RESOURCE ALLOCATION IN COORDINATED AND UN-COORDINATED WIRELESS SYSTEMS WITH GREEDY OR NON-GREEDY USERS

By

JASVINDER SINGH

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

Prof. Christopher Rose

and approved by

New Brunswick, New Jersey October, 2007

ABSTRACT OF THE DISSERTATION

Resource Allocation in Coordinated and Un-coordinated Wireless Systems with Greedy or Non-greedy Users

by Jasvinder Singh Dissertation Director: Prof. Christopher Rose

In this thesis, we investigate wireless resource optimization problems arising in the context of unlicensed bands. The first half of this thesis assumes a multiple access channel communication model (many transmitters talking to a single receiver), while the second half assumes an interference channel (a collection of multiple interfering transmitter-receiver pairs). The problems considered can further be classified based on the level of coordination available among the devices, and based on the optimization objectives of the devices. The devices are either greedy (they choose their actions to maximize their own utility), or are non-greedy (they choose their actions to maximize their own utility). We investigate four resource optimization scenarios that encompass the possibilities of presence or absence of coordination infrastructure as well as the cases of greedy and non-greedy devices.

As an example of the greedy uncoordinated scenario, we consider application of interference avoidance algorithms in generalized CDMA systems. We introduce variants of standard interference avoidance procedures which produce more easily tracked incremental codewords, and study the response of the system to abrupt changes in the background interference as might be encountered in a practical system.

Next we consider a sensor network scenario where multiple sensors are transmitting correlated symbols to a common receiver, with an objective of minimizing the total mean square error (TMSE) in the symbol estimates at the receiver. The source-channel separation theorem for the point to point case does not hold for this problem and the optimal communication scheme is unknown. We propose a CDMA based transmission scheme that exploits the correlation between the sensors' symbols and facilitates statistical cooperation among the sensors through the choice of their codewords. We give an analytical characterization of the TMSE-minimizing codeword set for this scheme, and compare its performance both with a separation-based scheme, and an information-theoretic upper bound.

The second half of this thesis considers scheduling problems for interfering links where each link employs an ON-OFF modulation scheme in each time slot. First we consider the case of non-greedy interfering links and come up with a distributed scheduling solution that requires no coordination among the links. We prove the convergence of our distributed scheduling algorithm and compare its performance with centralized scheduling. After this, we consider the case of greedy interfering links that coordinate with each other through a mediating authority called the "spectrum server". Each link reports the set of links that interfere with it, based on which the spectrum server constructs the interference graph and finds an optimal schedule for the links (maximizing a certain global objective). Since the links are greedy, they will choose their reports to maximize their individual utilities. We therefore investigate the following natural question: - "Under what conditions is it realistic to assume that the links will send truthful reports?" and present some preliminary results.

Acknowledgements

I would like to express my sincere gratitude to Prof. Christopher Rose for his invaluable advice and constant support at every step of my PhD program. Without his unfailing support, patience, and belief in me, this thesis would not have been possible. I have learnt a great deal from him by observing his ability to formulate imprecise research questions as precise mathematical problems. His style of attacking the most interesting research problems, and his emphasis on the formulation part, rather than the simpler analysis part, has helped me develop my own research tastes. I would also like to thank Prof. Narayan Mandayam, Prof. Roy Yates, and Prof. Predrag Spasojevic for being on my thesis committee, and for their guidance and support during my PhD program. I learnt a lot from them through individual interactions, their courses, and by observing and reflecting on their questions and comments during several WINLAB seminars. Prof. Spasojevic's zeal for pure research and his emphasis on understanding the mathematical structure of problems has always inspired me. I was also deeply influenced by Prof. Yates's and Prof. Mandayam's ability of abstracting out the non-essential parts of any problem and asking the most important questions.

I would also like to thank Prof. Dipankar Raychaudhuri for his help and advice during my PhD program. It has been a privilege and a rewarding learning experience working alongside distinguished researchers and technologists like Prof. Larry Greenstein, Dr. Rich Howard, and Dick Frenkiel. The great support staff at WINLAB including Ivan, Kevin, James, Noreen, Melissa, Elaine, have all helped making my stay at WINLAB very comfortable.

I also want to thank Dr. Krishna Balachandran for serving as an external member on my thesis committee, and for inviting me for a fruitful summer internship at Bell Labs. Thanks are also due to Dr. Joseph Kang, and Dr. Kemal Karakayali for making the internship, a great learning and fun-filled experience for me.

A big part of my PhD learning experience has been interaction and collaboration with fellow students at WINLAB. In particular my discussions with Hithesh, Joy, Chandru, Furuzan, Neville and many others have benefitted me a lot. It has been a privilege to work alongside remarkable colleagues like Ruoheng, Lalitha, Kemal, Omer, Sachin, Pandurang, Suhas, Sanjit, Kishore, Suli, Zhibin, Sylvia, Baik, and so many others. I also thank great friends and former colleagues Pavan, Praveen and Rito, for their everlasting friendship and support. In addition, I thank all my friends outside WINLAB including Sudhanshu, Gatla, Diwakar, Priyanka, and so many others for making my life in US a pleasant experience.

Finally I owe a great debt to my parents, and my brother and sister, for their invaluable support all along. Without their support, I would have never been able to complete this task.

Dedication

To my beloved parents

Table of Contents

Abstract								
Ac	Acknowledgements							
De	Dedication							
1.	Intro	oduction						
2.	Disti	ributed	Incremental Interference Avoidance	5				
	2.1.	System	Model	7				
	2.2.	Codew	ord Steering Schemes	9				
		2.2.1.	Gradient Descent	9				
		2.2.2.	Lagged IA	10				
	2.3.	The Re	cceiver Filter	13				
		2.3.1.	Matched Filter	13				
		2.3.2.	Linear MMSE Filter	14				
		2.3.3.	MMSE Decision Feedback Filter	15				
	2.4.	Numer	ical Results	16				
		2.4.1.	Comparison of Receiver Structures	16				
		2.4.2.	Optimizing Codeword Steering	17				
		2.4.3.	Abrupt Interference Insults	18				
		2.4.4.	SINR Based Codeword Update	20				
		2.4.5.	Feedback Channel Capacity Effects	21				
		2.4.6.	Prior Feedback Bounds	24				
	2.5. Conclusion and Discussion			27				
	2.6.	Append	dix: proof of equation (2.26) (Eigenvalue/Eigenvector Update)	28				
3	A CI	DMA he	ased Scheme for Transmission of Correlated Gaussian Symbols over a					
Ga	ussia	n Multi	ple Access Channel	30				

3.1.	System	Model	31			
3.2.	Total power constraint: Optimal Transmitter Codewords, Power Levels and					
	Receiver Structure					
	3.2.1.	Constructing the Optimal Sequences	37			
	3.2.2.	An Alternate Derivation of TMSE Minimizing Codewords/Power levels	38			
	3.2.3.	Comparison of Optimal Codewords with Random/Orthogonal Codewords	39			
3.3.	Individ	ual Power Constraints	41			
	3.3.1.	Analytical solution for $M = 2$	42			
3.4.	The total power-TMSE trade-off function					
	3.4.1.	Inner Bound: Uncoded transmission over parallel channels with opti-				
		mal power allocation across channels	44			
	3.4.2.	Outer Bound: Point to point system	44			
		Rate-distortion for Correlated Gaussian Random Variables under the				
		assumption of Joint Encoding	45			
		Cooperative Gaussian Multiacess Vector Channel Capacity	45			
	3.4.3.	A Separation based scheme for $M = 2$ case	46			
		Rate-distortion for Correlated Gaussian Random Variables under the				
		assumption of Independent Encoders	46			
		Bounding the Capacity of the Gaussian Multiple Access Channel with				
		Correlated Inputs	47			
	3.4.4.	Performance Comparison	47			
		Low SNR case	47			
		High SNR case	49			
3.5.	Conclu	sion and Discussion	49			
3.6.	Append	dix	50			
	3.6.1.	Majorization: Definitions and Some Key Results	50			
	3.6.2.	Cooperative Gaussian Multiacess Vector Channel Capacity	51			
	3.6.3.	Bivariate Gaussian Rate-Distortion with Independent Encoders	52			

		3.6.4.	Bounding the Capacity of the Gaussian Multiple Access Vector Chan-		
			nel with Correlated Inputs	54	
4.	Ran	ndom Access for Variable Rate Links			
	4.1.	System	n Model	56	
	4.2.	Rate R	egions	58	
		4.2.1.	Centralized scheduling	58	
		4.2.2.	Random Access Scheme	59	
	4.3.	Charac	terization of \mathcal{R}_L^P	60	
		4.3.1.	$L = 2 \ldots \ldots$	60	
			Low Interference Case : $\alpha + \beta \ge 1$	60	
			High Interference Case : $\alpha+\beta<1$	61	
		4.3.2.	$L = 3 \ldots \ldots$	63	
	4.4.	Distrib	uted Algorithm	65	
	4.5.	Conclu	sion and Discussion	67	
	4.6.	Appen	dix	68	
	4.7.	. Derivation of the pareto-boundary of the rate region for $L = 2$ case		68	
		4.7.1.	Low Interference Case : $\alpha + \beta \geq 1$	69	
		4.7.2.	High Interference Case : $\alpha+\beta<1$	69	
	4.8.	Proof o	of lemma 8	69	
5.	Gree	edy Use	rs and Resource Allocation Advisory Services	71	
	5.1.	I. System Model and Problem Statement		71	
		5.1.1.	Server Optimization Problem	72	
			Characterization of the Optimal Solution	72	
		5.1.2.	User Optimization Problem	73	
			Truthful Reporting Problem Statement	74	
	5.2.	Optimal Reporting Strategy for Some Special Server Objective Functions			
		5.2.1.	Scheduling for maximizing weighted sum of time allocations of users .	74	

	5.2.2	. Scheduling for ensuring weighted proportional fairness across user time	
		allocations	5
	5.3. Conc	lusion and Discussion	7
6.	Conclusion	ns and Future Work	9
References		84	
Cı	ırriculum V	7 itae	9

Chapter 1

Introduction

With the ever increasing diversity of scenarios and applications for which wireless communication systems are being deployed, the problem of optimizing radio resources comes in various forms. At one extreme, we might have very low power, low data rate sensor devices operating in a wide-band regime [1]. In this scenario, a trivial solution to the problem of radio interference between devices could be to allocate orthogonal frequency bands to each device ¹ and reduce the problem to the well studied point to point communication channel [2]. In that case, a more important problem would be to use the battery power as efficiently as possible for probing and communicating over the uncertain channel [3]. On the other extreme, we might have a high density of independent, but mutually interfering, high data rate links, for example, in the unlicensed bands [4]. This scenario can be modeled as an interference channel. In this second scenario, scalability of the spectrum coordination algorithm with device density is of utmost importance and therefore we need solutions that are distributed and have very low protocol overhead while having as high spectrum efficiency as possible. Another space of solution possibilities emerges for this scenario if we assume the presence of a Spectrum Server [5,6] that gathers locally observable information from devices in a region, processes it to form a global view of the interference in the spectrum space, and dispatches advice to devices for reducing interference. A third scenario might be an isolated network where multiple devices communicate with a common receiver (and potentially interfere with each other's data streams at the receiver) but any external interference can be ignored. For example, uplink of a cellular system after ignoring the out of the cell interference, or laptops equipped with wireless cards communicating with a single AP, or even sensors reporting their observations to a central authority. Scenarios of this type are usually modeled as a multiple access channel.

In this thesis, we investigate wireless resource optimization problems arising in the context

¹Though this might not be the optimal solution for some objective functions as will be seen in chapter 3.

of unlicensed bands. The first half of this thesis assumes a multiple access channel communication model (many transmitters talking to a single receiver), while the second half assumes an interference channel (a collection of multiple interfering transmitter-receiver pairs). The problems considered can further be classified based on the level of coordination available among the links, and based on the optimization objectives of the links. The devices are either greedy (they choose their actions to maximize their own utility), or are non-greedy (they choose their actions to maximize/satisfy a common social objective). We investigate four resource optimization scenarios that are given below along with examples of previous work for each scenario (see also Figure 1.1).

- 1. Uncoordinated greedy (strategic) links, trying to maximize their individual utilities water-filling in Gaussian multiaccess or interference channel [7–9].
- 2. Uncoordinated non-greedy (truthful) links that honestly follow a set of spectrum etiquettes to enable coexistence – the 802.11 standard, spectrum etiquette protocols [10].
- Coordinated non-greedy users who act under a centralized control. The centralized controller dictates the spectrum usage of each user, either based on some global objective – the Network Utility Maximization (NUM) framework [11].
- 4. Greedy users who find it mutually beneficial to coordinate, maybe directly among themselves, or through a mediator – users bid for spectrum by announcing their valuation of spectrum to a centralized *spectrum server* [12] and the spectrum server designs appropriate pricing mechanisms to ensure that each user bids truthfully [13, 14].

In general, greedy and uncoordinated action by each device might lead to socially inefficient resource allocation. However, Interference Avoidance (IA) algorithms present an exception in that, the greedy objective of maximizing SINR by each user matches the global objective of system capacity maximization for generalized CDMA systems. Each user greedily adapts his codeword in response to feedback from the receiver and can achieve the globally optimal solution in a distributed way. In chapter 2 of this thesis, we investigate practical online implementation of interference avoidance algorithms. We introduce variants of standard interference avoidance procedures which produce more easily tracked incremental codewords, and study the response of the system to abrupt changes in the background interference as might be encountered in a practical system. This problem scenario can be classified as a (G,NC) or (Greedy, Non-coordinated), since only a minimal coordination infrastructure (global feedback from receiver) is assumed, and the users act greedily (to maximize their SNR).

