interference power seen at the base station of the cell to be optimally sectorized. It is a function of ϕ, the *direction of interference*; i.e. the value of the OCI function at ϕ_k is denoted by $I_k = I(\phi_k)$ and is the contribution of the kth out of cell interferer which has angular distance ϕ_k to the cell to be optimally sectorized. Thus, the OCI function has nonzero value only at a set of angle values ϕ_k where $k = 1, \cdots, L$. Figure 5.4 shows a sample OCI profile. We will assume that $I(\phi)$ is known at the base station of the hot spot cell.

The level of OCI that any given user in the hot spot cell depends on the area covered by its sector. In particular, the OCI that user $j \in s_i$ experiences in sector i which covers a region from $\alpha^{i-1} = \sum_{j=1}^{i-1} \theta_j$ to $\alpha^i = \sum_{j=1}^{i} \theta_j$ can be expressed as

$$I^i = \sum_{\{k|\alpha^{i-1} \leq \phi_k \leq \alpha^i\}} I_k \tag{5.16}$$

Recall that a_j denotes the index of the sector that user j belongs to. For any two users $j, k \in s_i$ in the same sector, the OCI are $I^{a_j} = I^{a_k} = I^i$.

Note that, since the OCI is a function of the sector width, optimal cell boundaries are no longer flexible as is the case in the single cell scenario. To limit the infinite degree of freedom in selection of sector boundary locations, we constrain the sector boundaries to lie only midway between any pair of neighboring terminals. This limitation will render the optimization problem manageable.

The transmit power optimization problem in the presence of OCI can be formulated as

$$\begin{aligned}
\min_{\boldsymbol{\theta},\mathbf{p}} \quad & \sum_{i=1}^{M} \sum_{j \in s_i(\boldsymbol{\theta})} p_j \qquad\qquad \textbf{(MTP)} \\
\text{s.t.} \quad & \gamma_j = \frac{p_j h_j}{\sum_{k \neq j} \delta_{kj} p_k h_k / G + I^{a_j} + \sigma^2} \geq \gamma^* \\
& j = 1, ..., N \\
& \mathbf{p} \geq \mathbf{0} \qquad \mathbf{1}^T \boldsymbol{\theta} = 2\pi
\end{aligned} \tag{5.17}$$

Notice that the OCI contributes the additional interference term in the denominator of the SIR in the above equation. Once again, observe that at the optimum solution of MTP, $\gamma_j = \gamma^*$ for all users and users belonging to the same sector should have equal received power at the base station. Then, (5.17) can be rewritten as,

$$\begin{aligned}
\min_{\boldsymbol{\theta},N_i} \quad & \sum_{i=1}^{M} \frac{\gamma^*(\sigma^2 + I^i)}{1 - \frac{\gamma^* N_i(\boldsymbol{\theta})}{G}} \sum_{j \in s_i(\boldsymbol{\theta})} \frac{1}{h_j} \qquad\qquad \textbf{(MTP)} \\
& 1 - \frac{\gamma^* N_i(\boldsymbol{\theta})}{G} \geq \mathbf{0} \qquad \sum_{i=1}^{M} N_i(\boldsymbol{\theta}) = N \qquad \mathbf{1}^T \boldsymbol{\theta} = 2\pi
\end{aligned} \tag{5.18}$$

A comparison between (5.11) and (5.18) reveals the difference between the optimization problem in a single cell and a multicell scenario. With OCI present, the total interference experienced in sector i is $\sigma^2 + I^i$ as opposed to the AWGN power of σ^2 when there is no out-of-cell interference. The alternative expression of the MTP in (5.18) allows us to observe that the formulation of the problem as a graph partitioning problem is still possible with slight changes to the methodology already explained in Section 5.4. In particular, the weight of a subset S of the set of vertices V should be modified as follows:

$$W(S) = \frac{\gamma^*(\sigma^2 + I^S)}{1 - \frac{\gamma^*|S|}{G}} \sum_{k:v_k \in S} w_k \tag{5.19}$$

where $w_k = \frac{1}{h_k}$ and I^S is the OCI associated with subset S. It remains to calculate I^S for each subset S in any given feasible partition. This calculation from (5.16) brings an additional computational cost to the GP problem, but the complexity of the problem remains polynomial.

In the next section, we will provide examples both for the single cell system and for a hot spot cell in a multicell setting and discuss the results from uniform and optimum sectorizations.

5.6 Numerical Results

We provide sectorization results by three methods: transmit power optimization, received power optimization, and uniform sectorization. The output of the algorithm given in Section 5.4.3 is the optimal partitioning (OP) that minimizes sum of transmit powers of all users in the cell. Sectorization that places equal number of terminals in each sector (EP) is the minimum received power result. Finally, in uniform sectorization (UP), sectors are of equal width and their orientation does not depend on the terminal locations.

First, we consider the single cell system. Figure 5.5 shows sector boundaries obtained by OP and EP. For this particular example, UP is not feasible, i.e. the system cannot support acceptable SIR to all users in the cell if uniform sectorization of the cell is employed. The adaptive sectorization enables the system to support all users. Recall that OP has complexity $\mathcal{O}(N^3M)$. Since EP is much easier to implement, one might wonder how EP performs in terms of total transmit power. We have compared the total transmit power values for all three methods in Table 5.1. It can be seen that, for this example, OP and EP result in quite different total transmit power values and OP

	Figure 5.5	Figure 5.6
TTP(OP)	6.3	3.9
TTP(EP)	9.1	4.2
TTP(UP)	Infeasible	5.8

Table 5.1: Total transmit powers [Watts] obtained from three methods.

	Figure 5.7 (no OCI)	Figure 5.8 (with OCI)
TTP(OP)	5.69	6.54
TTP(UP)	11.9	14.1
% improvement	52	54

Table 5.2: Total transmit powers [Watts] obtained from OP and UP in the multicell scenario.

brings 40% improvement in total transmit power (TTP) as compared to EP. This is an example where finding OP really pays off in terms of transmit power savings.

Next, we consider a scenario where the distribution of the terminals are uniform in area within the cell. The resulting sectorization arrangements are shown in Figure 5.6. For this example, the total transmit power of OP (TTP(OP)) is about 32% less than the total transmit power of UP (TTP(UP)). However, OP's improvement over EP is only 7% (Table 5.1). In general, in case of uniformly distributed terminals, TTPs with OP, EP and UP are expected to be close to each other. An extreme case where all three methods yield the same TTP is when the terminals lie along the perimeter of a circle at uniform angular spacing.

In the multicell scenario, we study sectorization of a center cell that is surrounded by interfering cells on a 9 cell grid. We examine the worst case scenario by assigning maximum transmit power to all the out-of-cell interferers. Maximum transmit power level is taken to be 1 Watt. We compare uniform sectorization with optimal sectorization obtained by solving the MTP problem explained in Section 5.5. Figure 5.7 shows sector boundaries for uniform sectorization (UP) and optimal sectorization (OP) when no out-of-cell interferers are active. Figure 5.8 shows UP and OP boundaries for the same system when all out-of-cell terminals transmit at maximum power. Notice that due to the uniform distribution of the out-of-cell interferers, the boundaries do not change much from one case to the other. In fact, only one user is shifted to a different sector. Total transmit power values are compared in Table 5.2. It is observed that although the total transmit powers increase in the presence of OCI, the result obtained by OP is still substantially better than that provided by UP. OP is clearly superior to the conventional method of UP. In fact, the improvement in terms of total transmit power in the case with OCI is more emphasized than the case without it.

5.7 Chapter Summary

In this chapter, we studied the joint total power optimization and adaptive sectorization problem for both the received and transmit power optimization cases. It is shown that the total received power is minimized by equalizing the number of users in each sector (EP) while the transmit power optimizing partition (OP) can be found with polynomial complexity using a graph theoretical approach. OP is expected to reduce total transmit power the most in scenarios similar to the one given in Figure 5.5 where terminal locations are non uniform in the cell. We also studied sectorization of a hot spot cell in the presence of out-of-cell interference. We used the same graph theoretic tools as in the previous case with slight modifications to the formulation under the assumption that the interference profile is available at the base station of the cell in question. Results obtained in both the single cell and the multicell scenarios show transmit power savings and strongly encourage the use of adaptive sectorization to support more users for a given SIR target.

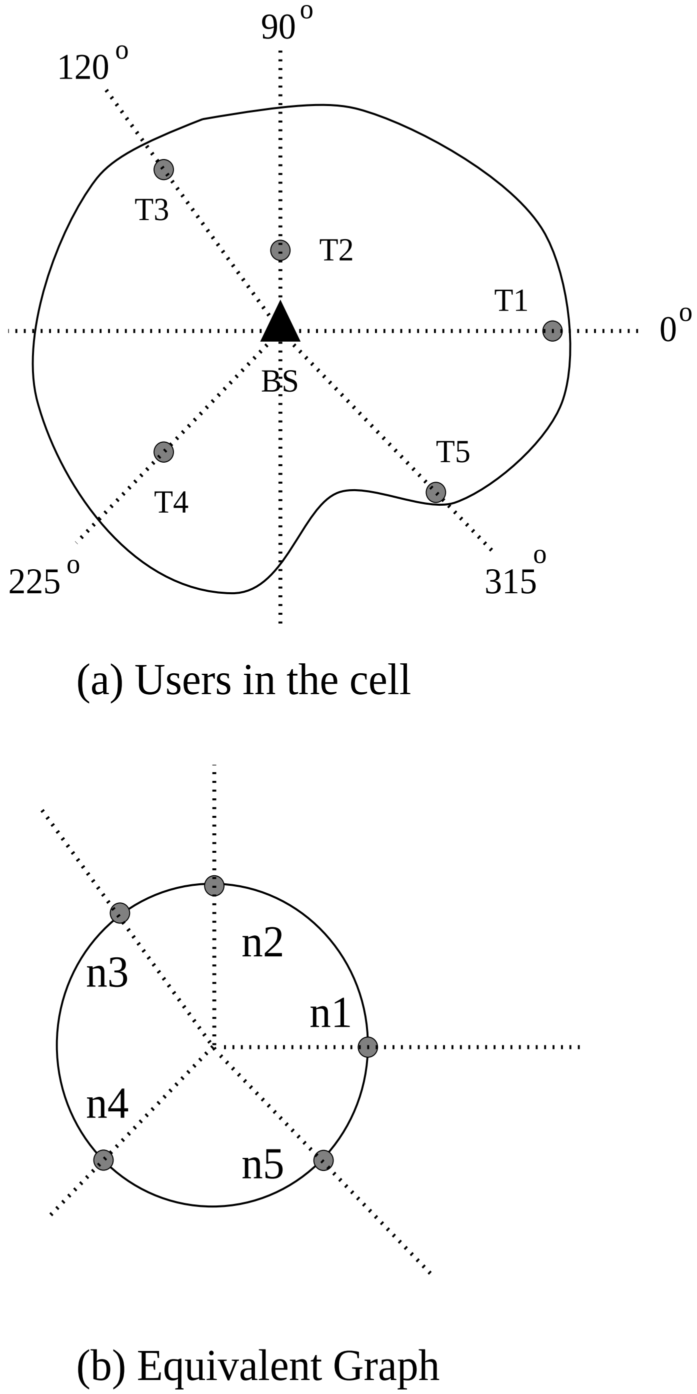

Figure 5.2: (a) Terminal locations within a cell. (b) The graph that is constructed based on the terminal locations. Node n_i in graph corresponds to terminal labeled Ti for all i.