In a CDMA system, the transmitter's symbols are assumed to be independent of each other and the work mentioned in chapter 2 maintains this assumption. However, there may be scenarios in which transmitters send correlated data to a receiver. For example, in the literature for sensor networks [1], one readily comes across a scenario where sensor nodes (analogous to transmitters in CDMA) measure a common physical phenomenon and send their observations (which are correlated) to a central repository. In Chapter 3, we consider such a sensor network scenario where multiple sensors are transmitting correlated symbols to a common receiver, with an objective to minimize the total mean square error (TMSE) in the symbol estimates at the receiver. The source-channel separation theorem for the point to point case does not hold for this problem and the optimal communication scheme is unknown [15]. We propose a CDMA based transmission scheme that exploits the correlation between the sensors' symbols and facilitates statistical cooperation among the sensors through the choice of their codewords. We give an analytical characterization of the TMSE-minimizing codeword set for this scheme, and compare its performance both with a separation-based scheme, and an information-theoretic upper bound. Due to the common global objective of the devices, and the centralized computation of optimal signatures, this problem scenario can be classified as (NG, C).

The second half of this thesis considers scheduling problems for interfering links where each link employs an ON-OFF modulation scheme in each time slot. In chapter 4, we consider the case of non-greedy links and come up with a distributed scheduling solution that requires no coordination among the links (NG, NC). Then, in chapter 5, we look at a scenario where greedy links coordinate with each other (G,C) through a mediating authority called the "spectrum server" and investigate the issue of truthful reporting.

In the case of non-greedy interfering links considered in chapter 4, an active link obtains a data rate determined by the interference from other active links in the network. We consider a random access scheme, each link decides to transmit in each time slot with certain probability, and compare its achievable throughput region with the upper bound given by centralized



Figure 1.1: Thesis Organization

scheduling. We then present a distributed implementation of the random access scheme that achieves all feasible rate vectors in the throughput region. The distributed algorithm consists of an iteration, where each link updates its transmission probability based only on its measured throughput, and is provably convergent under certain conditions.

In chapter 6, for the case of greedy links, interference is modeled as a directional interference graph with the links as vertices. An edge from vertex A to vertex B implies that link B can transmit successfully only if link A is not transmitting. Each link reports its incoming edges (interfering neighbors) to the server, based on which the server constructs the interference graph and constructs an optimal schedule for the links (that maximizes a certain global objective). Since the links are greedy, they will choose their reports to maximize their individual utilities. We therefore investigate the following natural questions: - "Under what conditions is it realistic to assume that the links will send truthful reports? If the links have an incentive to lie, can we design mechanisms and provide them with counter-incentives to enforce truthful reporting?" We formulate this problem and present some preliminary results based on our work.

Chapter 2

Distributed Incremental Interference Avoidance

The past decade has witnessed tremendous research activity in the area of joint signaturereceiver optimization for single-cell Direct-Sequence Code Division Multiple Access (DS-CDMA) systems [16–22]. In a single cell system with a fixed number of users, one can imagine solving the signature-receiver optimization problem off-line in a centralized manner at the receiver and feeding the optimal codewords back to the mobiles. However, this approach could lead to onerous feedback bandwidth requirements [23–25], if the frequency of users joining or leaving the network is high or the background interference changes rapidly. Therefore one must take limitations on the feedback channel capacity into account and quantize the codewords accordingly. Reference [25] provides an asymptotic analysis of limited feedback channel capacity and achievable SINR at the transmitters by considering a Random Vector Quantization (RVQ) scheme for codeword feedback. Since RVQ has very high computational complexity (codeword search complexity at each mobile is exponential in the the number of feedback bits), [25] also considers a reduced rank scheme where codewords lie in a lower-dimensional space assumed known both to the transmitters and the receiver.

In an unlicensed band where centralized control is absent or difficult to implement, computing and feeding back codewords to each individual user might not be feasible. In addition, dynamically changing interference levels might impose strict robustness requirements on the system that can only be satisfied by a distributed and on-line signature-receiver optimization algorithm. The enabling technology for such a system solution would be *software radios* [26] that promise to provide *on the fly* control over the modulation/demodulation schemes of the transceivers.

Interference Avoidance (IA) is a class of algorithms [27], [22], [28], [21] that greedily adapts the user codewords in response to feedback from the receiver and can achieve globally optimal solution in a distributed way. The basic idea of IA is simple. Each user waveform is represented as a linear combination of orthonormal basis functions which span the signal space.

The set of real valued coefficients used to represent the waveform is a *codeword* and each user changes its codeword greedily to improve SINR. We find it most useful to think in terms of the inverse SINR, given by the Rayleigh quotient [29, pp. 253] as

$$\chi_k = \frac{\mathbf{s}_k^\top \mathbf{R}_k \mathbf{s}_k}{\mathbf{s}_k^\top \mathbf{s}_k} \tag{2.1}$$

where \mathbf{R}_k is the interference plus noise covariance seen by codeword k at the receiver. Since it is known that the Rayleigh quotient of a matrix is minimized by the "minimum eigenvector" associated with the minimum eigenvalue, we can see that χ_k will be maximized if codeword \mathbf{s}_k is replaced by the minimum eigenvector of \mathbf{R}_k . Moving toward a state where all codewords are simultaneously minimum eigenvectors of their respective covariances is the overall goal of IA – and one which is achieved by iterative application of a variety of codeword update procedures [21,27,28].

However, as pointed out in [30], IA algorithms are unsuitable for direct on-line implementation because of either of the following assumptions.

- a) As the signature sequences are updated, the receiver filters are changed to be the corresponding matched filters instantaneously.
- b) The iterations are run off-line only in terms of the signature sequences, and once the signature sequences converge to an optimum set, corresponding matched filters are deployed as receivers.

Here we consider the scenario where users know their channels, the receiver covariance (or the received vector) is broadcast by the receiver and users employ some IA procedure to improve their codewords.

Since the receiver must adapt to codeword changes, and the only information the receiver has are user transmissions, we note that IA algorithms such as the MMSE [21] or eigen algorithm [28] can cause abrupt changes in codewords which might be difficult for receivers to track without disruption of associated data streams. To minimize such disruption we introduce the following two schemes which allow only incremental adjustments to codewords:

• Gradient Descent IA: codewords are adjusted to most rapidly reduce the inverse SINR.

• Lagged IA: codewords are adjusted in the direction of the optimal codeword.

Based on these methods, we propose a simple structure for practical distributed IA. Overall, the method is reminiscent of adaptive equalization and appears robust to reasonably abrupt changes in the interference environment as well as the amount of broadcast feedback provided by the transmitter.

2.1 System Model

We consider a system with block diagram as shown in FIGURE 2.1. Information symbols modulate transmitted codewords and receivers use separate filters for each user to produce an estimate of transmitted symbols for each user. Joint decoding is *not* assumed. The receiver has no *a priori* information about the transmitter codewords and starts with randomly selected filter coefficients. The transmitters send a known training sequence with which the receiver iteratively refines the receiver filters based on typical error minimization criteria. During the training phase, the system equations are,

$$\mathbf{r}(n) = \mathbf{S}(n)\mathbf{b}(n) + \mathbf{v}(n) \tag{2.2}$$

$$\mathbf{C}(n+1) = f[\mathbf{C}(n), \mathbf{r}(n), \mathbf{b}(n)]$$
(2.3)

where, n is the time index and $\mathbf{S}(n) = [\mathbf{s}_1(n), \mathbf{s}_2(n), ., \mathbf{s}_M(n)]$ is the codeword matrix $[L \times M]$ whose columns are the unit norm user codewords. $\mathbf{C}(n) = [\mathbf{c}_1(n), \mathbf{c}_2(n), ..., \mathbf{c}_M(n)]$ is the receiver filter matrix $[L \times M]$ whose columns are receiver filters corresponding to the transmitter codewords. $\mathbf{r}(n)$ is the received signal vector at the receiver $[L \times 1]$. $\mathbf{b}(n)$ is the vector containing symbols sent by each user $[1 \times M]$. $\mathbf{v}(n)$ is assumed zero mean and white with covariance $\mathbf{K}_{\mathbf{v}}(n, k) = N_0 \mathbf{I}$. $f[\cdot]$ is the receiver filter update scheme used.

After training, the receiver measures the received signal covariance \mathbf{R} and broadcasts it to all users periodically. Alternately, the receiver could more frequently broadcast the received vector $\mathbf{r}(n)$ and let each user construct receiver covariance estimates. We assume that each user knows its channel, so the feedback can be used by the transmitters to steer transmitted codewords toward higher SINR. The receiver decodes the symbols sent by the users and continues updating \mathbf{C} as it did during training but now assuming that the decoded symbols are correct. This is exactly analogous to the operation of a typical adaptive equalizer [31].



Figure 2.1: System Model

The system equations are,

$$\mathbf{r}(n) = \mathbf{S}(n)\mathbf{b}(n) + \mathbf{v}(n) \tag{2.4}$$

$$\widehat{\mathbf{b}}(n) = \operatorname{sgn}\left[\widetilde{\mathbf{b}}(n)\right] = \operatorname{sgn}\left[\mathbf{r}(n)^{\top}\mathbf{C}(n)\right]$$
 (2.5)

$$\mathbf{C}(n+1) = f[\mathbf{C}(n), \mathbf{r}(n), \mathbf{b}(n)]$$
(2.6)

where $\tilde{\mathbf{b}}(n)$ is the soft estimate of the transmitted symbol vector $\mathbf{b}(n)$ and $\hat{\mathbf{b}}(n)$ is the corresponding hard estimate. The received covariance is defined as

$$\mathbf{R}(n) = E\left[\mathbf{r}(n)\mathbf{r}^{\top}(n)\right]$$
(2.7)

and the covariance seen by a particular user k as

$$\mathbf{R}_k(n) = \mathbf{R}(n) - \mathbf{s}_k(n)\mathbf{s}_k(n)^\top$$
(2.8)

We need to estimate $\mathbf{R}(n)$ from time samples $\mathbf{r}(n)$. If the received vector sequence $\mathbf{r}(n)$ were stationary, then we could have replaced the ensemble average $E[\cdot]$ above by the time average. However, owing to the codeword updates, $\mathbf{r}(n)$ is no longer stationary, so an exponentially weighted average

$$\hat{\mathbf{R}}(n) = \frac{1}{N} \sum_{k=1}^{N} \xi^{N-k} \mathbf{r}(k) \mathbf{r}(k)^{\top}$$
(2.9)

seems more appropriate where ξ is a "forgetting factor" [32].

The above estimate can be computed recursively in the following manner

$$\hat{\mathbf{R}}(n) = (1 - \xi)\hat{\mathbf{R}}(n - 1) + \xi \mathbf{r}(n)\mathbf{r}(n)^{\top}$$
(2.10)

Likewise the individual user k covariance estimate is defined as

$$\hat{\mathbf{R}}_k(n) = \hat{\mathbf{R}}(n) - \mathbf{s}_k(n)\mathbf{s}_k(n)^\top$$
(2.11)

Finally, the codeword update equation is

$$\mathbf{s}_k(n+1) = g\left[\mathbf{s}_k(n), \mathbf{R}_k(n)\right] \quad k \in \{1, 2, ..., M$$
(2.12)

where $g[\cdot]$ is the codeword steering scheme used at the transmitters.

2.2 Codeword Steering Schemes

Our aim is to provide schemes which make small performance-improving adjustments to the codewords at each time step. Each user's greedy performance objective is SINR maximization or inverse SINR minimization. We consider the following two possible schemes.

2.2.1 Gradient Descent

The inverse SINR for the k^{th} user is

$$\chi_k = \frac{\mathbf{s}_k^\top \mathbf{R}_k \mathbf{s}_k}{\mathbf{s}_k^\top \mathbf{s}_k} \tag{2.13}$$

and its gradient with respect to the codeword components $\{s_{kj}\}$ is

$$\nabla \chi_k = \frac{2[\mathbf{s}_k^\top \mathbf{s}_k \mathbf{R}_k \mathbf{s}_k - (\mathbf{s}_k^\top \mathbf{R}_k \mathbf{s}_k) \mathbf{s}_k]}{(\mathbf{s}_k^\top \mathbf{s}_k)^2}$$
(2.14)

Therefore, the iteration $\mathbf{s}_k(n+1) = \mathbf{s}_k(n) - \nu \nabla \chi_k$, with ν a suitably small constant, increases SINR.

Now, even if we impose the unit power constraint on $s_k(n+1)$ by normalization, SINR still increases because normalization does not change the value of χ_k . So, our iteration becomes,

$$\mathbf{s}_k(n+1) = \frac{\mathbf{s}_k(n) - \nu \nabla \chi_k}{\|\mathbf{s}_k(n) - \nu \nabla \chi_k\|}$$
(2.15)

Now, since convergence is guaranteed if an algorithm decreases total squared correlation (TSC) [22, 27, 28] where

$$TSC(n) = Trace \left[\mathbf{R}(n)\mathbf{R}(n)^{\top} \right]$$
(2.16)

we need to show that increasing SINR decreases TSC. Writing

$$\mathbf{R}(n) = \mathbf{R}_k(n) + \mathbf{s}_k(n)\mathbf{s}_k^{\dagger}(n)$$
(2.17)

and noting that $\|\mathbf{s}_k(n)\| = \|\mathbf{s}_k(n+1)\| = 1$, it is easy to show [33] that

$$\Delta \text{TSC} = \text{TSC}(n+1) - \text{TSC}(n) = 2\Delta \chi_k \le 0$$
(2.18)

and the result follows.

Choice of ν dictates the rate at which the codewords change. Ideally we would like to chose the largest possible ν which still allows the receiver to accurately track codewords. However, convergence of gradient descent schemes require ν to be chosen sufficiently small. An improvement would be a convergence condition which did not depend on the *step size* parameter ν . Such a steering scheme is presented next.

2.2.2 Lagged IA

As seen previously, the optimal codeword for user k is given by the eigenvector corresponding to the minimum eigenvalue of the channel interference matrix \mathbf{R}_k . Let us denote this eigenvector by \mathbf{s}_k^* . Assuming codewords of other users remain fixed, user k can increase its SINR by steering its codeword toward \mathbf{s}_k^* using the iteration

$$\mathbf{s}_{k}(n+1) = \frac{\alpha \mathbf{s}_{k}(n) + m\beta \mathbf{s}_{k}^{*}}{\left|\alpha \mathbf{s}_{k}(n) + m\beta \mathbf{s}_{k}^{*}\right|}$$
(2.19)

 $\alpha, \beta \in \Re^+$. We have defined $m = \operatorname{sgn} [\rho_k(n)]$ with $\rho_k(n) = \mathbf{s}_k^\top(n)\mathbf{s}_k^*$ and note that $|\rho_k(n)| \le 1$ since $|\mathbf{s}_k^*| = |\mathbf{s}_k(n)| = 1$.

Equation (2.19) has a simple and intuitive geometric meaning: $\mathbf{s}_k(n+1)$ represents a step towards the closest optimal codeword $m\mathbf{s}_k^*$ along the arc joining $m\mathbf{s}_k^*$ and $\mathbf{s}_k(n)$. That is, $\mathbf{s}_k(n)$ and $m\mathbf{s}_k^*$ share the same half plane. Formally, we have

Theorem 1 User k's SINR will increase under the iteration of equation (2.19) where \mathbf{s}_k^* is a minimum eigenvector of \mathbf{R}_k .

Proof: \mathbf{R}_k is a covariance matrix and therefore positive-semidefinite with orthonormal eigenvectors which span the signal space. Let $[(\lambda_1, x_1), (\lambda_2, x_2), ..., (\lambda_L, x_L)]$ be the eigenvalues and

eigenvectors of \mathbf{R}_k such that $\lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_L$ ($\mathbf{s}_k^* = x_1$). Thus

$$\mathbf{s}_{k}^{\top}(n)\mathbf{R}_{k}\mathbf{s}_{k}(n) \ge (\mathbf{s}_{k}^{*})^{\top}\mathbf{R}_{k}\mathbf{s}_{k}^{*} = \lambda_{1}$$
(2.20)

since $|\mathbf{s}_k(n)| = |\mathbf{s}_k^*| = 1$. The change in inverse SINR is

$$\Delta \chi_{k} = \mathbf{s}_{k}^{\top}(n+1)\mathbf{R}_{k}\mathbf{s}_{k}(n+1) - \mathbf{s}_{k}^{\top}(n)\mathbf{R}_{k}\mathbf{s}_{k}(n)$$

$$= \frac{\left(\alpha \mathbf{s}_{k}(n) + m\beta \mathbf{s}_{k}^{*}\right)^{\top}}{\left|\alpha \mathbf{s}_{k}(n) + m\beta \mathbf{s}_{k}^{*}\right|}\mathbf{R}_{k}\frac{\left(\alpha \mathbf{s}_{k}(n) + m\beta \mathbf{s}_{k}^{*}\right)}{\left|\alpha \mathbf{s}_{k}(n) + m\beta \mathbf{s}_{k}^{*}\right|} - \mathbf{s}_{k}(n)^{\top}\mathbf{R}_{k}\mathbf{s}_{k}(n)$$
(2.21)

Let us use κ^2 to denote

$$\kappa^2 = |\alpha \mathbf{s}_k(n) + m\beta \mathbf{s}_k^*|^2 = \alpha^2 + \beta^2 + 2\alpha\beta|\rho(n)|$$
(2.22)

Then we can write

$$\kappa^{2} \Delta \chi_{k} = \alpha^{2} \mathbf{s}_{k}^{\top}(n) \mathbf{R}_{k} \mathbf{s}_{k}(n) - \kappa^{2} \mathbf{s}_{k}(n)^{\top} \mathbf{R}_{k} \mathbf{s}_{k}(n) + m\alpha \beta \mathbf{s}_{k}(n)^{\top} \mathbf{R}_{k} \mathbf{s}_{k}^{*} + m\alpha \beta \mathbf{s}_{k}^{*T} \mathbf{R}_{k} \mathbf{s}_{k}(n) + \beta^{2} \mathbf{s}_{k}^{*T} \mathbf{R}_{k} \mathbf{s}_{k}^{*} = (\alpha^{2} - \kappa^{2}) \mathbf{s}_{k}(n)^{\top} \mathbf{R}_{k} \mathbf{s}_{k}(n) + \lambda_{1}(\kappa^{2} - \alpha^{2}) = (\kappa^{2} - \alpha^{2}) \left(\lambda_{1} - \mathbf{s}_{k}(n)^{\top} \mathbf{R}_{k} \mathbf{s}_{k}(n)\right)$$
(2.23)

which for $\alpha, \beta > 0$ is always less than or equal to zero by equation (2.22) and equation (2.20) To ensure that codewords change incrementally we require $|\alpha| \gg |\beta|$. Increasing SINR implies TSC decrease as noted in the previous section, so convergence is guaranteed.

Note that the above convergence results are *deterministic* in nature i.e. they assume that \mathbf{R} is exact rather than an estimate. A rigorous stochastic convergence proof would require taking into account the closed loop nature of the system (due to $\mathbf{r}(n)$ feedback), and might possibly be developed along the lines of [34].

Since the codeword steering schemes are to be implemented at the mobiles, computational complexity is a very important issue. We note that the gradient descent iteration does not explicitly require calculation of the minimum eigenvector, and thus offers a computational advantage over the lagged IA scheme. In order to make the lagged IA scheme computationally efficient, we can use results from matrix perturbation theory [35]. Matrix perturbation theory allows us to compute the effect of small perturbations in a matrix to the resulting perturbations in its eigenvalues and eigenvectors without redoing the complete eigen-decomposition for the perturbed matrix. Specifically, let \mathbf{u}_{i0} and γ_{i0} be the *i*th eigenvector and eigenvalue of an $[L \times L]$

hermitian matrix \mathbf{A}_0 and let \mathbf{u}_{i1} and γ_{i1} be the corresponding eigenvector and eigenvalue of a perturbed hermitian matrix $\mathbf{A}_1 = \mathbf{A}_0 + \xi \mathbf{A}_p$ such that $\xi \ll 1$. Then the following first order Taylor series approximation can be made

$$\mathbf{u}_{i1} = \mathbf{u}_{i0} + \xi \mathbf{u}_{ip} \tag{2.24}$$

$$\gamma_{i1} = \gamma_{i0} + \xi \gamma_{ip} \tag{2.25}$$

where \mathbf{u}_{ip} and γ_{ip} are given by ([36], see appendix),

$$\gamma_{ip} = \mathbf{u}_{i0}^{\top} \mathbf{A}_p \mathbf{u}_{i0} \tag{2.26}$$

$$\mathbf{u}_{ip} = \sum_{j=1}^{L} \theta_{ji} \mathbf{u}_{j0} \tag{2.27}$$

where

$$\theta_{ji} = \mathbf{u}_{j0}^{\top} \mathbf{u}_{ip} = \frac{-\mathbf{u}_{j0}^{\top} \mathbf{A}_p \mathbf{u}_{j0}}{(\gamma_{j0} - \gamma_{i0})}$$
(2.28)

We are interested in efficiently computing the minimum eigenvector of $\hat{\mathbf{R}}_k(n)$ in terms of that of $\hat{\mathbf{R}}_k(n-1)$. Combining (2.10) and (2.11) gives

$$\hat{\mathbf{R}}_k(n) = \hat{\mathbf{R}}_k(n-1) + \xi \left(\mathbf{r}(n)\mathbf{r}(n)^\top - \hat{\mathbf{R}}_k(n-1) - \mathbf{s}_k(n-1)\mathbf{s}_k(n-1)^\top \right)$$
(2.29)

$$+\left(\mathbf{s}_{k}(n-1)\mathbf{s}_{k}(n-1)^{\top}-\mathbf{s}_{k}(n)\mathbf{s}_{k}(n)^{\top}\right)$$
(2.30)

The above equation does not lend itself to a straightforward application of matrix perturbation results, due to the presence of the term $\mathbf{s}_k(n-1)\mathbf{s}_k(n-1)^{\top} - \mathbf{s}_k(n)\mathbf{s}_k(n)^{\top 1}$. However, the problem can be circumvented by estimating $\hat{\mathbf{R}}_k(n)$ directly at the k^{th} mobile, instead of first estimating $\hat{\mathbf{R}}(n)$.

Letting $\mathbf{r}_k(n) = \mathbf{r}(n) - b_k(n)\mathbf{s}_k(n)$ we can write

$$\mathbf{R}_{k}(n) = E\left[\mathbf{r}_{k}(n)\mathbf{r}_{k}(n)^{\top}\right]$$
(2.31)

Mobile k computes $\mathbf{r}_k(n)$ after receiving the feedback vector $\mathbf{r}(n)$, and then estimates $\hat{\mathbf{R}}_k(n)$ recursively in the following manner.

$$\hat{\mathbf{R}}_k(n) = \hat{\mathbf{R}}_k(n-1) + \xi(\mathbf{r}_k(n)\mathbf{r}_k(n)^\top - \hat{\mathbf{R}}_k(n-1))$$
(2.32)

¹One must show that this term is of order $O(\xi)$ in order to apply the matrix perturbation results

Now let $[(\lambda_1(n), x_1(n)), (\lambda_2(n), x_2(n)), ..., (\lambda_L(n), x_L(n))]$ be the eigenvalues and eigenvectors of $\mathbf{R}_k(n)$ such that $\lambda_1(n) \leq \lambda_2(n) \leq ... \leq \lambda_L(n)$ ($\mathbf{s}_k^* = x_1(n)$). Then using (2.26), we can write

$$\mathbf{x}_{i}(n) = \sum_{j=1}^{L} \frac{\mathbf{x}_{i}(n-1)^{\top} \left[\hat{\mathbf{R}}_{k}(n-1) - \mathbf{r}_{k}(n)\mathbf{r}_{k}(n)^{\top} \right] \mathbf{x}_{i}(n-1)}{[\lambda_{j}(n-1) - \lambda_{i}(n-1)]} \mathbf{x}_{i}(n-1)$$
(2.33)

and

$$\lambda_i(n) = \mathbf{x}_i(n-1)^\top \left[\mathbf{r}_k(n) \mathbf{r}_k(n)^\top - \hat{\mathbf{R}}_k(n-1) \right] \mathbf{x}_i(n-1)$$
(2.34)

Substituting $\mathbf{s}_k^* = x_1(n)$ in (2.19), we find that the calculation of minimum eigenvector has been replaced by simple multiplications and divisions and concomitantly reduced complexity.

2.3 The Receiver Filter

We consider adaptive formulations of three types of receiver filters – matched, linear MMSE, and Decision Feedback MMSE. For the first two filters, we investigate two types of adaptive algorithms [32]: LMS (Least Mean Squared) and RLS (Recursive Least Squares). The LMS, or Least Mean Squared algorithm is an approximation of the steepest descent algorithm which uses an instantaneous estimate of the gradient vector. The estimate of the gradient is based on sample values of the received vector and an error signal. The algorithm iterates over each coefficient in the filter, moving it in the direction of the approximated gradient.

On the other hand, RLS, or Recursive Least Squares is an exact algorithm in the sense that at each time instant the filter coefficients are optimal for the given observations. It uses the matrix inversion identity [32] to efficiently compute the optimal filter at time n in terms of the optimal filter at time (n - 1) and the observations at time n. The RLS algorithm uses $O(L^2)$ operations per iteration as opposed to O(L) used by LMS but has a faster convergence rate.

2.3.1 Matched Filter

For optimal codeword ensembles, the matched filter ($\mathbf{c}_i = \mathbf{s}_i$) is the optimal linear receiver [37]. The LMS filter update equation can be obtained by using gradient descent to minimize $E[\mathbf{e}(n)^2]$, where

$$\mathbf{e}(n) = \mathbf{r}(n) - \mathbf{C}(n)\mathbf{b}(n) \tag{2.35}$$

so that the filter update equation is

$$\mathbf{C}(n+1) = \mathbf{C}(n) + \mu \mathbf{e}(n)\mathbf{b}(n)^{\top}$$
(2.36)

for some suitable constant μ . Note that after training, we use $\mathbf{b}(n) = \hat{\mathbf{b}}(n)$ and continue updating **C** in the same manner.

The RLS algorithm includes a forgetting factor ξ that allows it to work for non-stationary signals. The error squared term that has to be minimized after time instant n is given by $\frac{1}{n} \sum_{k=1}^{n} \xi^{n-k} |\mathbf{e}(k)|^2$. The optimal filter after n time instants is given by

$$\mathbf{C}(n+1) = \mathbf{Q}(n)^{-1}\mathbf{P}(n) \tag{2.37}$$

where

$$\mathbf{Q}(n) = \sum_{k=1}^{n} \xi^{n-k} \mathbf{b}(k) \mathbf{b}(k)^{\top}$$
(2.38)

$$\mathbf{P}(n) = \sum_{k=1}^{n} \xi^{n-k} \mathbf{r}(k) \mathbf{b}(k)^{\top}$$
(2.39)

Since $\mathbf{Q}(n) = \xi \mathbf{Q}(n-1) + \mathbf{b}(n)\mathbf{b}(n)^{\top}$, using the matrix inversion identity [32] we can write

$$\mathbf{Q}(n)^{-1} = \xi^{-1} \mathbf{Q}(n-1)^{-1} - \frac{\xi^{-2} \mathbf{Q}(n-1)^{-1} \mathbf{b}(n) \mathbf{b}(n)^{\top} \mathbf{Q}(n-1)^{-1}}{(1+\xi^{-1} \mathbf{b}(n)^{\top} \mathbf{Q}(n-1)^{-1} \mathbf{b}(n))}$$
(2.40)

Also $\mathbf{P}(n) = \xi \mathbf{P}(n-1) + \mathbf{r}(n)\mathbf{b}(n)^{\top}$ and therefore C(n+1) can be computed efficiently.

2.3.2 Linear MMSE Filter

The MMSE filter for the k^{th} user is defined as the vector \mathbf{c}_k which minimizes $E[e_k(n)^2]$ where $e_k(n)$ is now defined as,

$$e_k(n) = b_k(n) - \mathbf{r}(n)^{\top} \mathbf{c}_k(n)$$
(2.41)

Note that in the above equation, $\mathbf{r}(n)^{\top}\mathbf{c}_k(n)$ represents a soft estimate for $b_k(n)$ and hence $e_k(n)$ is the symbol estimation error for user k. The LMS filter update equation in this case is

$$\mathbf{C}(n+1) = \mathbf{C}(n) + \mu \mathbf{e}(n)\mathbf{r}(n)$$
(2.42)

where,

$$\mathbf{e}(n) = \mathbf{b}(n) - \mathbf{r}(n)^{\top} \mathbf{C}(n)$$
(2.43)



Figure 2.2: Block diagram of multiuser decision-feedback detector.