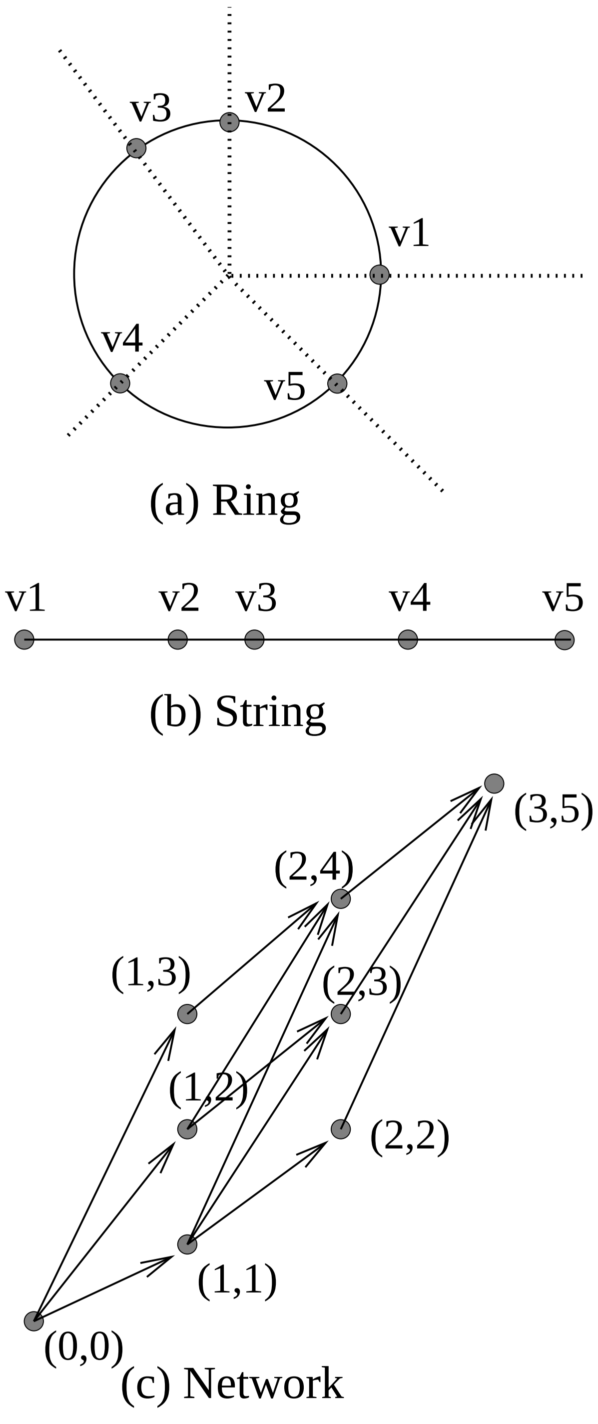

Figure 5.3: (a) Ring that is constructed based on terminal locations within cell (b) String obtained by removing e_5 (c) Network constructed from string that is to be partitioned into 3 subsets.

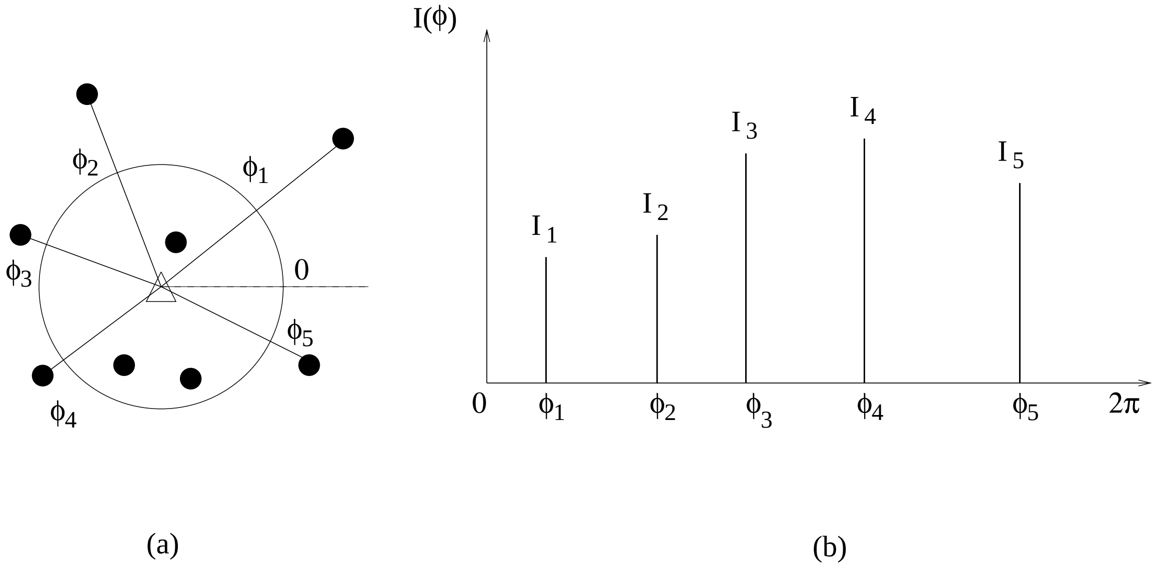

Figure 5.4: (a) Terminals outside the coverage area of the cell cause interference at the base station at the given angular positions. (b) The level of interference caused by each user maps to an interference profile where interference power at the base station is plotted as a function of the direction of interference.

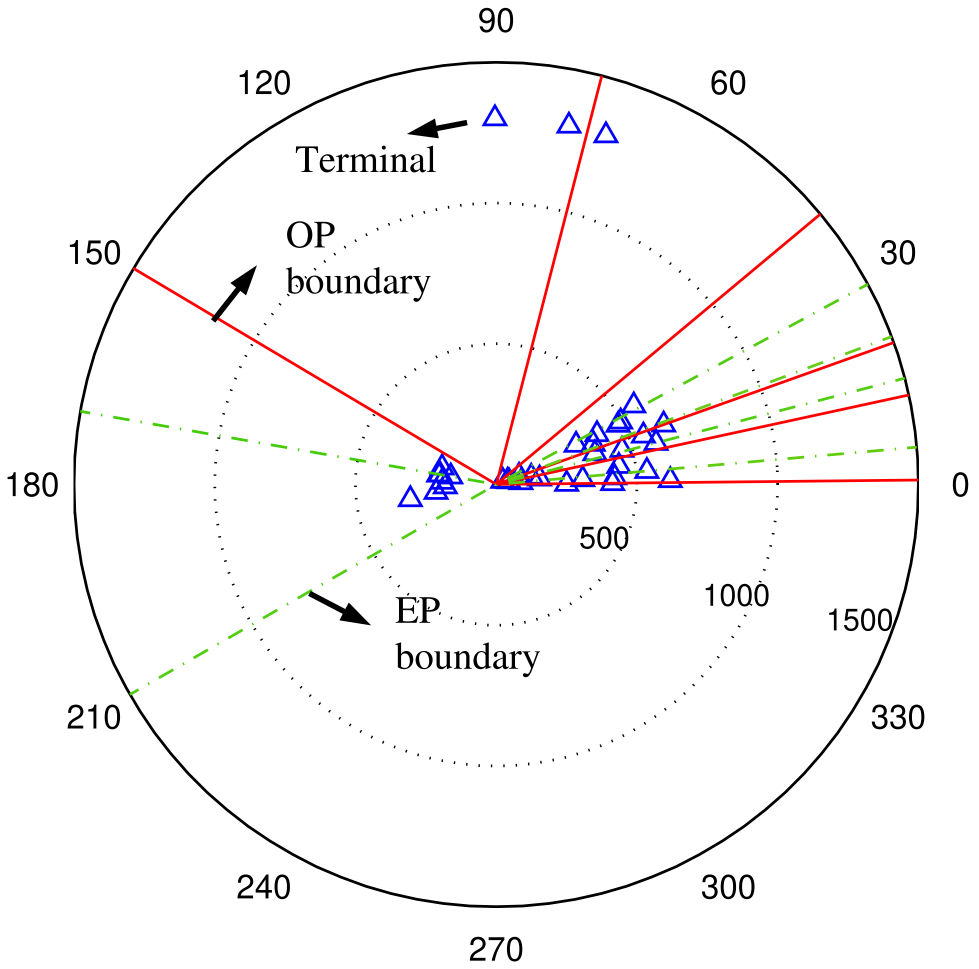

Figure 5.5: Sector boundaries of OP and EP in a cell where the number of terminals, $N = 36$, number of sectors, $M = 6$, processing gain, $G = 64$, target SIR, $\gamma^* = 7dB$ and noise power, $\sigma^2 = 10^{-13}$.