After training, we use $\mathbf{b}(n) = \widehat{\mathbf{b}}(n) = \operatorname{sgn}[\widetilde{\mathbf{b}}(n)]$ and continue updating C in the same manner. The RLS update equations in this case are given by

$$\mathbf{Q}(n)^{-1} = \xi^{-1} \mathbf{Q}(n-1)^{-1} - \frac{\xi^{-2} \mathbf{Q}(n-1)^{-1} \mathbf{r}(n) \mathbf{r}(n)^{\top} \mathbf{Q}(n-1)^{-1}}{(1+\xi^{-1} \mathbf{r}(n)^{\top} \mathbf{Q}(n-1)^{-1} \mathbf{r}(n))}$$
(2.44)

$$\mathbf{P}(n) = \xi \mathbf{P}(n-1) + \mathbf{b}(n)\mathbf{r}(n)^{\top}$$
(2.45)

2.3.3 MMSE Decision Feedback Filter

Here we consider a decision feedback (DF) based parallel interference canceler [38]. The general structure of a multiuser decision-feedback detector is given in FIGURE 2.2 which describes both successive and parallel interference cancellation as well as more general detection schemes [38]. The receiver consists of a feedforward filter F and a backward (feedback) filter B which together represent filter C of our system model.

Since here the received powers for all users are equal, parellel interference cancellation makes more intuitive sense than successive interference cancellation. For parallel operation, the multiuser detector takes the form as shown in FIGURE 2.3 where *B* is constrained to have all zero diagonal elements [38]. Note that *F* in FIGURE 2.2 corresponds to the product X(I+B)in FIGURE 2.3. Tentative bit estimates $b_1(n)$ are first obtained using the linear MMSE filter i.e. $b_1(n) = r(n)^{\top}X(n)$ and then fed back to refine the estimates.

For the form shown in FIGURE 2.3, the LMS filter update equations can be found by using gradient descent to minimize $E[e(n)^2]$ where e(n) is now defined as,

$$e(n) = b(n) - [\mathbf{F}(n)\mathbf{r}(n) - \mathbf{B}(n)b_1(n)]$$
(2.46)



Figure 2.3: Parallel DFD.

The feedforward filter update equation is,

$$\mathbf{F}(n+1) = \mathbf{F}(n) + \mu_F e(n) \mathbf{r}(n)^{\top}$$
(2.47)

The k^{th} row of the backward filter is updated as follows.

$$\mathbf{B}_{ki}(n+1) = \mathbf{B}_{ki}(n) - \mu_B e_k(n) b_{1i}(n) \qquad \forall i(i \neq k)$$
(2.48)

The linear MMSE filter X can then be computed as,

$$\mathbf{X}(n+1) = (I + \mathbf{B}(n+1))^{-1} \mathbf{F}(n+1)$$
(2.49)

2.4 Numerical Results

2.4.1 Comparison of Receiver Structures

We first consider users with static randomly chosen codewords (no codeword adaptation) and compare the performance of various receiver structures. FIGURE 2.4 shows the BER vs SNR curves for an underloaded system (4 users in 12 dimensions) with three different receiver structures, namely, matched filter, linear MMSE and DF MMSE. FIGURE 2.5 repeats the same experiment with an overloaded system (14 users in 12 dimensions). Note that the same transmitter codeword sets are used in all three receiver filter schemes. As might be expected for unadapted codewords, MMSE filtering is superior to matched filtering since it mitigates the effects of any randomly high correlation between user codewords. DF MMSE outperforms linear MMSE at high SNR's but gives degraded performance at low SNR's owning to poor initial bit



Figure 2.4: Matched filter and MMSE filter comparison. BER vs SNR plot for 4 *static* users in 12 dimensions. $\mu = 0.001, \mu_F = 0.01, \mu_B = 0.01$.

estimates. Also, we note that the performance gain of DF MMSE over linear MMSE increases as the system load $(\frac{M}{L})$ is increased. These results are in agreement with those obtained in [38]. Since the linear MMSE filter provided uniformly good performance at low computational complexity all subsequent experiments use the linear MMSE receiver filter structure.

2.4.2 Optimizing Codeword Steering

We consider an overloaded system (14 users in 12 dimensions) and look at the codeword steering performance for both gradient descent and lagged IA. The choice of ξ in equation (2.10) depends on how fast the statistics of the $\mathbf{r}(n)$ process are changing. If the statistics are changing very slowly and the process is almost stationary, then a small value of ξ should be used. We heuristically pick ξ as 0.02 in our experiments. We seek minimum codeword convergence time without incurring high BER at the receiver and chose a nominal value for SNR as 20 dB since we expect such systems to be interference rather than noise limited. The initial codewords are chosen randomly as before. We vary the steering step size control parameters (ν for gradient



Figure 2.5: Matched filter and MMSE filter comparison. BER vs SNR plot for 14 *static* users in 12 dimensions. $\mu = 0.001, \mu_F = 0.01, \mu_B = 0.01$.

descent, α for lagged IA with β fixed) and measure the BER while the codewords are being adapted. If the steering step size is too large, receiver cannot track the codewords which results in a high BER. On the other hand, if the step size is too small, codewords converge very slowly to their optimal values and high correlation leads to high BER during the finite measurement window. Suitable steering step control parameters for this system were found to be in the range $10^{-3} < \nu < 10^{-2}$ and $40 < \alpha < 300$.

2.4.3 Abrupt Interference Insults

We also considered the effect of adding static interference (in the form of a new user codeword) to the system after the user codewords have stabilized in a WBE set. First, we consider an underloaded system with 4 users in 12 dimensions. As before, the SNR for each user is chosen to be 20 dB. We used the following values for steering step control parameters: ($\nu = 0.005$, $\alpha = 100$, $\beta = 1$). FIGURE 2.6 shows the average inverse SIR variation with time. The first 10000 bit intervals comprise the post training interval where transmitters are adapting their codewords



Figure 2.6: Average inverse SIR variation for 4 users in 12 dimensions. New user added after 10000 bit intervals. top plot: gradient descent with $\nu = 0.005$. bottom plot: lagged IA with $\alpha = 100, \beta = 1$.

using IA. Since this is an underloaded system, we expect the inverse SIR value to become small (nearly orthogonal user codewords). At t = 10001, a user with a random but subsequently fixed codeword is introduced. From t = 10001 onward we see that IA reduces the average SIR to near zero within 500 symbol intervals for gradient descent and 1000 symbol intervals for lagged IA. FIGURE 2.6 also compares the performance for exact covariance feedback and for estimates of different qualities ($\xi = 0.02, 1$). Note that $\xi = 1$ corresponds to an instantaneous – and therefore highly volatile – estimate ie. $\mathbf{R}(n) = \mathbf{r}(n)\mathbf{r}(n)^{\top}$.

FIGURE 2.7 shows the same plots for an overloaded system (14 users in 12 dimensions). For exactly known and average covariance before the introduction of static interference, the average inverse SIR value is approximately $0.1667 = \frac{14-12}{12}$), the theoretically optimal value associated with a Welch bound equality codeword set (see [28]). The performance of the instantaneous covariance feedback is slightly worse, but not unreasonable.

Adding sudden interference does not greatly increase the average SIR and more importantly,



Figure 2.7: Average inverse SIR variation for 14 users in 12 dimensions. New user added after 10000 bit intervals. top plot: gradient descent with $\nu = 0.005$. bottom plot: lagged IA with $\alpha = 100, \beta = 1$.

does not greatly disrupt the data streams of other users – at least in as much as no retraining was required, even with the imprecise instantaneous covariance. Following the interference insult, the system quickly settles down to the theoretically minimum inverse SIR of $\frac{15-12}{12} = 0.25$ in the case of exact and averaged covariance feedback, and once again, performs a bit more poorly for instantaneous covariance feedback.

We also plot the variance of inverse SIR among users vs time in FIGURE 2.8 which suggests that codeword adaptation rapidly equalizes user SIRs and indirectly corroborates convergence to approximately optimal codeword ensembles.

2.4.4 SINR Based Codeword Update

For standard iterative interference avoidance algorithms [21, 28], codeword update by user k results in a lower TSC only if \mathbf{R}_k does not change during the codeword update. If more than one user updates codewords at the same time (synchronous update), then we can say little analytic



Figure 2.8: Variance of inverse SIR for 4 users in 12 dimensions. New user added after 10000 bit intervals. top plot: gradient descent with $\nu = 0.005$. bottom plot: lagged IA with $\alpha = 100, \beta = 1$.

about the resulting TSC. However in a truly uncoordinated and distributed implementation of IA, users will probably update codewords whenever SINR falls below some threshold, and multiple users might choose to update their codewords at the same time. Empirically we have observed that under such *SINR based update*, codeword steering with sufficiently small step sizes leads to lower TSC – though not necessarily the optimum value – for randomly chosen initial codewords. FIGURES 2.9, 2.10, 2.11 provide comparison between TSC convergence rates (in the form of average inverse SIR variation with time) for single user codeword update and for SINR based codeword update. Note that the above-mentioned figures also illustrate the issue of limited feedback channel capacity (to be discussed in the next section).

2.4.5 Feedback Channel Capacity Effects

All the prior experiments assume the availability of a perfect noiseless (infinite capacity) feedback channel. In this experiment, we consider a noisy (finite capacity) feedback channel and empirically estimate the minimum feedback capacity required for good system performance.

The capacity C_g of a channel with additive Gaussian noise is given by,

$$C_g = \frac{1}{2}log_2(1+SNR)$$
bits (2.50)

Since a vector $\mathbf{r}(n)$ is fed back, we can model the feedback channel as a set of L identical parallel channels, one for each dimension. Different noise levels result in different capacities (bits per dimension, or $\frac{b}{n}$ for short) for these channels. The SNR for each dimension of the vector feedback channel is given by,

$$SNR = \frac{\frac{1}{L}E[\mathbf{r}^{\top}\mathbf{r}]]}{N_f} = \frac{M}{LN_f}$$
(2.51)

where, N_f is the feedback channel noise power per dimension.

Figure 2.9 shows the behavior of an underloaded system (12 users in 24 dimensions) for different values of feedback capacity both for the case of single user update and SINR based update. We can see that increasing $\frac{b}{n}$ above 1 does not confer much advantage in convergence time. For the single user update case, convergence requires around 2000 bit intervals. In the SINR-based update scheme, each user updates its codeword if its SINR is below 7 dB i.e. 1 dB below the maximum 8 dB value which was chosen as the SNR for each user.² We can see that all users are able to achieve an SINR above 7 dB after ≈ 1000 bit intervals.

Figure 2.10 shows the corresponding plot for an overloaded system (16 users in 12 dimensions). Here also, we see that feedback of ≈ 1 bit per received vector dimension seems sufficient and convergence requires ≈ 3000 iterations in the case of single user codeword update. For obtaining the SIR based update plot, we set the minimum desired SINR threshold equal to 6 dB (SNR for each user was chosen as 20 dB which results in a maximum possible SINR = 7.2636 corresponding to a WBE codeword set choice). Although the average inverse SIR falls very sharply and reaches a steady state in less than 500 iterations, the steady state SINR for many users remains lower than 6 dB since the codewords don't seem to converge (the average inverse SIR does not stabilize) as they did in the underloaded case. This fact is verified by FIGURE 2.12 which shows the number of simultaneous user codeword updates as a function of time.

²SNR was chosen as 8 dB for comparisons with previous work considered in next section.



Figure 2.9: Effect of varying feedback channel capacity on system performance for 12 users in 24 dimensions. *SNR* of each user = 8 dB. The lagged IA scheme was used with parameters: $(\alpha = 100, \beta = 1)$. SINR Threshold = 7 dB. $\frac{b}{n}$ = bits per dimension.

FIGURE 2.11 shows the inverse SIR plots for higher number of users, keeping the users per dimension constant. The number of feedback bits required remains constant, however the convergence time for the single user case increases with the number of users as expected. With a constant steering step size, convergence time increases with the number of users since their codewords have to be steered sequentially.

SINR based codeword update again reduces the average inverse SIR rapidly but is unable to achieve the 6 dB SINR condition for many users. Figure 2.12 shows the number of users updating their codewords in each bit interval for SIR based update both for the underloaded (M = 12, L = 24) and overloaded (M = 16, L = 12; M = 32, L = 24) scenarios. We can conclude that in general if multiple users update their codewords simultaneously, the users won't be able to attain their desired SINR's. One simple heuristic by which users can avoid such semi-synchronous codeword updates is to wait for a random (geometrically distributed) number of bit intervals before updating whenever a need for update arises. The parameter controlling



Figure 2.10: Effect of varying capacity of the feedback channel on system performance for 16 users in 12 dimensions. SNR of each user = 20 dB. The lagged IA scheme was used with parameters: ($\alpha = 100, \beta = 1$). SINR Threshold = 6 dB. $\frac{b}{n}$ = bits per dimension.

the geometric distribution can be chosen to achieve a compromise between the convergence speed and steady state SINR value.

FIGURE 2.13 shows the average inverse SIR variation for an overloaded system (M = 16, L = 12) when the above mentioned heuristic is employed.

2.4.6 Prior Feedback Bounds

The problem of codeword optimization through the use of a limited bandwidth feedback channel has also been explored in [23, 25]. In [25], the receiver feeds back the optimal codeword for each user after quantizing it to B bits. Random vector quantization (RVQ) gives the upper bound on performance under a limited bandwidth constraint. [25] compares the performance of quantized reduced-rank codeword optimization with RVQ and for an underloaded system, shows the effect of varying B on SIR improvement for a single user. On the other hand, [23]



Figure 2.11: Effect of varying capacity of the feedback channel on system performance for 32 users in 24 dimensions. SNR of each user = 20 dB. The lagged IA scheme was used with parameters: ($\alpha = 100, \beta = 1$). SINR Threshold = 6 dB. $\frac{b}{n}$ = bits per dimension.

treats the issue of codeword quantization emperically and presents performance results for uniform and non-uniform quantization. In this section, we compare the feedback capacity requirement of our scheme with those of [23, 25].

For a system load M/L of 0.5, the reduced-rank codeword optimization scheme in [25] requires around 1 bit per user per dimension to improve the SINR to 7 dB, when the SNR of each user is 8 dB. Thus the total number of feedback bits required (N_B) for reduced-rank codeword optimization for the underloaded system we considered earlier (M = 12, L = 24) is equal to $M \times L = 288$ bits. Results from [23] indicate that 4 - 5 bits per codeword dimension per user are sufficient for near-optimal performance which gives a total feedback bit budget of $(N_B = 4 \times M \times L = 1152)$. For our scheme, N_B equals $T \times L \times C_g$ where, T is the number of bit intervals elapsed before the average SINR comes within 1 dB of the optimal value, C_g is the channel capacity per dimension in bits. FIGURE 2.9 shows the inverse SIR variation with time for the underloaded system under consideration. An easy calculation shows that the above



Figure 2.12: Variation of number of users simultaneously updating codewords in a bit interval vs time. The three plots correspond to the underloaded (M = 12, L = 24) and overloaded (M = 16, 32, L = 12, 24) systems considered earlier.

condition on average SINR is equivalent to saying that the average inverse SIR falls below 0.0486. From the figure, $T \approx 1000$ for the SIR based update scheme if we choose $C_g = 1.0$ which gives $N_B \approx 24000$ bits.

The large difference in the N_B values above stems from the fact that our scheme makes only *incremental changes* in user codewords so that the receiver can track them, whereas in [25], since the receiver already knows the optimal codewords, it can tolerate abrupt codeword changes. Also, the results in [25] were derived for the asymptotic case where both M and Ltend to ∞ with $\frac{M}{L} = 0.5$.

In addition, use of RLS filters instead of the LMS filtering scheme used here would reduce the number of required feedback bits. However, neither scheme, even in combination, is likely to match the feedback performance of centralized schemes like the reduced-rank codeword optimization [25].


Figure 2.13: M = 32, L = 24. If every user waits for a random number (geometrically distributed with parameter p = 1/8) of bit intervals before updating its codeword after the need arises, then the steady state SINR can be improved dramatically (compare with previous SINR based update for (M = 32, L = 24) when this heuristic was not used) with modest increase in convergence time. SNR of each user = 20 dB. The lagged IA scheme was used with parameters: ($\alpha = 100, \beta = 1$). SINR Threshold = 6 dB. b/n = bits per dimension.

2.5 Conclusion and Discussion

We have analyzed and simulated distributed interference avoidance (IA) based on covariance feedback broadcast from the receiver and incremental codeword changes by each user. The feedback could be a covariance matrix estimate from the receiver, or a sequence of received vectors $\mathbf{r}(n)$ to allow estimates to be constructed by each transmitter. The receiver tracks codeword changes by adapting the associated filters under a symbol error criterion. With perfect covariance feedback, the distributed method is equivalent in terms of codeword ensemble performance to centralized methods where codewords are computed by the receiver and distributed to transmitters. Lagged IA, however, shows a greater sensitivity than gradient descent IA to covariance uncertainty.

We considered matched, MMSE and decision feedback (DF) codeword filters and found

that the poor BER of the matched filter, especially as the number of active users was increased, militated against its use. MMSE and decision feedback filters had comparable performance with MMSE being better for lower SINR and DF better for higher SNR.

Our experiments with asynchronous (single user) codeword updates suggest that a global feedback (broadcast) of 1 bit per codeword dimension per bit interval is adequate for achieving near optimal IA performance. This figure, multiplied by the total number of iterations before convergence is declared, gives the total number of bits required for feedback and thereby a rough measure of feedback performance. In practice, instead of doing single user codeword updates, one can employ SINR based updates along with a geometric wait time heuristic (FIG-URE 2.13), and cut down on the number of iterations required for convergence.

In contrast, centralized schemes seem to require much less feedback capacity [25]. So perhaps distributed interference avoidance methods will be most useful in cases where centralized methods cannot be implemented. However, also we note that the machinery used for both codeword tracking and codeword update is strongly reminiscent of adaptive equalization for which a large body of work and hardware methods exist. Thus, one could imagine reaping the sum/user capacity increases of interference avoidance in a closed loop fashion with minimal development of new adaptive methods and hardware.

2.6 Appendix: proof of equation (2.26) (Eigenvalue/Eigenvector Update)

 $\mathbf{A}_1 \mathbf{u}_{i1} = \gamma_{i1} \mathbf{u}_{i1}$ can be rewritten as

$$\left(\mathbf{A}_{0} + \xi \mathbf{A}_{p}\right)\left(\mathbf{u}_{i0} + \xi \mathbf{u}_{ip}\right) = \left(\gamma_{i0} + \xi \gamma_{ip}\right)\left(\mathbf{u}_{i0} + \xi \mathbf{u}_{ip}\right)$$
(2.52)

Ignoring the $O(\xi^2)$ terms and comparing constants and coefficients of ξ on both sides we get

$$\mathbf{A}_0 \mathbf{u}_{i0} = \gamma_{i0} \mathbf{u}_{i0} \tag{2.53}$$

$$\mathbf{A}_0 \mathbf{u}_{i1} + \mathbf{A}_1 \mathbf{u}_{i0} = \gamma_{i0} \mathbf{u}_{i1} + \gamma_{i1} \mathbf{u}_{i0}$$
(2.54)

Equation (2.53) does not provide any new information. Note that $\{\mathbf{u}_{i0} : i = 1, ..., L\}$ is an orthonormal set because \mathbf{A}_0 is Hermitian. Multiplying equation (2.54) by \mathbf{u}_{j0}^{\top} for j = 1, ..., L on the left, and observing that $\mathbf{u}_{i0}^{\top}\mathbf{A}_0 = \gamma_{i0}\mathbf{u}_{i0}^{\top}$, we obtain

$$\gamma_{i1} = \mathbf{u}_{i0}^{\top} \mathbf{A}_1 \mathbf{u}_{i0} \tag{2.55}$$

$$(\gamma_{j0} - \gamma_{i0}) \mathbf{u}_{j0}^{\top} \mathbf{u}_{i1} = -\mathbf{u}_{j0}^{\top} \mathbf{A}_1 \mathbf{u}_{i0}, j \neq i$$
(2.56)

The eigenvalue/eigenvector update equations (2.26) can now be deduced from (2.55).

Chapter 3

A CDMA based Scheme for Transmission of Correlated Gaussian Symbols over a Gaussian Multiple Access Channel

Energy-efficient transmission of correlated data is one of the central problems in the design of wireless sensor networks. For example, in a typical sensor network [1], a group of sensors measure a common physical phenomenon and send their observations to a central repository. Since sensors are generally assumed to be deployed in very large numbers, measurements from spatially closer sensors will have a high degree of correlation. These sensor nodes usually have a non–replenishable source of energy, therefore it is highly desirable to keep the transmission powers at their minimum levels. Schemes that exploit the underlying correlation structure of the data in reducing the transmission power requirements are especially desirable. In this chapter, we consider a CDMA (code-division-multiple-access) based transmission scheme that exploits the correlation structure of sources by facilitating statistical cooperation among the sensors.

The information-theoretic capacity of a single cell symbol synchronous white Gaussian noise CDMA system was derived in [16] and it was shown that the capacity is a function of the correlations between transmitter signature waveforms. For an average power constraint on symbols of all transmitters, Massey *et al.* [17] showed that the capacity maximizing codewords for the single cell symbol synchronous system are same as the Welch Bound Equality (WBE) sequences. Viswanath *et al.* [18] generalized the result to the case where the transmitter power constraints are unequal. Further extensions include the colored noise case [19] and joint optimizations of codeword/power levels for fading channels [39].

In a CDMA system, the transmitter's symbols are assumed to be independent of each other and all the work above maintains this assumption. In this chapter, we assume a sensor network model where nodes observe correlated Gaussian symbols and make use of the correlation structure to design signature waveforms (codewords) to transmit their observations to a common receiver. We first formulate and analytically solve the problem of finding the optimal codewords and power levels where optimality is defined as minimization of the TMSE (total mean squared error) in the reproduced observations at the receiver under a total power constraint. Then, we consider the case of individual power constraints and give analytical solutions for certain special cases. Following this, we characterize the performance of our transmission scheme in the form of a Total power-TMSE tradeoff function. We compare the performance of our scheme with an information theoretic outer bound, and with another separation based scheme.

Throughout this chapter, the following notational guidelines will be used: Bold uppercase letters are used for matrices, bold lowercase letters for vectors and lowercase letters for scalars. Uppercase letters are also used to denote system constants e.g. M, the number of transmitters.

3.1 System Model

Let us assume that there are M sensor nodes (transmitters) and node i observes a time sequence of zero mean i.i.d. Gaussian symbols $\mathbf{b}_i^n = [b_i(1), b_i(2), \dots, b_i(n)]$. At time instant j, the M symbols observed by the sensors, $[b_1(j), b_2(j), \dots, b_M(j)]$ have a covariance matrix \mathbf{B} with unit diagonal elements, which is the same for all time instances j. At time instant n, encoder i produces an output $\mathbf{x}_i(n) = \mathbf{f}_i(\mathbf{b}_i^n) = (f_{i1}(\mathbf{b}_i^n), f_{i2}(\mathbf{b}_i^n), \dots, f_{iL}(\mathbf{b}_i^n))$. Assuming a finite dimensional signal space, the signature waveforms of transmitters can be described as Ldimensional vectors i.e., $\mathbf{x}_i(n) \in \mathbf{R}^L$ (Figure 3.1), and $f_{ij}(\mathbf{b}_i^n)$'s are deterministic zero mean functions. The average transmission power used by the i^{th} transmitter is given by

$$p_{i} = \lim_{n \to \infty} \frac{1}{n} \sum_{j=1}^{n} E\left[\|\mathbf{x}_{i}(j)\|^{2} \right]$$
(3.1)

The received signal at time instant n is given by

$$\mathbf{y}(n) = \sum_{i=1}^{M} \mathbf{x}_i(n) + \mathbf{z},$$
(3.2)

where \mathbf{z} is a zero-mean Gaussian noise vector with covariance matrix $\sigma^2 \mathbf{I}_L$. The decoder processes the received signal $\mathbf{y}^n = {\mathbf{y}(1), \mathbf{y}(2), \dots, \mathbf{y}(n)}$ to produce symbol estimates $\hat{\mathbf{b}}^n = {\{\hat{\mathbf{b}}_1^n, \hat{\mathbf{b}}_2^n, \dots, \hat{\mathbf{b}}_M^n\}}$ such that $\hat{\mathbf{b}}_i^n = \mathbf{g}_i(\mathbf{y}^n)$. The average distortion in the i^{th} symbol is given by

$$\mathbf{MSE}_{i} = \lim_{n \to \infty} \frac{1}{n} E\left[\left\| \widehat{\mathbf{b}}_{i}^{n} - \mathbf{b}_{i}^{n} \right\|^{2} \right]$$
(3.3)



Figure 3.1: General Problem Setup: M sensor nodes measure correlated Gaussian symbols, independently encode their symbols, and attempt to communicate them to a central receiver through a Vector Gaussian Multiple Access Channel.

Given a set of constraints on maximum values of p_i 's, finding out the values of MSE_i 's that can be achieved for the general encoder-decoder structure given above is an open problem. Reference [15] recently considered the scalar version of this problem (with individual power constraints on transmitters) and showed that uncoded transmission is optimal if the power budget is below a certain threshold. Another related open problem is the CEO problem [40] where the sensors measure an independent noisy copy of the same signal and try to estimate the value of this signal at the receiver. In this chapter, we impose a special structure on the encoders where each sensor node transmits symbols using unit–norm codewords of length L, i.e. $\mathbf{x}_i(n) = \sqrt{p_i}\mathbf{s}_i b_i(n)$ (Figure 3.1). Dropping the time index n, (3.2) becomes,

$$\mathbf{y} = \mathbf{S}\mathbf{P}^{\frac{1}{2}}\mathbf{b} + \mathbf{z},\tag{3.4}$$

where

- \mathbf{P} : diag (p_1, p_2, \ldots, p_M)
- p_i : transmit power of i^{th} transmitter
- $\mathbf{S}: L \times M$ matrix $[\mathbf{s}_1, \mathbf{s}_2, \dots, \mathbf{s}_M]$
- $\mathbf{s}_{i\;L} I_{4} imes 1$ unit norm signature codeword of i^{th} transmitter
- **b** : Gaussian symbol vector $\sim \mathcal{N}(0, \mathbf{B})$
- \mathbf{z} : zero-mean Gaussian noise with variance $\sigma^2 \mathbf{I}_L$

 $\mathbf{B} = \mathbf{E} \left[\mathbf{b} \mathbf{b}^{\top} \right]$ is defined as the symbol correlation matrix normalized to unit norm diagonal entries.



Figure 3.2: CDMA based transmission scheme: Each sensor encodes its symbol by multiplying it with a length L signature sequence of unit norm and scaling the encoder output to have power p_i . The optimal decoder at the receiver (in terms of minimizing TMSE) turns out be a linear MMSE filter.