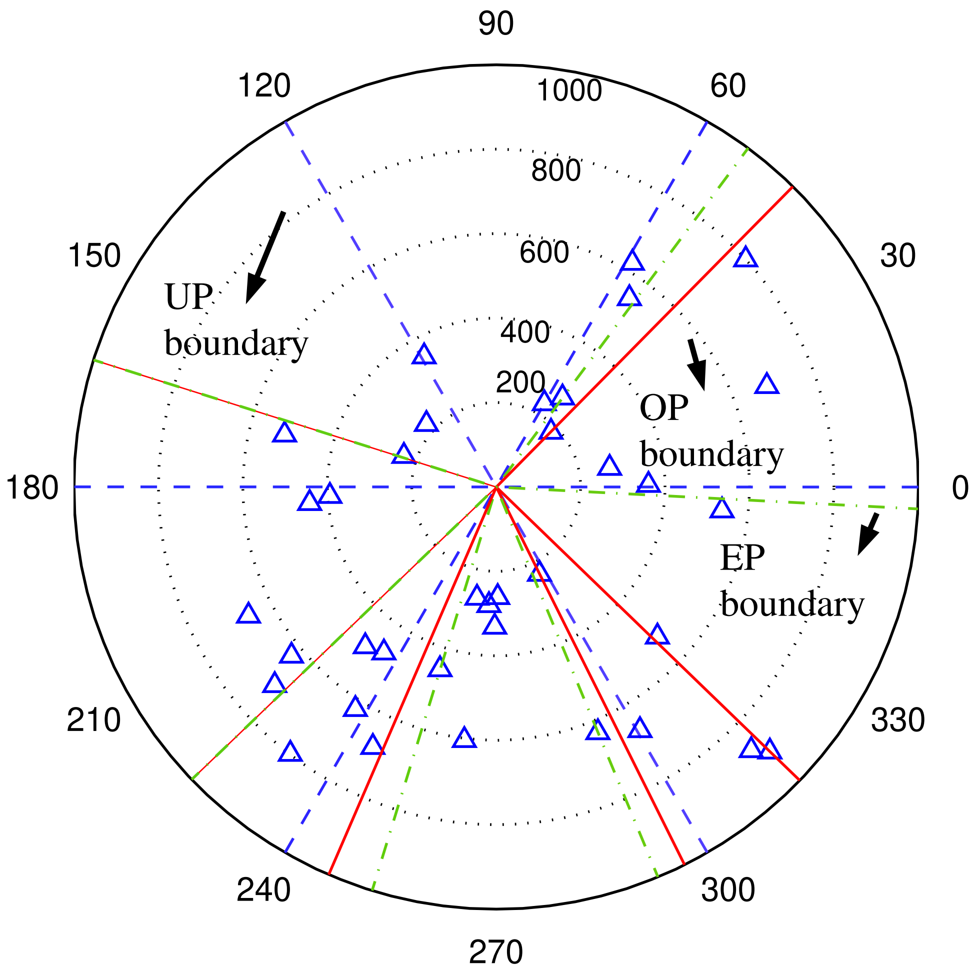

Figure 5.6: Sector boundaries of OP, EP and UP in a cell where the number of terminals, $N = 36$, number of sectors, $M = 6$, processing gain, $G = 64$, target SIR, $\gamma^* = 7dB$ and noise power, $\sigma^2 = 10^{-13}$. The distribution of terminals is uniform in area.

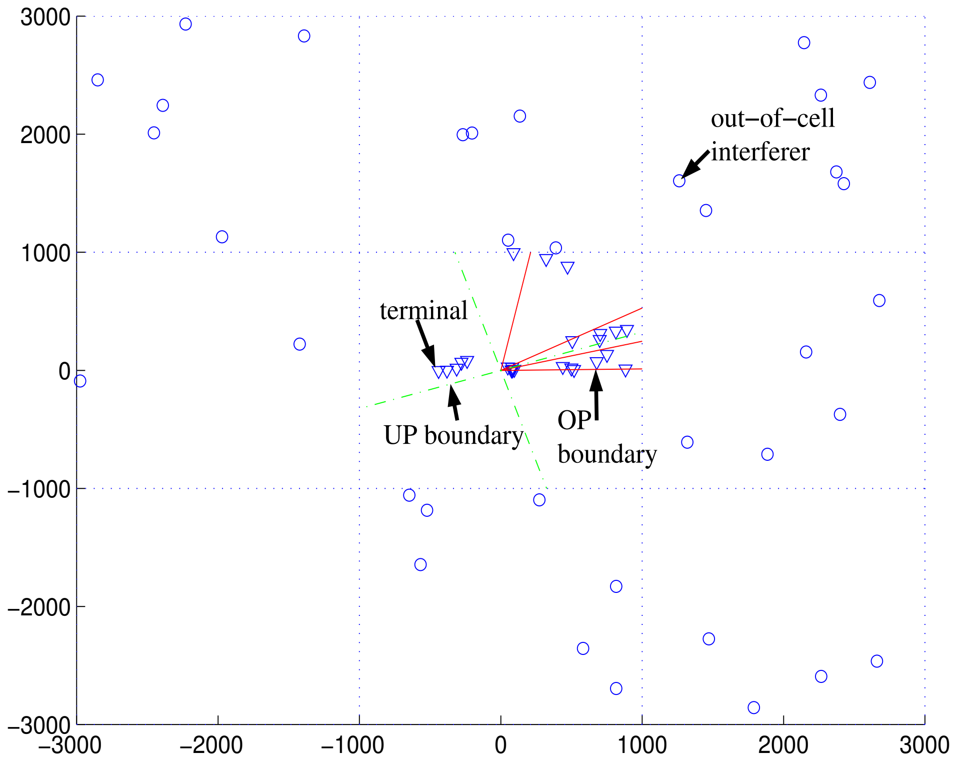

Figure 5.7: Sector boundaries of OP and UP in a cell with no OCI. Number of terminals, $N = 24$, number of sectors, $M = 4$, processing gain, $G = 64$, target SIR, $\gamma^* = 7dB$ and noise power, $\sigma^2 = 10^{-13}$.

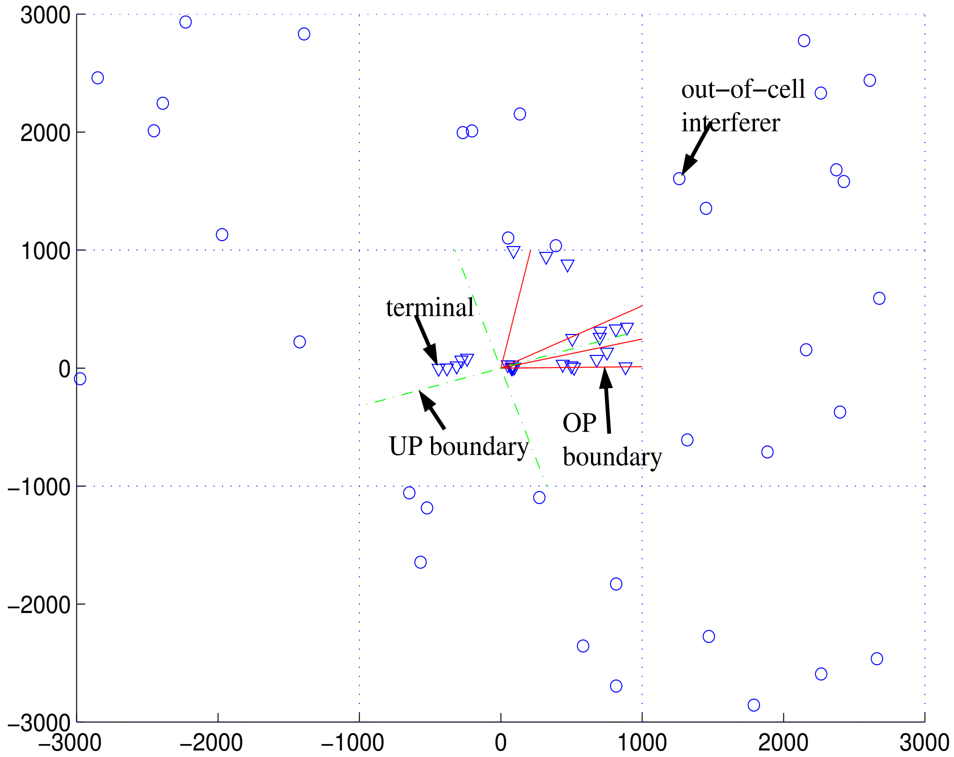

Figure 5.8: Sector boundaries of OP and UP in a cell with OCI. Number of terminals, $N = 24$, number of out-of-cell interferers, $L = 36$, number of sectors, $M = 4$, processing gain, $G = 64$, target SIR, $\gamma^* = 7dB$ and noise power, $\sigma^2 = 10^{-13}$. The out-of-cell terminals transmit with maximum power. The distribution of out-of-cell terminals is uniform. Same setting as in previous figure.

Chapter 6

Multiaccess Capacity of Packet-Switched CDMA

6.1 Introduction

This thesis so far has dealt with interference management issues for circuit–switched CDMA systems. In this last part of the thesis, we will consider a simple packet–switched CDMA system with a connectionless architecture and investigate its user capacity.

Packet based data communication services is anticipated for next generation CDMA systems [6]. Considerable effort has been directed towards the performance analysis and establishing efficient protocols for packet CDMA networks to date (see for example, [3] and references therein). These studies all dealt with issues after the timing parameters and the activity status of users are obtained. This is a valid assumption for systems where users have long and/or frequent communication sessions and thus can afford to first establish connection with the system and then keep this dedicated connection, essentially leading to a circuit switched architecture. However, for packet based services with relatively short sessions, the aforementioned model is not appropriate. Instead, a system with no dedicated resources should be considered.

In a packet switched CDMA system with *connectionless* architecture, no connection is to be kept between a user and the base station except when the user is sending information. User recognition and synchronization has to be established every time a user needs to transmit information. Another important feature of a connectionless packet CDMA system is that users do not have assigned signature sequences and have to send information using one of the predetermined signature sequences. Thus, the potential user population may be very large compared to the processing gain of the system. It is notable that existing analyses on such systems often ignore the fact that user acquisition (user recognition and timing acquisition) has to be achieved for every transmission or state that it can be achieved readily [4,5,17,30,57]. Thus, the capacity of such a system is observed to be interference limited.

In contrast to the observation that synchronization is readily achievable for connectionless packet CDMA systems, for circuit switched CDMA, there is a body of research that concentrates on user parameter acquisition and it has been well known that timing acquisition can be capacity limiting [37]. Timing acquisition work initially concentrated on determining the delay of a single user assuming a Gaussian disturbance, then on determining the delay of a single user in the presence of multiaccess interference [8,37,52]. Finally extensive effort has been directed towards estimation of all users' delays (and amplitudes) jointly [9, 38, 40, 63]. These studies assume that each user whose delay is estimated is known to be active. When this is not the case, as we have in connectionless packet switching CDMA, the number of active users may need to be estimated as well [82, 83]. Also, a dedicated signature sequence per user is assumed suggesting an inherently circuit switched structure. Although it is possible to modify the multiuser timing acquisition algorithms such as given in [9,47,63] to address the case where users do not have dedicated signatures, the overhead requirements on the order of hundreds of bits that these algorithms need to work well is likely to be intolerable for a short packet data system. Finally, receivers which can detect one new user's activity at a time along with estimation of its parameters have been constructed in [25]. This work is founded on the assumption that at most one new user attempts to access the system during the detection period which is not appropriate for the multiuser connnectionless architecture.