Since the received vector is jointly Gaussian with the data symbols, the optimal decoder for each b_i for minimum MSE_i, turns out to be a linear receiver filter. Let c_i be the receiver filter corresponding to the i^{th} transmitter, then the symbol estimate is given by

$$\widehat{b}_i = \mathbf{y}^\top \mathbf{c}_i. \tag{3.5}$$

The mean square error (MSE) corresponding to the i^{th} transmitter is given by,

$$MSE_{i} = E\left[\left(\widehat{b_{i}} - b_{i}\right)^{2}\right],$$
(3.6)

which allows us to define total MSE as

$$\mathsf{TMSE} = \sum_{i=1}^{M} \mathsf{MSE}_i \tag{3.7a}$$

$$=\sum_{i=1}^{M} \mathbf{c}_{i}^{\top} \left(\mathbf{S} \mathbf{P}^{\frac{1}{2}} \mathbf{B} \mathbf{P}^{\frac{1}{2}} \mathbf{S}^{\top} + \sigma^{2} \mathbf{I}_{L} \right) \mathbf{c}_{i} + M - 2\sum_{i=1}^{M} \mathbf{c}_{i}^{\top} \mathbf{S} \mathbf{P}^{\frac{1}{2}} \mathbf{E} \left[\mathbf{b} b_{i} \right]$$
(3.7b)

= Trace
$$\left[\mathbf{C}^{\top} \mathbf{S} \mathbf{P}^{\frac{1}{2}} \mathbf{B} \mathbf{P}^{\frac{1}{2}} \mathbf{S}^{\top} \mathbf{C} + \sigma^{2} \mathbf{C}^{\top} \mathbf{C} - 2 \mathbf{C}^{\top} \mathbf{S} \mathbf{P}^{\frac{1}{2}} \mathbf{B} + \mathbf{I}_{M} \right].$$
 (3.7c)

Note that TMSE is the trace of the MSE matrix

$$\mathbf{E} = E[(\hat{\mathbf{b}} - \mathbf{b})(\hat{\mathbf{b}} - \mathbf{b})^{\top}]$$
(3.8a)

$$= \left(\mathbf{C}^{\top} \mathbf{S} \mathbf{P}^{\frac{1}{2}} \mathbf{B} \mathbf{P}^{\frac{1}{2}} \mathbf{S}^{\top} \mathbf{C} + \sigma^{2} \mathbf{C}^{\top} \mathbf{C} - 2 \mathbf{C}^{\top} \mathbf{S} \mathbf{P}^{\frac{1}{2}} \mathbf{B} + \mathbf{I}_{M} \right).$$
(3.8b)

3.2 Total power constraint: Optimal Transmitter Codewords, Power Levels and Receiver Structure

Since the network lifetime of a sensor network is inversely proportional to the battery powers of the sensors, it is crucial to design transmission schemes that conserve battery power. Here we consider a constraint on total transmission power used by the sensors. Physically, this problem might correspond to a scenario where we are interested in deploying identical sensors in dense groups such that all members in a group measure the same symbol value and use the same codewords for transmission. Each group of sensors could then be seen as a single sensor whose total transmission power would be equal to the aggregate power of sensors in the group. The problem of finding the optimal number of sensors in each group would then translate to a power allocation problem under a total power budget.

The optimization problem can be stated as follows,

$$\min_{\mathbf{S},\mathbf{P},\mathbf{C}} \mathsf{TMSE} \tag{3.9a}$$

s.t. diag
$$(\mathbf{S}^{\top}\mathbf{S}) = \mathbf{1}$$
, Trace $[\mathbf{P}] = \mathbf{P}_{tot}$. (3.9b)

It is well–known [41] that the structure of the optimum linear receiver that minimizes the MSE is the MMSE receiver. For this problem, the expression for the optimum receiver was obtained by setting $\partial (\text{TMSE}) / \partial \mathbf{C} \mid_{\mathbf{C}=\mathbf{C}^{\star}} = 0$. The solution is found to be

$$\mathbf{C}^{\star} = \left(\mathbf{S}\mathbf{P}^{\frac{1}{2}}\mathbf{B}\mathbf{P}^{\frac{1}{2}}\mathbf{S}^{\top} + \sigma^{2}\mathbf{I}_{L}\right)^{-1} \left(\mathbf{S}\mathbf{P}^{\frac{1}{2}}\mathbf{B}\right).$$
(3.10)

Substituting (3.10) in (3.7), the TMSE expression reduces to

TMSE =
$$M$$
 - Trace $\left[\mathbf{B} \mathbf{P}^{\frac{1}{2}} \mathbf{S}^{\top} \left(\sigma^{2} \mathbf{I}_{L} + \mathbf{S} \mathbf{P}^{\frac{1}{2}} \mathbf{B} \mathbf{P}^{\frac{1}{2}} \mathbf{S}^{\top} \right)^{-1} \mathbf{S} \mathbf{P}^{\frac{1}{2}} \mathbf{B} \right]$ (3.11a)

$$= M - \operatorname{Trace}\left[\frac{\mathbf{B}\mathbf{P}^{\frac{1}{2}}\mathbf{S}^{\top}}{\sigma^{2}}\left\{\mathbf{I}_{L} - \frac{\mathbf{S}\mathbf{P}^{\frac{1}{2}}\mathbf{B}\mathbf{P}^{\frac{1}{2}}\mathbf{S}^{\top}}{\sigma^{2}} + \left(\frac{\mathbf{S}\mathbf{P}^{\frac{1}{2}}\mathbf{B}\mathbf{P}^{\frac{1}{2}}\mathbf{S}^{\top}}{\sigma^{2}}\right)^{2} - \cdots\right\}\mathbf{S}\mathbf{P}^{\frac{1}{2}}\mathbf{B}\right]$$
(3.11b)

$$= \sigma^{2} \operatorname{Trace}\left[\left(\sigma^{2} \mathbf{B}^{-1} + \mathbf{P}^{\frac{1}{2}} \mathbf{S}^{\top} \mathbf{S} \mathbf{P}^{\frac{1}{2}} \right)^{-1} \right].$$
(3.11c)

Note that $\mathbf{SP}^{\frac{1}{2}}\mathbf{BP}^{\frac{1}{2}}\mathbf{S}^{\top}$ is positive definite, which implies that $\left(\mathbf{SP}^{\frac{1}{2}}\mathbf{BP}^{\frac{1}{2}}\mathbf{S}^{\top} + \sigma^{2}\mathbf{I}_{L}\right)$ is invertible. Also, it has been assumed in the above analysis that \mathbf{B}^{-1} exists. However, it will later be argued that invertibility of \mathbf{B} is not necessary since it does not affect the structure of the optimum codewords.

Now for finding the optimal codewords, assume that $L \leq M$. Later we show that for the L > M case the optimal codewords can be found out in a manner similar to the $L \leq M$ case. Let $\mathbf{B} = \mathbf{U}_1 \mathbf{\Sigma}_1 \mathbf{U}_1^{\top}$ and $\mathbf{A} = \mathbf{SP}^{\frac{1}{2}} = \mathbf{U}_2 \mathbf{\Sigma}_2 \mathbf{V}_2^{\top}$ where $\mathbf{\Sigma}_1 = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_M)$ such that, $\lambda_1 > \lambda_2 > \dots > \lambda_M$ and $\mathbf{\Sigma}_2 = [\text{diag}(\mu_1, \mu_2, \dots, \mu_L), \mathbf{0}_{L \times (M-L)}].$

Note that S and $P^{\frac{1}{2}}$ can be obtained from A as the normalized columns and norms of columns of A respectively.

Then the optimization problem (3.9) can be rewritten as,

$$\min_{\mathbf{A}\in\mathcal{R}^{L\times M}}\operatorname{Trace}\left[\left(\sigma^{2}\mathbf{B}^{-1}+\mathbf{A}^{\top}\mathbf{A}\right)^{-1}\right]$$
(3.12a)

s.t. tr(
$$\mathbf{A}^{\top}\mathbf{A}$$
) = $\sum_{j=1}^{L} \mu_j^2 = P_{\text{tot}}.$ (3.12b)

Lemma 1 If G and H are $n \times n$ Hermitian matrices, then

 $det(\mathbf{G} + \mathbf{H}) \leq \prod_{i=1}^{n} (\lambda_{[i]}(\mathbf{G}) + \lambda_{[n+1-i]}(\mathbf{H}))$ where $\{\lambda_{[i]}(\mathbf{G})\}$ is the ordered sequence of eigenvalues.

Proof: Marshall and Olkin [42, Lemma 9.G.4].

Lemma 2 Let \mathcal{A} be the set of all $L \times M$ matrices. For all $\mathbf{A} \in \mathcal{A}$, there exists $\widetilde{\mathbf{A}} \in \mathcal{A}$ such that $TMSE(\widetilde{\mathbf{A}}) \leq TMSE(\mathbf{A})$ and $\widetilde{\mathbf{A}}^{\top}\widetilde{\mathbf{A}}$ commutes with \mathbf{B} .

Proof: Define a function $\theta(\mathbf{A}) = \det (\sigma^2 \mathbf{B}^{-1} + \mathbf{A}^\top \mathbf{A})$. Choose $\mathbf{G} = \sigma^2 \mathbf{B}^{-1}$ and $\mathbf{H} = \mathbf{A}^\top \mathbf{A}$ following a similar argument as in [19]. Define $\widetilde{\mathbf{A}} = \mathbf{A}\mathbf{Q}$, where \mathbf{Q} is an orthogonal matrix chosen so that $\sigma^2 \mathbf{B}^{-1}$ and $\widetilde{\mathbf{A}}^\top \widetilde{\mathbf{A}}$ commute and the eigenvector corresponding to the i^{th} largest eigenvalue of $\sigma^2 \mathbf{B}^{-1}$ is the same as that corresponding to the $(n + 1 - i)^{th}$ largest eigenvalue of $\widetilde{\mathbf{A}}^\top \widetilde{\mathbf{A}}$. Note that $\widetilde{\mathbf{A}} \in \mathcal{A}$ since $\operatorname{tr}(\widetilde{\mathbf{A}}^\top \widetilde{\mathbf{A}}) = \operatorname{tr}(\mathbf{Q}^\top \mathbf{A}^\top \mathbf{A}\mathbf{Q}) = P_{\text{tot}}$. Now using Lemma 1, $\theta(\widetilde{\mathbf{A}}) \geq \theta(\mathbf{A})$. Since $\theta(\mathbf{A})$ is Schur–concave and TMSE is Schur–convex in the eigenvalues of $(\sigma^2 \mathbf{B}^{-1} + \mathbf{A}^\top \mathbf{A})$ (see Appendix 3.6.1), it follows that TMSE($\widetilde{\mathbf{A}}$) ≤ TMSE(\mathbf{A}). ■

Lemma 2, combined with the fact that two matrices commute if and only if they share the same eigenvectors [29], restricts the optimization space to that subset of \mathcal{A} for which the condition $\mathbf{V}_2 = \mathbf{U}_1$ holds. Note that the above condition is sufficient but not necessary. An alternate proof that doesn't use the concepts of majorization [42] (but is more involved) can be developed along the lines of [43].

Substituting $V_2 = U_1$ in (3.11a), the following two cases arise.

1. $M \ge L$

TMSE =
$$\sigma^2 \text{Trace} \left[\left(\sigma^2 \mathbf{U}_1 \boldsymbol{\Sigma}_1^{-1} \mathbf{U}_1^{\top} + \mathbf{U}_1 \boldsymbol{\Sigma}_2^{\top} \boldsymbol{\Sigma}_2 \mathbf{U}_1^{\top} \right)^{-1} \right]$$
 (3.13)

$$= \sigma^2 \sum_{i=1}^{L} \frac{1}{\frac{\sigma^2}{\lambda_i} + \mu_i^2} + \sum_{i=L+1}^{M} \lambda_i.$$
(3.14)

The Lagrangian corresponding to the optimization problem (3.12) can be written as follows,

$$\mathcal{L}\left(\mu_{1}^{2},\ldots,\mu_{L}^{2},\beta\right) = \text{TMSE} + \beta\left(\sum_{i=1}^{L}\mu_{i}^{2} - P_{\text{tot}}\right).$$

Using Kuhn-Tucker conditions [44], this leads to the following optimal solution,

$$\mu_i = \max\left(0, \alpha\right),\tag{3.15}$$

where α is chosen such that

$$\sum_{i=1}^{M} \mu_i^2 = P_{tot}.$$
(3.16)

2. M < L:

It can be verified that only the first $M \mu_i$ s need to be optimized, and the remaining (L - M) eigenvalues may be set to zero for obtaining the optimal solution.



Figure 3.3: Water-filling is achieved by distributing the sum of the eigenvalues of $\mathbf{A}^{\top}\mathbf{A}$ over the eigenvalues of \mathbf{B}^{-1} .

In other words, for any M and L, the optimal solution corresponds to water-filling (Fig. 1) the smallest $K = \min(L, M)$ eigenvalues of \mathbf{B}^{-1} with those of $\mathbf{A}^{\top}\mathbf{A}$, and aligning the eigenvectors of $\mathbf{A}^{\top}\mathbf{A}$ and \mathbf{B} as described in the proof of Lemma 2. Intuitively this corresponds to allocating power along directions carrying maximum information about \mathbf{B} .

The above analysis assumed that **B** is invertible. However, the result holds even for a noninvertible **B** since it can be made invertible by adding an infinitesimally small perturbation matrix (while ensuring that **B** is still a correlation matrix). As a result, previously non-zero eigenvalues of \mathbf{B}^{-1} will suffer very little change, while the other eigenvalues (previously zero) will now attain large finite values, but the corresponding dimensions will be avoided by the water-filling solution [45].

3.2.1 Constructing the Optimal Sequences

From Section 3.2, $\mathbf{A} = \mathbf{SP}^{\frac{1}{2}} = \mathbf{U}_2 \boldsymbol{\Sigma}_2 \mathbf{V}_2^{\top}$. Equation (3.15) gives the structure of $\boldsymbol{\Sigma}_2$ and Lemma 2 gives the structure of \mathbf{V}_2 . Note that any orthogonal matrix can be chosen for \mathbf{U}_2 and so there exists a whole class of signature sets. We illustrate the structure of optimal codewords

with the following examples:

Example: Consider the case of when symbols from different transmitters are uncorrelated and let L < M. Then $\mathbf{B} = \mathbf{I}_M$. The optimal codeword structure turns out to be,

$$\mathbf{V}_2 = \mathbf{I}_M \tag{3.17}$$

$$\boldsymbol{\Sigma}_{2} = \left[\operatorname{diag}\left(\sqrt{\frac{P_{\operatorname{tot}}}{L}}, \sqrt{\frac{P_{\operatorname{tot}}}{L}}, \cdots, \sqrt{\frac{P_{\operatorname{tot}}}{L}} \right), \boldsymbol{0}_{L \times (M-L)} \right].$$
(3.18)

Since U_2 can be chosen arbitrarily, the optimal codewords are given by,

$$\mathbf{A} = \left[\sqrt{\frac{P_{\text{tot}}}{L}} \mathbf{U}_{2}, \mathbf{0}_{L \times (M-L)} \right]$$
(3.19)

which means that TMSE can be minimized by letting only the first L transmitters transmit in orthogonal channels with equal power. However it can be easily verified that if the M transmitters are given equal power and the sequences are chosen as WBE sequences [17] i.e. $SS^T = \frac{M}{L}I_L$, even then the same minimum TMSE can be obtained. This illustrates the fact that the conditions given in this chapter for deriving optimal codewords are *sufficient but not necessary* and there can be other constructions of optimal codewords.

Example: Consider the case when **B** has identical elements and hence $rank(\mathbf{B}) = 1$. The optimal solution would correspond to all M transmitters using identical codewords with equal powers. An alternative approach based on separation principle would be to do source coding first and then transmitting the compressed data. This approach suggests that only one transmitter transmits with power equal to P_{tot} . It can be easily verified that to achieve the same distortion, our scheme would confer an M-fold power savings over the second scheme.

3.2.2 An Alternate Derivation of TMSE Minimizing Codewords/Power levels

Note that the TMSE expression depends only on the eigenvalues of the matrix $(\sigma^2 \mathbf{B}^{-1} + \mathbf{A}^\top \mathbf{A})$. TMSE is a Schur–convex function of the eigenvalues of $(\sigma^2 \mathbf{B}^{-1} + \mathbf{A}^\top \mathbf{A})$ (see Appendix 3.6.1). Minimizing TMSE is equivalent to maximizing det $(\sigma^2 \mathbf{B}^{-1} + \mathbf{A}^\top \mathbf{A})$, a Schur–concave function. Making use of the fact that det $(\mathbf{I} + \mathbf{A}\mathbf{B}) = \det(\mathbf{I} + \mathbf{B}\mathbf{A})$, and expressing $\sigma^{-2}\mathbf{B} = \mathbf{G}^\top \mathbf{G}$ and $\mathbf{X} = \mathbf{A}^\top \mathbf{A}$, we can write the following maximization problems that are equivalent to TMSE minimization.

$$\max_{\mathbf{A}} \frac{1}{2} \log \left[\det \left(\sigma^2 \mathbf{I}_L + \mathbf{A} \mathbf{B} \mathbf{A}^\top \right) \right] - \frac{L}{2} \log \sigma^2$$
(3.20a)

s.t. Trace
$$\begin{bmatrix} \mathbf{A}^{\top} \mathbf{A} \end{bmatrix} = P_{tot}$$
 (3.20b)

$$\mathbf{A} \ge 0. \tag{3.20c}$$

The second optimization is given by,

$$\max_{\mathbf{X}} \log \det \left(\mathbf{I}_M + \mathbf{G} \mathbf{X} \mathbf{G}^\top \right)$$
(3.21a)

s.t. Trace
$$[\mathbf{X}] = P_{tot}$$
 (3.21b)

$$X > 0.$$
 (3.21c)

The optimization problem (3.20) looks somewhat similar to the information theoretic sum capacity maximization problem for a white Gaussian synchronous CDMA system [16]. The information theoretic optimal way of sending correlated data over a multiple access channel is still unknown [46]. However for our scheme, if we aim to maximize the mutual information between the transmitted symbol vector and the received signal, then we end up with the first optimization problem given above.

The second optimization problem given above (3.21) is mathematically equivalent to the problem obtained when one tries to maximize the capacity of a point to point system with parallel channels having correlated Gaussian noise [47]. This problem can be solved analytically to obtain the optimal codeword and power levels for our original TMSE minimization problem. Recall that $\mathbf{B} = \mathbf{U}_1 \boldsymbol{\Sigma}_1 \mathbf{U}_1^{\mathsf{T}}$, therefore $\mathbf{G} = \sigma^{-1} \boldsymbol{\Sigma}_1^{1/2} \mathbf{U}_1^{\mathsf{T}}$. Let $\tilde{\mathbf{X}} = \mathbf{U}_1^{\mathsf{T}} \mathbf{X} \mathbf{U}_1$, then (3.21) can be re-written as

$$\max_{\tilde{\mathbf{X}}} \log \det \left(\mathbf{I}_M + \boldsymbol{\Sigma}_1^{1/2} \tilde{\mathbf{X}} \boldsymbol{\Sigma}_1^{1/2} \right)$$
(3.22a)

s.t. Trace
$$\left[\tilde{\mathbf{X}}\right] = P_{tot}$$
 (3.22b)

$$\tilde{\mathbf{X}} > 0. \tag{3.22c}$$

Using Hadamard's inequality [2, Theorem 16.8.2], one can see that choosing $\tilde{\mathbf{X}}$ to be a diagonal matrix is optimal. Rewriting the above problem in terms of diagonal elements of $\tilde{\mathbf{X}}$ and solving using a Lagrangian multiplier, we get the same water-filling solution obtained earlier. Also, it can be seen that left eigenvectors of \mathbf{A} (denoted by \mathbf{U}_2 earlier) can be chosen arbitrarily and an optimal value of \mathbf{V}_2 is \mathbf{U}_1 .

3.2.3 Comparison of Optimal Codewords with Random/Orthogonal Codewords

Figures 3.4 - 3.7 illustrate the reduction in TMSE achieved by using optimal codewords and power levels over random codewords and random power levels or orthogonal codewords with



Figure 3.4: M = L = 4, SNR = 10 dB

Figure 3.5: M = L = 4, SNR = 0 dB



Figure 3.6: M = L = 4, SNR = 10 dB

Figure 3.7: M = L = 4, SNR = 0 dB

equal power levels. In order to better understand the effect of increasing correlation on the reduction in TMSE, a special structure of B (given below) is used in Figures 3.4 and 3.5.

$$\mathbf{B} = \begin{bmatrix} 1 & \rho & \cdots & \rho \\ 0 & 1 & \cdots & \vdots \\ \vdots & \vdots & \ddots & \rho \\ \rho & \cdots & \rho & 1 \end{bmatrix}$$
(3.23)

In practice, the sensors will transmit at low powers to conserve battery and increase network lifetime. It can be observed that the reduction in TMSE is more pronounced at 0 dB than at 10 dB, thus emphasizing the importance of choosing codewords optimally.

Figures 3.6 and 3.7 repeat the above experiment with a general B, and again one can observe

3.3 Individual Power Constraints

In physical scenarios where we do not have a very dense deployment of sensors, it is often more meaningful to consider individual power constraints on sensors' transmissions when formulating the TMSE minimization problem. For simplicity, lets just consider the case when each sensor has an identical unit power constraint i.e. $\mathbf{P} = \mathbf{I}_M$. The optimization problem can then be written as,

$$\min_{\mathbf{S},\mathbf{C}} \mathsf{TMSE} \tag{3.24a}$$

s.t. diag
$$(\mathbf{S}^{\top}\mathbf{S}) = \mathbf{1}, \mathbf{P} = \mathbf{I}_{\mathbf{M}}.$$
 (3.24b)

The optimization of the receiver filters does not depend on the power constraints and the optimum receiver filter in this case is same as that obtained for the total power constraint case and is a linear MMSE filter. Substituting in the TMSE expression and simplifying gives

TMSE =
$$\sigma^2 \text{Trace}\left[\left(\sigma^2 \mathbf{B}^{-1} + \mathbf{S}^\top \mathbf{S}\right)^{-1}\right].$$
 (3.25)

We can convert the above problem into a standard convex optimization problem that can be efficiently solved using numerical techniques. Note that the TMSE expression depends only on the eigenvalues of the matrix $(\sigma^2 \mathbf{B}^{-1} + \mathbf{S}^\top \mathbf{S})$. Let $\lambda (\sigma^2 \mathbf{B}^{-1} + \mathbf{S}^\top \mathbf{S})$ be a set denoting the vector of eigenvalues of the above matrix under the constraints given in (3.24). Then TMSE is a Schur-convex function of the eigenvalues of $(\sigma^2 \mathbf{B}^{-1} + \mathbf{S}^\top \mathbf{S})$. Minimizing TMSE is same as maximizing the following Schur-concave function: det $(\sigma^2 \mathbf{B}^{-1} + \mathbf{S}^\top \mathbf{S})$. Let $\sigma^{-2} \mathbf{B} = \mathbf{G}^\top \mathbf{G}$ and $\mathbf{X} = \mathbf{S}^\top \mathbf{S}$, then the TMSE minimization problem under individual power constraints is equivalent to the following problem,

$$\min_{\mathbf{X}} -\log \det \left(\mathbf{I}_M + \mathbf{G} \mathbf{X} \mathbf{G}^\top \right)$$
(3.26a)

s.t. diag
$$(\mathbf{X}) = \mathbf{1}$$
 (3.26b)

$$\mathbf{X} > 0. \tag{3.26c}$$

The log det problem in (3.26) is a well studied convex optimization problem [47] with no known analytical solution. However, efficient numerical procedures based on interior point

methods [47] can be used to solve this problem. We shall now describe some special cases for which an analytical solution can be found.

The Lagrangian function corresponding to the optimization problem (3.26) can be written down as

$$\mathcal{L}(\boldsymbol{\Psi}, \lambda) = -\log \det \left(\mathbf{I}_M + \mathbf{G} \mathbf{X} \mathbf{G}^\top \right) - \lambda^\top \operatorname{diag}(\mathbf{X}) + \operatorname{Trace}\left[\mathbf{X} \boldsymbol{\Psi} \right]$$
(3.27)

where λ is a vector with non-negative components and Ψ is a positive semidefinite matrix. The associated Karush-Kuhn-Tucker (K.K.T.) conditions are given below

diag
$$(\lambda^{\star}) = \mathbf{G}^{\top} \left(\mathbf{I}_{M} + \mathbf{G} \mathbf{X}^{\star} \mathbf{G}^{\top} \right) \mathbf{G} + \mathbf{\Psi}^{\star}$$
 (3.28a)

$$\mathbf{X}^{\star} \ge 0, \operatorname{diag}\left(\mathbf{X}^{\star}\right) = \mathbf{1} \tag{3.28b}$$

$$\Psi^{\star} \ge 0, \lambda^{\star} \ge 0, \operatorname{tr} \left(\mathbf{X}^{\star} \Psi^{\star} \right) = 0.$$
(3.28c)

The above K.K.T. conditions don't give much insight into the analytical solution. However, for the special case when $\mathbf{X}^* > 0$ in the optimal solution, matrix $\boldsymbol{\Psi} = 0$, and the simplified K.K.T. conditions allow us to find an analytical solution as shown below

diag
$$(\lambda^{\star}) = \mathbf{G}^{\top} \left(\mathbf{I}_{M} + \mathbf{G} \mathbf{X}^{\star} \mathbf{G}^{\top} \right)^{-1} \mathbf{G}$$
 (3.29a)

$$\mathbf{X}^{\star} \ge 0, \text{diag}\left(\mathbf{X}^{\star}\right) = \mathbf{1} \tag{3.29b}$$

$$\lambda^{\star} \ge 0. \tag{3.29c}$$

From the above conditions, we can conclude that

diag
$$(\lambda^{\star}) = \left(\mathbf{G}\mathbf{G}^{\top} + \mathbf{X}^{\star}\right)^{-1}$$
 (3.30)

or diag
$$\left(\lambda^{\star^{-1}}\right) = \left(\sigma^2 \mathbf{B}^{-1} + \mathbf{X}^{\star}\right)$$
 (3.31)

or
$$\mathbf{X}^{\star}(i,j) = -\sigma^2 \mathbf{B}^{-1}(i,j)$$
 if $i \neq j$. (3.32)

3.3.1 Analytical solution for M = 2

Let $B = [1, \rho; \rho, 1]$ (W.L.O.G. assume $\rho \ge 0$) and $\mathbf{s}_1^\top \mathbf{s}_2 = \rho_s$. The TMSE expression can be written as,

$$TMSE == \frac{1}{\sigma^2} \operatorname{Trace} \left\{ \begin{bmatrix} \frac{1}{1-\rho^2} + \frac{1}{\sigma^2} & \frac{-\rho}{1-\rho^2} + \frac{\rho_s}{\sigma^2} \\ \frac{-\rho}{1-\rho^2} + \frac{\rho_s}{\sigma^2} & \frac{1}{1-\rho^2} + \frac{1}{\sigma^2} \end{bmatrix}^{-1} \right\}.$$
 (3.33)

Then we need to minimize the above expression w.r.t. ρ_s .

Consider the following result from the theory of majorization [42, Lemma 9.C.1.c]:-

Lemma 3 Consider the partitioned Hermitian matrix given below

$$\mathbf{H}_{\theta} = \begin{bmatrix} H_{11} & \theta H_{12} \\ \theta H_{21} & H_{22} \end{bmatrix}$$
(3.34)

then
$$\lambda(\mathbf{H}_{\theta_1}) \prec \lambda(\mathbf{H}_{\theta_2})$$
 (3.35)

for
$$0 \le \theta_1 \le \theta_2 \le 1$$
. (3.36)

where \prec is the majorization operator defined in Appendix 3.6.1. Since TMSE is a Schurconvex function of the eigenvalues of $(\sigma^2 \mathbf{B}^{-1} + \mathbf{S}^\top \mathbf{S})$, the above lemma can be used to find the optimal value of ρ_s . If $\rho \sigma^2 / (1 - \rho^2) \leq 1$, then $\rho_s^* = \rho \sigma^2 / (1 - \rho^2)$, else $\rho_s^* = 1$.

The minimum value of TMSE can be calculated as

$$TMSE^{\star} = \begin{cases} \frac{2\sigma^{2}(1-\rho^{2})}{\sigma^{2}+1-\rho^{2}} & \text{if } \frac{\rho\sigma^{2}}{1-\rho^{2}} \leq 1\\ \frac{2(\sigma^{2}+1-\rho^{2})}{2(1+\rho)+\sigma^{2}} & \text{otherwise.} \end{cases}$$
(3.37)

It turns out that for M = 2, TMSE^{*}, for the case of equal individual power constraints is the same as that obtained under a total power constraint.