While the existing packet switching literature evaluates system capacity and finds it to be interference limited after access has been achieved by all active users, it appears that in a *connectionless* scheme where users do not keep active connections, the multi user access and acquisition procedures may need careful design to have acceptable system capacity. The motivation of this chapter therefore is to investigate the number of users that can be served by a simple connectionless system and whether this system is limited by the access procedure.

6.2 System Model

We consider a CDMA system where multiple users can attempt to access the system at the same time. As explained in the previous section, we assume a connectionless architecture. Further, we assume the simplest possible system, one where every user attempts to access the system using a common signature sequence (access channel). Note that if all users were synchronized, this system would not have the multiple access capability. Similar to Spread ALOHA [4], the asynchronous nature of the transmissions will enable this system to be capable of serving multiple users.

The communication format is to have access periods (frames) where the information transfer of the users that successfully access the system takes place. The users enter the system, transmit their messages and leave before the system announces the next access start time, thus no other connection is present or established during the service of these users. We assume that all users have acquired the base station's pilot signal and are tuned to the downlink paging channel where they can receive broadcast messages. The start time of the access message is broadcasted from the base station along with other necessary access parameters. The delay uncertainty of the users accessing the system thus comes from their transmission (propagation) delays relative to the broadcast of the base station and possible delays that the new users may add to their transmissions to further randomize the time of signal arrival to the base station . We assume these delays to be less than 1 bit period for each new user (see Figure 6.1).

We assume the packet that each user sends includes a preamble (a sequence of 1s) that we use to detect the user's activity and estimate its propagation delay followed by the information bits. Like the signature sequence, the preamble is the same for all users. If the user's presence is *detected* by the system during the access phase and its information is decoded correctly, the user receives an acknowledgment that its transmission was successful. Otherwise, the user will have to try to retransmit this packet at a later access period. Although we will concentrate on single packet transmission per user per access period, it is straight forward to extend this model to one where a data frame consisting of a few data packets are to be transmitted (a short message) by each active user following the initial packet. In this case, the initial packet plays the role of the identifying packet and users whose first packets are successful go forward with the rest of the packets (see Figure 6.2). The multiaccess capacity of this system is identical to the single packet model.

The model considered suggests a *two-stage* receiver whose initial stage works on the transmitted preambles to detect the activity status of the users and is followed by a detector which will decode the active users' information (see Figure 6.3). In this chapter, we will concentrate on how to construct the first stage, i.e. the detection of the presence of the active users. This will be done by detecting the activities (if any) at or around certain delay values determined by the users' propagation delays. The detector for the second stage is an asynchronous multiuser detector and is constructed using the information about the activity status of the users supplied by the first stage, i.e. if users' activity is detected at or around a delay value, a corresponding receiver filter is designed to decode the incoming identification information and/or suppress interference.

The performance of the first stage is of vital importance to the system since the performance of the second stage detector hinges upon the correctness of the information supplied by the first stage. A false alarm event, the event that the system declares one (or more) user(s) present at a particular delay when there is none, implies waste of resources for the second stage since it may require the detector to try to decode fictitious users and to suppress their actually non-existing interference to other users. A miss event which contributes to the rate at which the system fails to capture a user is also highly undesirable since an active user will not enter the system and its interference will not be cancelled while other users are being decoded during the second stage.

6.3 Multiuser Access Detection

The base station has to detect the activity of *a random number* of users along with the delays of each of these users to be able to decode their packets. This has to be done using the preamble (sequence of 1s) at the beginning of the packet each active user transmits. We term this problem of multiple user acquisition the "multiuser access detection" problem and the corresponding first stage of the detector, the Multiuser Access Detector (MUAD). Since the first access stage uses a preamble of all 1's, the received signal during the first stage of access is

$$r(t) = \sum_{i=1}^{N_A} \sqrt{q_i} s_a(t - \tau_i) + n(t) \qquad t \in [0, LT_b] \tag{6.1}$$

where N_A is the number of active users, q_i and τ_i are the received power and the delay of the user i, and $n(t)$ is the zero mean white Gaussian noise with power spectral density σ^2, $L > 1$ is the length of the preamble in bits, and T_b is the bit duration. We assume $0 \leq \tau_i \leq T_{max} < T_b$. The accessing signature sequence $s_a(t)$ can be expressed as

$$s_a(t) = \sum_{i=0}^{G-1} c(i) \frac{1}{\sqrt{G}} p(t - iT_c) \tag{6.2}$$

where G is the processing gain, T_c is the chip duration, $c(i) \in \{-1, 1\}$ is the i^{th} chip value, and $p(t)$ is the chip waveform normalized to have unit energy. Throughout this chapter, we will assume for simplicity that $p(t)$ is rectangular. The received signal is observed from the start of the access message with 1-bit delay, thus for a total of $L - 1$ bits. Since $0 \leq \tau_i \leq T_{max} < T_b$, observing the signal with 1-bit delay ensures the capture of at least one bit period where all new users are actively sending their access preamble. Note that during each observed bit interval, the contribution of each active terminal consists of the access signature sequence circularly shifted by that terminal's delay value (see Figure 6.4).

For further processing, the received signal is discretized by projecting onto a sequence of chip waveforms delayed by multiples of T_c and then sampling at the chip rate (T_c samples per second). The idea then is to separate users (by processing the discrete signal) whose delays are sufficiently apart from each other. At this point, one can envision sampling the filter output every δ second (a fraction of T_c) in the hopes of resolving users that are closer in their delay values. For rectangular chip waveforms, sampling at a higher rate effectively would increase the dimension of the signal space and would improve resolvability. However, it is observed in [39] that this higher dimensional space is an artifact resulting from the use of rectangular pulses. In practice, when chip pulses are bandlimited, there should be no such improvement in resolvability. Since our reasoning for using rectangular pulses is strictly for analytical ease, we will purposefully refrain from this faster processing.

Let us move on to the discrete signal representation. It is easy to see that with chip waveforms time limited to duration T_c, one can express a chip matched filtered signal delayed by any delay value as a combination of two *adjacent* vectors as explained below.

Let us define $s_j^\alpha(t)$ as the circularly shifted version of the basic access signature sequence by $(j+\alpha)\ T_c$ where $\alpha \in [0,1]$ is the delay mismatch expressed as a fraction of one chip. The chip matched filtered version of $s_j^\alpha(t)$, $\mathbf{s}_j^\alpha$ can be written as

$$\mathbf{s}_j^\alpha = R(\alpha)\ \mathbf{s}_j + R(1-\alpha)\ \mathbf{s}_{j+1} = (1-\alpha)\ \mathbf{s}_j + \alpha\ \mathbf{s}_{j+1} \tag{6.3}$$

where $R(\alpha) = 1 - |\alpha|$ is the autocorrelation function of the rectangular chip waveform and $\mathbf{s}_j$ is the access signature sequence circularly shifted by j chips, i.e. $\mathbf{s}_0 = [c(0), \cdots, c(G-1)]$ and $\mathbf{s}_j = [c(G-j), \cdots, c(G-1), c(0), c(1), \cdots, (G-1-j)]$ for $1 \leq j < G$. Associated with each active user k, there is a fractional mismatch α_k $(0 \leq \alpha_k < 1)$ such that

$$\tau_k = (i_k + \alpha_k)T_c \qquad i_k = \lfloor \tau_k / T_c \rfloor \tag{6.4}$$

Recall that the received signal is observed starting at time $t = T_b$ and the contribution of user k within each bit of the observed signal is the access signature sequence circularly shifted by τ_k. Let us assume $\lceil T_{max}/T_c \rceil = M - 1$ where $M \leq G$ and define B_i as the set of users whose delays (τ_k) are in the interval $[(i-1)T_c, iT_c]$ $(i = 1, \cdots, M-1)$. The received signal in the observation window $t \in [T_b, L\ T_b]$ then can be expressed as

$$r(t) = \sum_{i=1}^{M-1} \sum_{k \in B_i} \sqrt{q_k} s_{i-1}^{\alpha_k}(t) + n(t) \tag{6.5}$$

Now, using (6.3), we can express the discretized received signal in the j^{th} observation interval $([j\ T_b, (j+1)\ T_b])$ as

$$\mathbf{r}_j = \mathbf{S}\mathbf{v} + \mathbf{n}_j \tag{6.6}$$

where $\mathbf{n}_j$ is a white Gaussian vector, $\mathbf{S}$ is the matrix with columns $\mathbf{s}_0, ..., \mathbf{s}_{M-1}$ and

$$v_i = \begin{cases} \sum\limits_{k\in B_i} (1-\alpha_k)\sqrt{q_k}, & i=1 \\ \sum\limits_{k\in B_{i-1}} \alpha_k\sqrt{q_k} + \sum\limits_{k\in B_i} (1-\alpha_k)\sqrt{q_k}, & 2\le i<M \\ \sum\limits_{k\in B_{i-1}} \alpha_k\sqrt{q_k}, & i=M \end{cases} \tag{6.7}$$

As can be seen from (6.7), users that belong to $B_{(i-1)}$ and B_i contribute to v_i with associated weights that depend on their delay mismatch parameters. Note that it is easy to extend this model to $M = G+1$ ($\lceil T_{max}/T_c \rceil = G$) by taking into account the fact that in addition to what is given in (6.7), v_1 and v_M will be one and the same and will have contributions from users in B_{M-1} and B_1. Fixing the maximum number of users to N at each interval, we can express $\mathbf{v}$ as

$$\mathbf{v} = \mathbf{ADb} \tag{6.8}$$

where

$$\mathbf{A} = \begin{bmatrix} 1-\alpha_1 & 1-\alpha_2 & \cdots & 0 & \cdots & 0 \\ \alpha_1 & \alpha_2 & \cdots & 1-\alpha_{N+1} & \cdots & 0 \\ 0 & 0 & \cdots & \alpha_{N+1} & \cdots & 0 \\ \vdots & \vdots & & \vdots & & \vdots \\ 0 & 0 & \cdots & 0 & \cdots & 1-\alpha_{NM} \\ 0 & 0 & \cdots & 0 & \cdots & \alpha_{NM} \end{bmatrix} \tag{6.9}$$

is the $M \times NM$ matrix that contains the fractional mismatch values of all possible users (active or not), $\mathbf{D}$ is the $NM \times NM$ diagonal matrix with the amplitudes as the diagonal entries and $\mathbf{b}$ is the binary $(0-1)$ NM-vector whose first N components represent the activity status of the possible users that belong to B_1, the next N components represent the users in B_2 and so on. In particular, $b_{(i-1)N+j} = 1$ means that the j^{th} user in B_i is active. Collecting multiple observations and defining $\mathbf{r} = \frac{1}{L-1}\sum_{j=1}^{L} \mathbf{r}_j$, the joint maximum likelihood detector for recognition of active users and estimation of their amplitudes and delays is given by

$$[\hat{\mathbf{A}}, \hat{\mathbf{D}}, \hat{\mathbf{b}}] = \arg\min_{\mathbf{A},\mathbf{D},\mathbf{b}} \mathbf{b}^\top\mathbf{D}^\top\mathbf{A}^\top\mathbf{S}^\top\mathbf{SADb} - 2\mathbf{r}^\top\mathbf{SADb} \tag{6.10}$$

It is not hard to see that if $\mathbf{A}$ and $\mathbf{D}$ were exactly known, to estimate $\mathbf{b}$ would require to solve a $0-1$ quadratic program which is N-P hard. When estimation of $\mathbf{A}$ is also required, even with perfect knowledge of the amplitude levels, the mismatch values do not have a unique solution if there is more than one active user in a given interval. Let us concentrate on the case of equal (known) received power, i.e. $\mathbf{D} = \sqrt{q}\mathbf{I}$ (A similar argument can be made for the unequal power case). Due to the discrete nature of $\mathbf{b}$, the solution to (6.10) has to be found by estimating $\mathbf{A}$ for every distinct possibility of users' activities represented by the $\mathbf{b}$ vectors. First, we observe that order of users in an interval is not important. For instance, if b_i and b_{i+1} are exchanged the same cost value is attained by exchanging the corresponding columns of $\mathbf{A}$, i.e. $\mathbf{a}_i$ and $\mathbf{a}_{i+1}$. Secondly, given $\mathbf{b}$, the minimizer of (6.10) can be found in the variable

$$\tilde{\mathbf{v}} \ = \ \mathbf{A}\mathbf{b} = \sum_{i=1}^{NM} b_i \mathbf{a}_i = \sum_{i, b_i = 1} \mathbf{a}_i$$

and not in individual $\mathbf{a}_i$'s. Now recall the structure of $\mathbf{A}$ in (6.9). Since only the sum of the mismatch values of the users show up in $\mathbf{v}$, if there is more than one column that corresponds to the same interval, it is not possible to uniquely identify the individual mismatch values. The conclusion here is that by using the discrete delay filters we can have a maximum likelihood estimate of how many active users arrive at each interval and the sum of the mismatch values of these users. This has to be done by solving $(N+1)^M$ optimization problems. Furthermore, if the number of users arriving at each interval is more than one, then we have effectively a collision since their individual delays can not be identified.

Since the above optimization is prohibitively complex for large M, one should consider suboptimum methods. As a simple solution, we consider a decorrelating multiuser access detector that will provide an unbiased estimate for v. The idea is to eliminate users that are outside $[(i-2)T_c, iT_c]$ at each detector output y_i. The output of the detector is

$$\mathbf{y} = \frac{1}{(L-1)} (\mathbf{S}^\top \mathbf{S})^{-1} \mathbf{S}^\top \mathbf{r} = \mathbf{v} + \tilde{\mathbf{n}} \tag{6.11}$$

where $\tilde{\mathbf{n}}$ is $\mathbf{0}$-mean Gaussian noise vector with covariance $\frac{\sigma^2}{(L-1)}\mathbf{\Gamma}^{-1}$ where we denote $\mathbf{\Gamma} = \mathbf{S}^\top\mathbf{S}$. Since in general, v_i contains contributions from an unknown number of users each of which has an unknown delay offset, it is difficult to solve for the exact number of active users. Instead, we can simply try to detect if there are *any* active users around the i^{th} delay value by designing a simple threshold test. If we ignore the correlations between filter outputs, we can design a Neyman-Pearson type binary hypothesis test for y_i with the hypotheses:

- H_{i1} : Activity detected at the i^{th} filter
- H_{i0} : No activity at the i^{th} filter

The threshold is set for each test so that a fixed false alarm probability performance (α_F) is achieved. Specifically, for y_i we set the threshold Λ_i such that

$$\alpha_F = \int_{\Lambda_i}^{\infty} \frac{1}{\sqrt{2\pi\Gamma_{ii}^{-1}\sigma^2/(L-1)}} e^{-x^2/(2\Gamma_{ii}^{-1}\sigma^2/(L-1))} dx \tag{6.12}$$

The resulting threshold Λ_i and the test are

$$\Lambda_i = Q^{-1}(\alpha_F)\sqrt{\frac{\Gamma_{ii}^{-1}\sigma^2}{(L-1)}} \quad , \quad\quad y_i \underset{H_{0i}}{\overset{H_{1i}}{\gtrless}} \Lambda_i \tag{6.13}$$

where $Q(x)$ is the standard normal complementary CDF.

So, for activity to be detected at the output of the i^{th} decorrelator, the output should exceed a factor that depends on the specified false alarm rate times the *enhanced* noise magnitude at that decorrelator. If activity is detected at the output of the i^{th} decorrelator, the second stage detector should use this information to try and decode the identification information of a user around the corresponding delay value. The second stage detector, as explained before is a multiuser bit detector.

6.4 System Capacity

The capacity is defined as the average number of users that can successfully access the system during an access attempt. This depends on the performance of the MUAD and the performance of the second stage bit detector which clearly is coupled with that of the MUAD under given system parameters which include the user arrival rate, MUAD false alarm rate and the quality of service requirement for reliable connection to the system. Since the exact analytical expression for the success probability of a user is complex, we have simulated a DS/CDMA system with processing gain $G = 31$ to observe the average number of users that the system can accommodate. The accessing signature sequence is chosen to be a basic M-sequence. The system has Poisson arrivals with mean λ and users which cannot access the system are immediately cleared. We assume equal received power for all users with values $10, 13, 15$ and 20dB and the background noise power σ^2 to be unity (0dB). In all experiments, the thresholds for the decorrelator outputs of the first stage are set such that the false alarm rate of each test is $\alpha_F = 0.01$. The second stage receivers constructed are one-shot asynchronous decorrelators. For all experiments, the reliable connection requirement is that the signal-to-interference ratio (SIR) of the user has to be above 5dB.

The system we compare against is a slotted ALOHA system with M parallel orthogonal (narrowband) channels where $T_{max}/T_c = M - 1$. Assuming the channels are in frequency domain, we have an equivalent system that occupy the same bandwidth that of the CDMA system if we set $M = G$. It is easy to see that with the given received power values and SIR target, the ALOHA channels are only limited by collisions. We again assume unsuccessful access attempts are immediately cleared. Recall that we assume uniform delay values for the CDMA users and thus the probability of a user being in any one chip interval is $1/M$. In the equivalent ALOHA model, we assume the probability of a user transmitting in a particular channel is $1/M$. This enables us to decompose the aggregate arrival rate into M independent λ/M Poisson streams each of which is the traffic load of one ALOHA channel. Thus, the sum capacity is M times that of one ALOHA channel. The total throughput (average number of users that can access the system) is $\lambda e^{-\lambda/M}$. The maximum is M/e.

Our experiments showed that in the absence of a fine tracking algorithm before the second stage, the system capacity is very low. This is due to the sensitivity of the bit detectors used in the second stage to the difference between the exact delays of the active users and the coarse delay estimates supplied by the first stage. Furthermore, even in the presence of tracking, where we assumed users can be tracked perfectly when alone in a chip interval, we observed the capacity is lower than sum capacity of M parallel slotted ALOHA channels (Figures (6.5) and (6.6)). This effect is even more pronounced when the delay space covers the whole bit interval where the equivalent

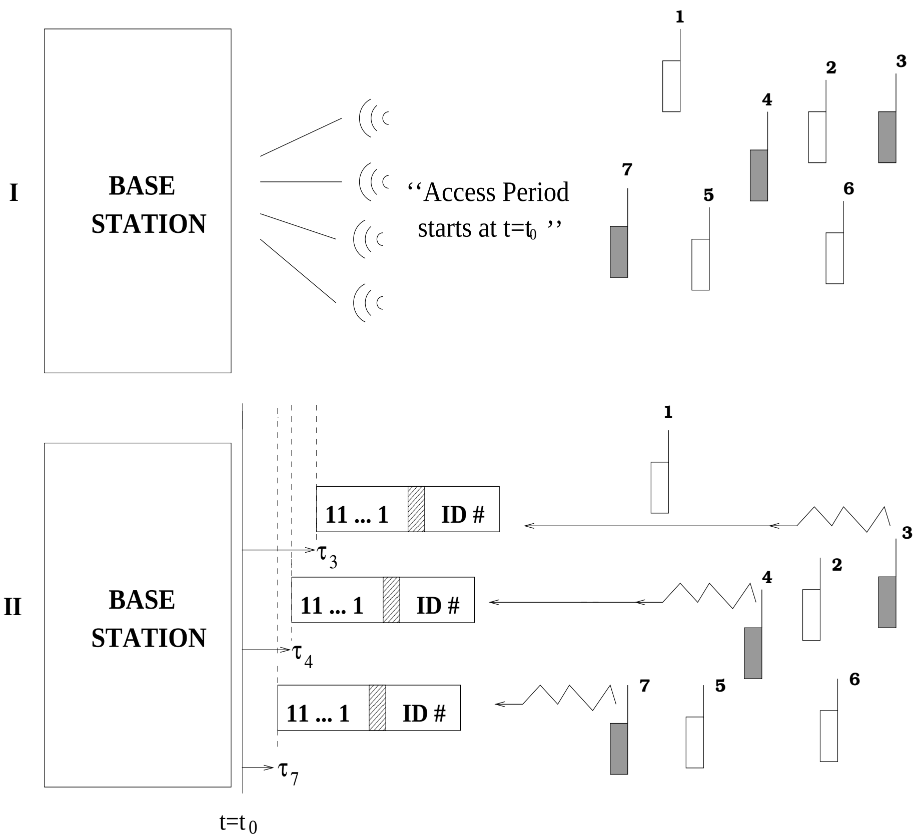

Figure 6.1: Multiaccess Model for the packet switched CDMA system

ALOHA system with 31 orthogonal channels can support about one more user than the packet CDMA system even with very high received power values (Figure (6.6)).

6.5 Chapter Summary

In this chapter, we considered a simple connectionless packet-switched CDMA system and investigated the number of users this system can support taking into account the fact that the base station has to acquire the activity and delays of each of the users during each transmission. For the simple system we considered, our observations lead us to believe that the access process is capacity limiting for a packet switching CDMA system. This is in contrast to earlier literature where capacity of random access CDMA is characterized essentially ignoring the effects of user acquisition. To increase the capacity of this system, effective retry strategies have to be designed.

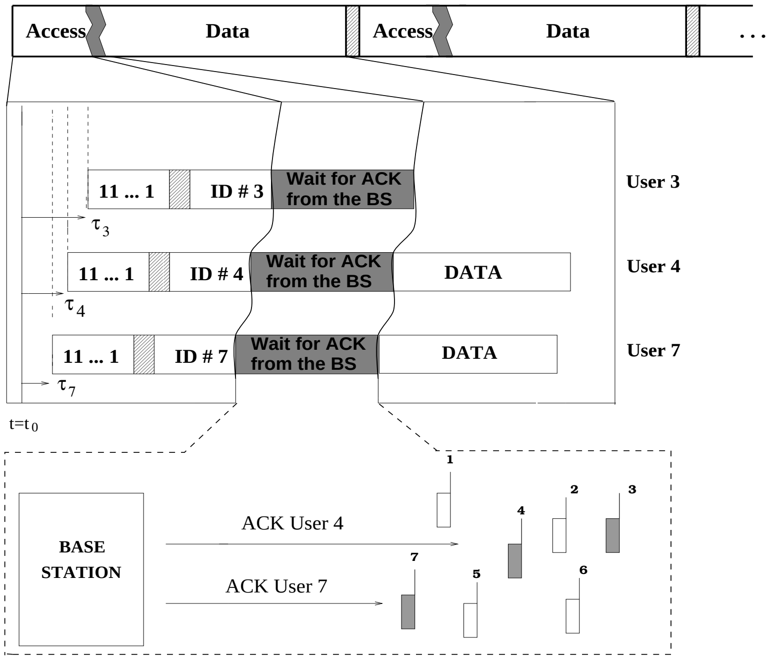

Figure 6.2: Communication Model for the packet switched CDMA system

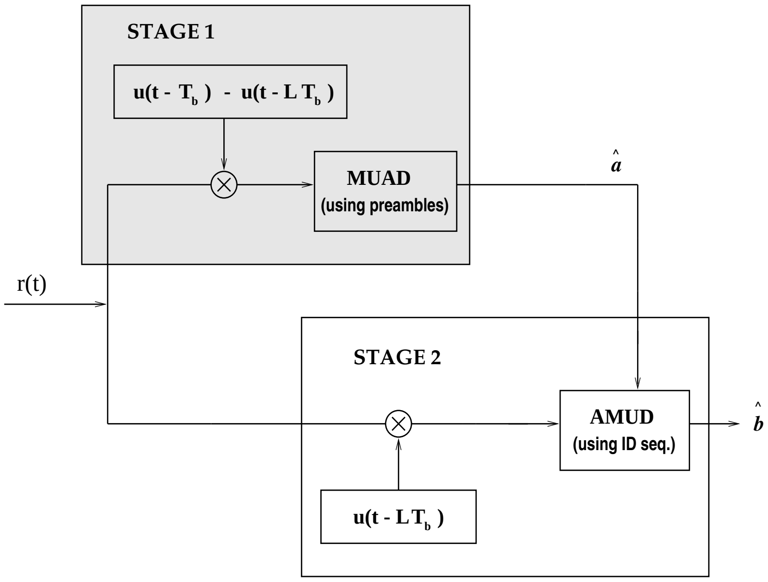

Figure 6.3: Two-stage detector for the packet switched CDMA system. MUAD: Multiuser Access Detector, AMUD: Asynchronous Multiuser Detector, $u(.)$ is the step function $(u(t) = 1, \quad t > 0)$

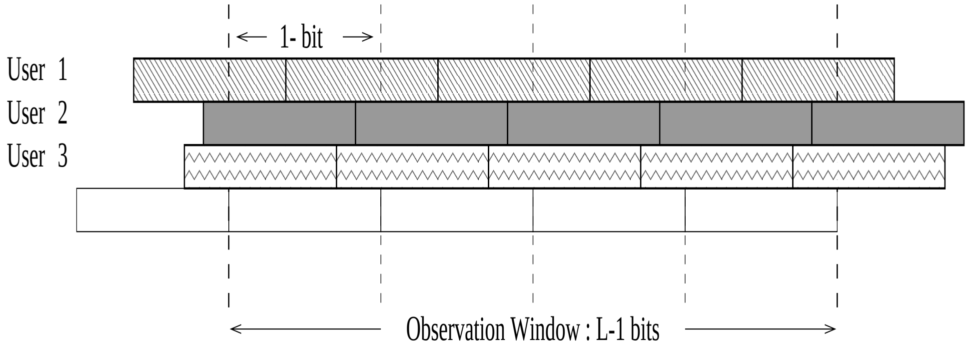

Figure 6.4: Observation of the access preamble by the multiuser access receiver

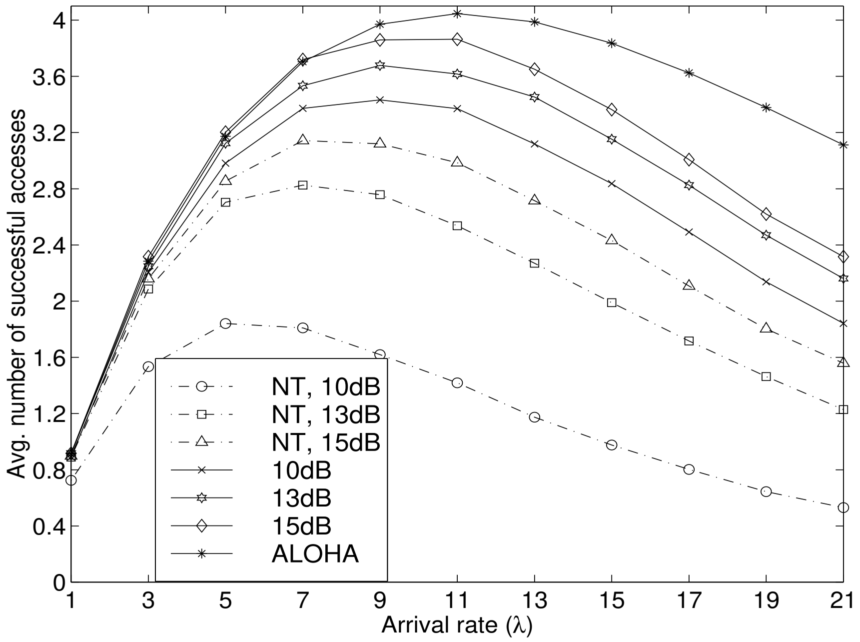

Figure 6.5: Comparison of access capacity for received power values $10, 13, 15$dB and M orthogonal ALOHA channels, NT:No tracking, M=11, L=11 (Longer preambles exhibited similar performance)

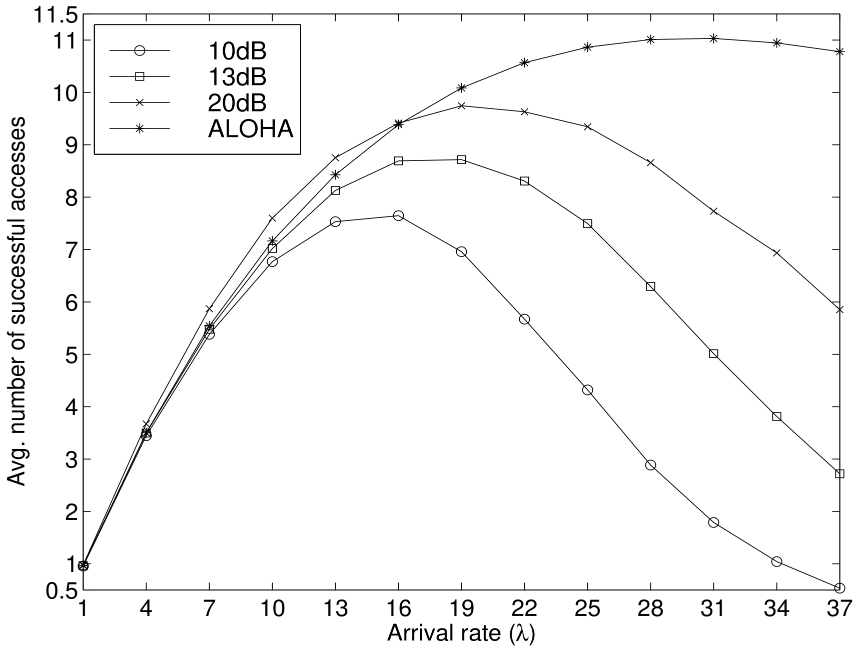

Figure 6.6: Comparison of access capacity for received power values $10, 13, 20$dB and M orthogonal ALOHA channels, M=G=31, L=11

Chapter 7

Conclusion and Future Work

7.1 Thesis Summary

In this thesis, we have studied capacity enhancement techniques of circuit-switched CDMA systems via lessening the near-far effect using various methods of interference management. The interference management techniques considered in this thesis are: multiuser detection, receiver beamforming, transmit power control and adaptive cell sectorization.

We have investigated approximate solutions for the optimum multiuser detection problem using nonlinear programming methods and showed that various suboptimum multiuser detectors proposed in the literature can be seen as approximations to the optimum multiuser detector. We also proposed a simple iterative detector whose behavior is similar to that of the MMSE detector. Using well known nonlinear programming techniques, the convergence behavior of the iterative detectors considered is easily established.

We have derived combined multiuser detection (temporal filtering) and receiver beamforming (spatial filtering) structures to further suppress the interference each user experiences. The filter structures considered are two dimensional linear filter structures which are all based on minimum mean squared error criteria, but differ in how the MMSE problems are defined in the temporal and spatial domains, i.e., jointly or in cascade. It is shown that while the joint optimum MMSE filter achieves the maximum signal-to-interference ratio (SIR) among all possible matrix filters, the constrained optimum MMSE filter which results in a single temporal and single spatial filter, outperforms all combined single-user/multiple-user approaches and cascaded optimization approaches either uniformly or asymptotically. The constrained optimum MMSE filter is near-far resistant in all but very highly loaded systems and enjoys low complexity.

This thesis also investigates the capacity increase that is possible by combining power control with intelligent temporal and spatial receiver filter design. The problem of minimizing the total transmit power of all users while guaranteeing to give each user at least its required QoS level (SIR) is solved when joint temporal-spatial MMSE filters (in unconstrained and constrained filter domains) are employed at the base-station. Two-step iterative power control algorithms that converge to the optimum powers and the joint temporal and spatial receiver filters in the corresponding filter domains are given. A power control algorithm with a less complex filter update procedure is also given. We observe that significant savings in total transmit power are possible if filtering in both domains is utilized compared with conventional power control and joint optimal power control and filtering in only one domain.

When the base station uses only single user receivers (matched filters), the best way of partitioning the cell into M non-overlapping sectors in response to user density is studied. The solution of the minimum total transmit power problem is found using graph theoretical methods, in particular by finding the optimum partition on the ring that corresponds to the system. The resulting arrangements is observed to bring power savings over uniformly partitioning the cell as well as sectorizing the cell such that an equal number of users reside in each sector.

This thesis also investigates the multiuser access capacity of a connectionless packet-switched CDMA system. No dedicated connection is kept between users and the system. Users have to access the system using the same signature sequence and have to be acquired each time they need to send packets. Due to the asynchronous nature of the system, multiple users can be accommodated using appropriate receiver processing even when a single signature is available. However, with short data packets and limited receiver processing complexity, we observe that the average number of users that can be served by this system is quite low. The limits on the capacity is imposed by the fact that multiple users have to recognized and acquired for each packet they need to send. Capacity enhancement is needed for this system.

Most of the research results in this thesis have been presented previously in various conferences, see [59, 88–92].

7.2 Future Directions

The nonlinear programming approach in chapter 2 can be used to derive other types of detectors as well. For instance, it is easy to show that the hypercube type relaxation is the optimum detector when the uplink gains of the users are assumed completely random in $[0, 1]$ and unknown [92]. For real channels, the uplink gains are likely to have much smaller values than 1. Using this fact, detectors for channels with uncertainties can be designed with smaller box boundaries.

The iterative algorithms presented in chapter 4 that update the transmit powers and the temporal-spatial filters of the users assume a static environment. The convergence is proved assuming that the system parameters such as the activity status of the users and the channel conditions will not change until the particular algorithm converges. The behavior of such an algorithm under dynamically changing system need to be investigated. Also, the algorithms are deterministic in the sense that perfect measurements are assumed. The investigation of convergence conditions when stochastic observations are used is an open problem.

In chapter 5, the optimum cell sectorization algorithm found has computational complexity $\mathcal{O}(N^3M)$ where N is the number of users and M is the number of sectors. More efficient graph theoretic methods could be developed to bring the complexity further down.

Finally, in chapter 6, we have provided an introduction to the capacity characterization of a simple packet-switched CDMA system. There may be a number of ways to increase the capacity of this CDMA system. Higher layer solution possibilities include developing a frame structure with more than one access period before data transmission, designing retry strategies that will separate the users further apart in delay space and/or distribute the traffic load between multiple access attempts. For example, since the mobile already has the delay estimate for the downlink (through base station pilot acquisition), it can use this information to change its delay during retransmissions.

References

[1] The cdma2000 ITU–R RTT candidate submission. Standards Proposal. 1998.

[2] TIA/EIA/IS-95-A mobile station–base station compatibility standard for dual-mode wideband spread spectrum cellular system. 1993.

[3] Code division multiple access networks - IV. *IEEE Journal on Selected Areas in Communications*, 14(9), December 1996.

[4] N. Abramson. Fundamentals of packet multiple access for satellite networks. *IEEE Journal on Selected Areas in Communications*, 10(2), February 1992.

[5] N. Abramson. Wideband random access for the last mile. *IEEE Personal Communication Magazine*, December 1996.

[6] F. Adachi, M. Sawahashi, and H. Suda. Wideband DS-CDMA for next-generation mobile communication systems. *IEEE Communications Magazine*, 36(9):56–69, September 1998.

[7] M. S. Bazaraa, J. J. Jarvis, and H. D. Sherali. *Linear Programming and Network Flows*. John Wiley and Sons Ltd, 2nd edition, 1990.

[8] S. E. Bensley and B. Aazhang. Maximum likelihood estimation of a single user's delay for code division multiple access communication systems. In *Conference on Information Science and Systems*, pages 998–1003, 1994.

[9] S. E. Bensley and B. Aazhang. Subspace-based channel estimation for code division multiple access communication systems. *IEEE Transactions on Communications*, 44(8):1009–1020, August 1996.

[10] X. Bernstein and A. M. Haimovich. Space-time optimum combining for CDMA communications. *Wireless Personal Communications*, (3):73–89, 1996.

[11] D. Bertsekas. *Nonlinear Programming*. Athena Scientific, 1995.

[12] D. Bertsekas and R. Gallager. *Data Networks*. Prentice-Hall, 1987.

[13] D. Bertsekas and J. Tsitsiklis. *Parallel and Distributed Computation*. Prentice-Hall, 1989.

[14] G. K. Chan. Effects of sectorization on the spectrum efficiency of cellular radio systems. *IEEE Transactions on Vehicular Technology*, 41(3):217–225, August 1992.

[15] E. Dahlman, P. Beming, F. Ovesjo J. Knutsson, M. Persson, and C. Roobol. WCDMA – the radio interface for future mobile multimedia communications. *IEEE Transactions on Vehicular Technology*, 47(4):1105–1118, November 1998.

[16] E. Dahlman, B. Gudmundson, M. Nilsson, and J. Skold. UMTS/IMT-2000 based on wideband CDMA. *IEEE Communications Magazine*, 36(9):70–80, September 1998.

[17] S. Dastangoo, S. Resheff, and B. R. Vojcic. Performance analysis of multi-code spread ALOHA system: Finite and infinite population. In *IEEE International Symposium on Personal, Indoor and Mobile Radio Communications*, 1998.

[18] 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.

[19] U. Fawer and B. Aazhang. A multiuser receiver for code division multiple access communications over multipath channels. *IEEE Transactions on Communications*, 43(2/3/4):1556–1565, February/March/April 1995.

[20] J. A. Fessler and A. O. Hero. Space-alternating generalized expectation-maximization algorithm. *IEEE Transactions on Signal Processing*, 42(10):2664–2677, October 1994.

[21] 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.

[22] M. R. Garey and D. S. Johnson. *Computers and Intractability.* W.H. Freeman and Company, 1979.

[23] V. K. Garg and E. L. Sneed. Digital wireless local loop system. *IEEE Communications Magazine*, 34(10):112–115, October 1996.

[24] 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.

[25] K. W. Halford and M. Brandt-Pearce. New-user identification in a CDMA system. *IEEE Transactions on Communications*, 46(1):144–155, January 1998.

[26] 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.

[27] S. V. Hanly and D. N. C. Tse. Resource pooling and effective bandwidths in CDMA systems with multiuser receivers and spatial diversity. *IEEE Transactions on Information Theory.* Submitted, May 1999.http://degas.eecs.berkeley.edu/∼dtse/pub.html.

[28] P. Hansen. Methods of nonlinear 0-1 programming. *Annals of Discrete Mathematics 5: Discrete Optimization II.* P.L. Hammer, E.L. Johnson and B.H. Korte, Eds. North Holland, 1979.

[29] M. Honig, U. Madhow, and S. Verdú. Blind adaptive multiuser detection. *IEEE Transactions on Information Theory*, 41(4):944–960, July 1995.

[30] F. Khan, C. Roobol, and J. Larsson. Performance of a common channel packet access in WCDMA. In *IEEE International Symposium on Personal, Indoor and Mobile Radio Communications*, 1998.

[31] R. Kohno, H. Imai, M. Hatori, and S. Pasupathy. Combination of an adaptive array antenna and a canceller of interference for direct-sequence spread-spectrum multiple-access system. *IEEE Journal on Selected Areas in Communications*, 8(4):675–682, May 1990.

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

[33] E.L. Lawler. *Cobinatorial Optimization: Networks and Matroids*. Holt, Rinehart and Winston, 1976.

[34] T. Lengauer. *Combinatorial Algorithms for Integrated Circuit Layout*. John Wiley and Sons Ltd., 1990.

[35] 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.

[36] 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.

[37] U. Madhow and M. B. Pursley. Acquisition in direct-sequence spread-spectrum communication networks: An asymptotic analysis. *IEEE Transactions on Information Theory*, 39(3):903–912, May 1993.

[38] L. Mailaender and R. A. Iltis. Multiuser code acquisition using parallel decorrelators. In *Conference on Information Science and Systems*, pages 109–114, 1994.

[39] A. Mantravadi and V. V. Veeravalli. On discrete sufficient statistics for asynchronous band-limited CDMA systems. *IEEE Transactions on Commmunications*. Submitted December 1998. Available at http://www.ee.cornell.edu/∼venu/.

[40] S. Y. Miller and S. C. Schwartz. Parameter estimation for asynchronous multiuser communication. In *Conference on Information Science and Systems*, 1989.

[41] R. A. Monzingo and T. W. Miller. *Introduction to Adaptive Arrays*. Wiley, 1980.

[42] A. F. Naguib, A. J. Paulraj, and T. Kailath. Capacity improvement with base-station antenna arrays in cellular CDMA. *IEEE Transactions on Vehicular Technology*, 43(3):691–698, August 1994.

[43] S. G. Nash and A. Sofer. *Linear and Nonlinear Programming*. Preprint, 1995.

[44] L. B. Nelson and H. V. Poor. Iterative multiuser receivers for CDMA channels: An EM-based approach. *IEEE Transactions on Communications*, 44(12):1700–1710, December 1996.

[45] T. Ojanpera and R. Prasad. An overview of air interface multiple access for IMT-2000/UMTS. *IEEE Communications Magazine*, 36(9):82–95, September 1998.

[46] J. A. O'Sullivan. Alternating minimization algorithms: From Blahut-Arimoto to expectation maximization. *Codes, Curves and Signals: Common Threads in Communications*, pages 173–183. A. Vardy Ed. Kluwer Academic Publishers.

[47] M. Pallas and G. Jourdain. Active high resolution time delay estimation for large BT signals. *IEEE Transactions on Signal Processing*, 39(4):781–788, April 1991.

[48] B.-P. Paris. Finite precision decorrelating receivers for multiuser CDMA communication systems. *IEEE Transactions on Communications*, 44(4):496–507, April 1996.

[49] P. Patel and J. Holtzman. Analysis of a simple successive interference cancellation scheme in a DS/CDMA system. *IEEE Journal on Selected Areas in Communications*, 12(5):796–807, June 1994.

[50] A. J. Paulraj and C. B. Papadias. Space-time processing for wireless communications. *IEEE Signal Processing Magazine*, 14(6):49–83, November 1997.

[51] 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.

[52] A. Polydoros and C. L. Weber. A unified approach to serial search spread-spectrum code acquisition - part 1: General theory. *IEEE Transactions on Communications*, 32(5):542–549, May 1984.

[53] H. V. Poor and S. Verdú. Probability of error in MMSE multiuser detection. *IEEE Transactions on Information Theory*, 43(3):858–871, May 1997.

[54] S. Ramakrishna. Optimization of CDMA systems. Ph.D. Thesis, Dept. of Electrical and Computer Engineering, Rutgers University, Piscataway, NJ. May 1998.

[55] F. Rashid-Farrokhi, K. J. R. Liu, and L. Tassiulas. Transmit beamforming for cellular communication systems. In *Conference on Information Science and Systems*, pages 92–97, 1997.

[56] F. Rashid-Farrokhi, L. Tassiulas, and K. J. R. Liu. Joint power control and beamforming in wireless networks using antenna arrays. *IEEE Transactions on Communications*, 46(10):247–256, 1998.

[57] D. Raychaudhuri. Performance analysis of random access packet-switched code division multiple access systems. *IEEE Transactions on Communications*, 29(6):895–901, June 1981.

[58] C. Sankaran and A. Ephremides. Solving a class of optimum multiuser detection problems with polynomial complexity. *IEEE Transactions on Information Theory*, 44(5):1958–1961, September 1998.

[59] C. U. Saraydar and A. Yener. Capacity enhancement for CDMA systems through adaptive cell sectorization. In *IEEE Wireless Communications and Networking Conference, WCNC'99*, pages 1139–1143, September 1999.

[60] R. A. Scholtz. The spread spectrum concept. *IEEE Transactions on Communications*, 25(8):748–755, August 1977.

[61] B. Simeone. Optimal connected partitions on graphs. Tutorial Series at DIMACS Center, Rutgers University, NJ, http://rutcor.rutgers.edu/~boros/LSDO/BrunoSimeone.html, May 1999.

[62] C. De Simone, M. Lucertini, S. Pallottino, and B. Simeone. Fair dissections of spiders, worms, and caterpillars. *Networks*, 20(3):323–344, 1990.

[63] E. G. Strom, S. Parkvall, S. L. Miller, and B. E. Ottersen. Propagation delay estimation in asynchronous direct-sequence code-division multiple access systems. *IEEE Transactions on Communications*, 44(1):84–93, January 1996.

[64] V. G. Subramanian and U. Madhow. Blind demodulation of direct-sequence CDMA signals using an antenna array. In *Conference on Information Science and Systems*, 1996.

[65] S.Verdú. http://www.ee.princeton.edu/~verdu/mud/solutions.html. Multiuser Detection. Solutions to problems.

[66] S. Ulukus. Power control, multiuser detection and interference avoidance in CDMA systems. Ph.D. Thesis, Dept. of Electrical and Computer Engineering, Rutgers University, Piscataway, NJ. July 1998.

[67] S. Ulukus and R. D. Yates. Adaptive power control and MMSE interference suppression. *ACM Wireless Networks, Special Issue on Multiuser Detection in Wireless Communications*, 4(6):489–496, November 1998.

[68] S. Ulukus and R. D. Yates. Iterative signature adaptation for capacity maximization of CDMA systems. In *36th Annual Allerton Conference on Communications, Control and Computing*, September 1998.

[69] 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.

[70] S. Ulukus and R. D. Yates. Stochastic power control for cellular radio systems. *IEEE Transactions on Communications*, 46(6):784–798, June 1998.

[71] 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.

[72] 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.

[73] B. D. Van Veen and K. M. Buckley. Beamforming: A versatile approach to spatial filtering. *IEEE Acoustic, Speech and Signal Processing Magazine*, 5(2):4–24, April 1988.

[74] S. Verdú. Multi-user signal detection. Ph.D. Thesis, Department of Electrical and Computer Engineering, University of Illinois, Urbana Champaign. 1984.

[75] S. Verdú. Minimum probability of error for asynchronous gaussian multiple-access channels. *IEEE Transactions on Information Theory*, 32(1):85–96, January 1986.

[76] S. Verdú. Computational complexity of multiuser detection. *Algorithmica*, 4(4):303–312, 1989.

[77] S. Verdú. Multiuser detection. *Advances in Statistical Signal Processing*, 2:369–409, 1993. JAI Press.

[78] S. Verdú. *Multiuser Detection*. Cambridge University Press, 1998.

[79] P. Viswanath, V. Anatharam, and D. Tse. Optimal sequences, power control and capacity of spread-spectrum systems with multiuser receivers. *IEEE Transactions on Information Theory*, 45(6):968–983, September 1999.

[80] A. J. Viterbi. Spread spectrum communications - myths and realities. *IEEE Communications Magazine*, pages 11–18, May 1979.

[81] X. Wang and H. V. Poor. Space-time multiuser detection in multipath CDMA channels. *IEEE Transactions on Signal Processing*, 47(9):2356–2374, September 1999.

[82] M. Wax and T. Kailath. Detection of signals by information theoretic criteria. *IEEE Transactions on Acoustics, Speech, Signal Processing*, 33(4):387–392, April 1985.

[83] L. C. Whao, P.R. Krishnaiah, and Z. D. Bai. On detection of number of signals in presence of white noise. *Journal of Multivariate Analysis*, 20(1):1–25, October 1986.

[84] J. Winters, J. Salz, and R. D. Gitlin. Impact of antenna diversity on the capacity of wireless communication systems. *IEEE Transactions on Communications*, 42(2/3/4):1740–1751, January/February/March 1994.

[85] T.F. Wong, T. M. Lok, J. S. Lehnert, and M. D. Zoltowski. A linear receiver for direct-sequence spread-spectrum multiple-access systems with antenna arrays and blind adaptation. *IEEE Transactions on Information Theory*, 44(2):659–676, March 1998.

[86] 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.

[87] 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.

[88] A. Yener and R. D. Yates. Multiuser access detection for CDMA systems. In *Conference on Information Science and Systems*, pages 17–22, March 1998.

[89] A. Yener and R. D. Yates. Multiuser access capacity of packet switched CDMA systems. In *IEEE Vehicular Technology Conference*, pages 1846–1850, May 1999.

[90] A. Yener and R. D. Yates. Decorrelating acquisition and access for connectionless CDMA. In *Symposium for Interference rejection and Signal Separation*, March 2000.

[91] A. Yener, R. D. Yates, and S. Ulukus. Joint power control, multiuser detection and beamforming for CDMA systems. In *IEEE Vehicular Technology Conference*, pages 1032–1036, May 1999.

[92] A. Yener, R. D. Yates, and S. Ulukus. A nonlinear programming approach to CDMA multiuser detection. In *33rd Asilomar Conference on Signals, Systems and Computers*, October 1999.

[93] D. M. Young. *Iterative Solution of Large Linear Systems*. Academic Press, 1971.

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

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

[96] X. Zhang and D. Brady. Asymptotic multiuser efficiencies for decision-directed multiuser detectors. *IEEE Transactions on Information Theory*, 44(2):502–515, March 1998.

Vita

Aylin Yener

1991 B.S. in Electrical and Electronics Engineering, Bogazici University, Istanbul, Turkey.

1991 B.S. in Physics, Bogazici University, Istanbul, Turkey.

1991-93 Teaching Assistant, Department of Electrical and Computer Engineering, Rutgers University, Piscataway, NJ.

1993-97 Graduate Research Assistant, Department of Electrical and Computer Engineering, Rutgers University, Piscataway, NJ.

1994 M.S. in Electrical and Computer Engineering, Rutgers University, Piscataway, NJ.

1997 Member of Technical Staff, Lucent Technologies, Whippany, NJ.

1997-2000 Graduate Research Assistant, WINLAB, Department of Electrical and Computer Engineering, Rutgers University, Piscataway, NJ.

1997 A. Yener and C. Rose, "Genetic Algorithms Applied to Cellular Call Admission Problem: Local Policies", *IEEE Transactions on Vehicular Technology*, February 1997, 46(1), pp. 72-79.

1997 A. Yener, C. Rose and R. D. Yates, "Optimum Power Scheduling for CDMA Access Channels", in the Proceedings of *IEEE Global Telecommunications Conference, Globecom'97*, Phoenix, AZ, November 1997.

1998 A. Yener and C. Rose, "Highly Mobile Users and Paging: optimal polling strategies", *IEEE Transactions on Vehicular Technology*, November 1998, 47(4), pp.1251-1257.

1998 A. Yener and R. D. Yates, "Multiuser Access Detection for CDMA Systems", in the Proceedings of *Conference on Information Sciences and Systems, CISS'98*, Princeton, NJ, March 1998.

1999 A. Yener, R. D. Yates and S. Ulukus, "Joint Power Control, Multiuser Detection and Beamforming for CDMA Systems", in the Proceedings of *IEEE Vehicular Technology Conference, VTC'99*, Houston, TX, May 1999.

1999 A. Yener and R. D. Yates, "Multiuser Access Capacity of Packet Switched CDMA Systems", in the Proceedings of *IEEE Vehicular Technology Conference, VTC'99*, Houston, TX, May 1999.

1999 C. U. Saraydar and A. Yener, "Capacity Enhancement for CDMA Systems Through Adaptive Cell Sectorization", in the Proceedings of *IEEE Wireless Communications and Networking Conference, WCNC'99*, New Orleans, LA, September 1999.

1999 A. Yener, R. D. Yates and S. Ulukus, "A Nonlinear Programming Approach to CDMA Multiuser Detection", in the Proceedings of *33rd Asilomar Conference on Signals, Systems and Computers*, Pasific Grove, CA, October 1999.

2000 A. Yener and R. D. Yates, "Decorrelating Acquisition and Access for Connectionless CDMA", in the Proceedings of *Symposium for Interference rejection and Signal Separation, IRSS'00*, Newark, NJ, March 2000.

2000 Ph.D. in Electrical and Computer Engineering, Rutgers University, Piscataway, NJ.