3.4 The total power-TMSE trade-off function

The performance limits of any scheme for the general problem considered in this chapter can be charaterized in the form of a total power-TMSE tradeoff region that corresponds to the set of achievable ($P_{tot}, TMSE$) pairs for the problem. The total power-TMSE tradeoff function for a scheme captures the TMSE achieved by that scheme for a given total power budget. Similar tradeoff functions are considered in related work by Gastpar et al [48]. However, they consider a CEO type problem where the distortion (TMSE) is not measured directly in the observations but in the data embedded in the observations. Also, [48] primarily focusses on the scaling behavior (as number of sensors goes to infinity) of different transmission schemes while we are considering only finite number of sensors and exact comparison between the tradeoff functions in this section. Our encoding scheme can be interpreted as uncoded transmission scheme over the Gaussian vector channel with optimal power allocation over each channel and our results provide an achievable inner bound on total power-total distortion (TMSE) trade-off region for the above problem. This inner bound is stated explicitly in the next section.

3.4.1 Inner Bound: Uncoded transmission over parallel channels with optimal power allocation across channels

The achievable total power-TMSE trade-off for our scheme can be found by substituting the optimal value of $\mathbf{SP}^{\frac{1}{2}}$ in the TMSE expression and is given below,

$$\text{TMSE}_{1} = \sum_{i=k+1}^{M} \lambda_{i} + \frac{k^{2}}{SNR_{1} + \sum_{j=1}^{k} \frac{1}{\lambda_{j}}},$$
(3.38)

where $\{\lambda_1, \ldots, \lambda_M\}$ are the eigenvalues of **B** in decreasing order and k depends on P_{tot}/σ^2 in a manner given by the expression below

$$\sum_{i=1}^{k} \left(\frac{1}{\lambda_k} - \frac{1}{\lambda_i} \right) \le SNR_1 \le \sum_{i=1}^{k+1} \left(\frac{1}{\lambda_{k+1}} - \frac{1}{\lambda_i} \right).$$
(3.39)

Let $B = [1, \rho; \rho, 1]$ (W.L.O.G. assume $\rho \ge 0$), then the above expression reduces to

$$\text{TMSE}_{1} = \begin{cases} 1 - \rho + \frac{1 + \rho}{1 + (1 + \rho)SNR_{1}} & \text{if } SNR_{1} \le \frac{2\rho}{1 - \rho^{2}} \\ \frac{4(1 - \rho^{2})}{SNR_{1}(1 - \rho^{2}) + 2} & \text{otherwise.} \end{cases}$$
(3.40)

Note that the TMSE₁ expression given above for the M = 2 case suprisingly turns out to be the same as that obtained with individual power constraints of $P_{\text{tot}}/2$ at each encoder.

3.4.2 Outer Bound: Point to point system

An outer bound on the TMSE-Power tradeoff can be derived by considering a point to point system and applying the separation principle. The sensors jointly quantize their symbols and transmit the bits cooperatively over the Gaussian vector MAC channel. The outer bound can be derived by equating the rate-distortion function [2, Chapter 13] for the source coding problem with the capacity-cost function for the cooperative Gaussian vector MAC channel.

Rate-distortion for Correlated Gaussian Random Variables under the assumption of Joint Encoding

Lemma 4 Let $\mathbf{b} \sim \mathcal{N}(\mathbf{0}, \mathbf{B})$ be a $M \times 1$ Gaussian vector and let the distortion measure be $d(\mathbf{b}, \hat{\mathbf{b}}) = (\mathbf{b} - \hat{\mathbf{b}})^{\top} (\mathbf{b} - \hat{\mathbf{b}})$. Then the rate distortion function is given by

$$\begin{split} R(D) = &\sum_{i=1}^{M} \frac{1}{2} \log \frac{\lambda_i}{D_i} \\ D_i = \left\{ \begin{array}{l} \alpha, if\alpha < \lambda_i \\ \lambda_i, if\alpha \ge \lambda_i \end{array} \right\}. \end{split}$$

where, λ_i 's are the eigenvalues of **B** and α is choosen such that $\sum_{i=1}^{M} D_i = D$.

• **Proof:** See Cover and Thomas [2, Theorem 13.3.3]

Cooperative Gaussian Multiacess Vector Channel Capacity

Let us define the Cooperative Gaussian nultiaccess vector channel as follows

$$\mathbf{y} = \sum_{i=1}^{M} \mathbf{x}_{i} + \mathbf{z}$$
$$\mathbf{x}_{i} : L \times 1, E\left[\sum_{i=1}^{M} |\mathbf{x}_{i}|^{2}\right] \le P_{tot}$$
$$\mathbf{z} \sim \mathcal{N}(\mathbf{0}, \sigma^{2} I_{L})$$

Then we can state the following lemma

Lemma 5 The sum capacity of a Cooperative Gaussian multiaccess vector channel defined above is given by,

$$\sum_{i=1}^{M} R_i \le \frac{L}{2} \log \left(1 + \frac{MP_{tot}}{L\sigma^2} \right)$$

• **Proof:** See Appendix (3.6.2).

Combining the above two results, we get

$$\text{TMSE}_{2} = \sum_{i=r+1}^{L} \lambda_{i} + \frac{r(\prod_{i=1}^{r} \lambda_{i})^{1/r}}{(1 + \frac{MSNR_{2}}{L})^{L/r}}$$
(3.41)



Figure 3.8: Separation based Scheme: Each sensor first quantizes a sequence of length n of its symbols independently and then maps the quantized index to a Gaussian vector for transmission over the Multiple Access Channel.

where, r depends on $SNR_2 = P_{tot}/\sigma^2$ as given below:

$$\frac{L}{M} \left(\prod_{i=1}^{r} \left[\frac{\lambda_i}{\lambda_r} \right]^{1/L} - 1 \right) \le SNR_2 \le \frac{L}{M} \left(\prod_{i=1}^{r+1} \left[\frac{\lambda_i}{\lambda_{r+1}} \right]^{1/L} - 1 \right)$$
(3.42)

Example: Let $B = [1, \rho; \rho, 1]$ (W.L.O.G. assume $\rho \ge 0$), then the above expression reduces to

$$TMSE_{2} = \begin{cases} 1 - \rho + \frac{1+\rho}{(1+SNR_{2})^{2}} & \text{if } SNR_{1} \le \sqrt{\frac{1+\rho}{1-\rho}} - 1\\ \frac{2\sqrt{1-\rho^{2}}}{1+SNR_{2}} & \text{otherwise.} \end{cases}$$
(3.43)

3.4.3 A Separation based scheme for M = 2 case

In this section, we apply the separation principle to the M = 2 case with the constraint of having independent encoders at the sensors (Figure 3.8). We use a recent result by Wagner et al [49] on the rate-distortion function for a bivariate Gaussian source using two independent encoders and combine it with a bound on the capacity of the Gaussian vector channel with correlated inputs.

Rate-distortion for Correlated Gaussian Random Variables under the assumption of Independent Encoders

Wagner et al [49] recently solved the open problem of finding the rate-region for the quadratic Gaussian two-terminal source coding problem . In their paper [49], they characterized the complete achievable region $\{R_1(D_1, D_2), R_2(D_1, D_2)\}$, where R_1, R_2 are the rates at which two correlated Gaussian symbols (with covariance matrix $\mathbf{B} = [1, \rho; \rho, 1]$) must be encoded so

that they can be reproduced with mean squared distortions D_1, D_2 . For our purposes however, using the expression for optimal trade-off between $R = R_1 + R_2$ and $D = D_1 + D_2$ turns out to be sufficient. This point is elaborated in Appendix (3.6.3).

For the above **B**, if we substitute the optimal codewords and power levels for our scheme in the expression of MSE matrix (3.8) and evaluate the diagonal elements, we find that for all values of P_{tot} , $D_1 = D_2$. R(D) can be found by substituting $D_1 = D_2 = D/2$ in the (3.63) as elaborated in Appendix3.6.3. The result is

$$R(D) \ge \frac{1}{2}\log^{+}\left(\frac{2(1-\rho^{2})}{D^{2}}\left[1+\sqrt{1+\frac{D^{2}\rho^{2}}{(1-\rho^{2})^{2}}}\right]\right).$$
(3.44)

Bounding the Capacity of the Gaussian Multiple Access Channel with Correlated Inputs

It is shown in Appendix (3.6.4) that an upper bound on the capacity is given by

$$\frac{1}{2}\log\left(1+\frac{P_{\rm tot}(1+\rho)}{2\sigma^2}\right)^2.$$
(3.45)

Combining (3.45) and (3.44), we get the following outer bound on the achievable powerdistortion region,

$$SNR_{3} = \frac{2}{1+\rho} \left[\sqrt{\frac{2(1-\rho^{2})}{TMSE_{3}^{2}}} \left[1 + \sqrt{1 + \frac{TMSE_{3}^{2}\rho^{2}}{(1-\rho^{2})^{2}}} \right] - 1 \right].$$
 (3.46)

Solving for TMSE₃ in terms of SNR₃ we get

$$\text{TMSE}_{3} = \frac{2\sqrt{\left(1 + \frac{(1+\rho)}{2}\text{SNR}_{3}\right)^{2}(1-\rho^{2}) + \rho^{2}}}{\left(1 + \frac{(1+\rho)}{2}\text{SNR}_{3}\right)^{2}}$$
(3.47)

3.4.4 Performance Comparison

Low SNR case

Note that increasing SNR results in increasing values of k and r in equations (3.38) and (3.41). Consider the case when SNR is low enough to guarantee k = r = 1. To achieve the same distortion TMSE₁ = TMSE₂ = D, the SNR required by the CDMA based scheme and the SNR corresponding to the outer bound are related as:

$$1 + \lambda_1 \mathbf{SNR}_1 = (1 + \frac{M\mathbf{SNR}_2}{L})^L$$

Using the approximation, $(1 + x)^n \approx 1 + nx$ for small x, we get:

$$\frac{\mathrm{SNR}_2}{\mathrm{SNR}_1} \approx \frac{\lambda_1}{M} \tag{3.48}$$

It is easy to see that $\lambda_1 \in [1, M]$, therefore

$$\frac{\mathrm{SNR}_2}{\mathrm{SNR}_1} \in [1/M, 1]$$

At low SNRs, the total power-TMSE function for the separation based scheme (3.47) can be approximated as

$$TMSE_3 = 1 - \rho^2 + \frac{1 + \rho^2}{1 + (1 + \rho)\text{SNR}_3}$$
(3.49)

Comparing (3.49) with (3.40) reveals that at low SNRs, $SNR_1 \leq SNR_3$ for achieving the same TMSE.

To get some numerical intuition into the difference of performance of the above schemes, lets consider the following example: $\mathbf{B} = [1, 0.5; 0.5, 1]$. The maximum value of TMSE is 2 for this choice of **B**, and the minimum is 0. The eigenvalues of the **B** matrix are 1.5 and 0.5.

From (3.48), we can see that to achieve the same TMSE, the SNR required by our scheme is within 1.25 dB⁻¹ of that suggested by the outer bound. Using (3.49) and (3.43) one can see that the difference in SNR's required by the separation based scheme and our scheme is a function of SNR. To get some quantitative idea of the SNR gap between the schemes, lets arbritrarily assume a TMSE value of 1.9. Solving equations (3.49) and (3.43) gives $SNR_1 =$ 0.04761 and $SNR_3 = 0.05797$. Calculating the SNR gap in dB we find that our scheme performs 0.855 dB better than the separation based scheme. Though this is not a huge gain, one should remember that (3.49) just gives an upper bound on the performance of the separation based scheme and not the true performance. Therefore in reality, the gains might be higher. Also, a fair comparison between our scheme and the separation based scheme should take into account the system complexity in implementing these schemes. While the separation based

 $^{^{1}10\}log_{10}\left(\frac{2}{1.5}\right)$

scheme requires quantizers, sophisticated source and channel coders and decoders, our scheme is a essentially an un-coded transmission scheme and therefore should be much simpler to implement in practice.

High SNR case

Now consider the case when SNR is high enough to guarantee k = r = L. For achieving the same distortion $\text{TMSE}_1 = TMSE_2 = D$, SNR_1 and SNR_2 are related as:

$$SNR_2 = \frac{1}{M} (\prod_{i=1}^{L} \lambda_i)^{1/L} SNR_1 + \frac{1}{M} (\sum_{j=1}^{L} \frac{1}{\lambda_j}) (\prod_{i=1}^{L} \lambda_i)^{1/L} - \frac{L}{M}$$

Choosing high enough SNRs will ensure that:

$$\frac{SNR_2}{SNR_1} \approx \frac{1}{M} (\prod_{i=1}^L \lambda_i)^{1/L}$$

It can be easily shown that $(\prod_{i=1}^L \lambda_i)^{1/L} \in (0, \frac{M}{L}]$, therefore

$$\frac{SNR_2}{SNR_1} \in [0, 1/L]$$

At high SNRs, the total power-TMSE function for the separation based scheme (3.47) can be approximated as

$$TMSE_{3} = \frac{2\sqrt{1-\rho^{2}}}{1+\left(\frac{1+\rho}{2}\right)SNR_{3}}$$
(3.50)

By comparing equation (3.50) with equations (3.40) and (3.43), one can see that at high SNRs,

$$SNR_3 = \sqrt{\frac{1-\rho}{1+\rho}}SNR_1 = \frac{2}{1+\rho}SNR_2$$
 (3.51)

At high SNR, (3.51) shows that the SNR gap between our scheme and the separation based scheme is unbounded (as ρ tends to 1).

3.5 Conclusion and Discussion

This chapter proposed a CDMA based transmission scheme for the problem of transmitting correlated Gaussian symbols over a Gaussian multiple access vector channel. The signature sequences and receiver structure that minimize the total mean square error in the reproduced

symbols at the receiver, for a given total power budget at the transmitters, were then derived for this scheme. For the case of individual power constraints at the transmitters, the problem of finding the optimal signature sequences can be formulated as a convex optimization problem, which can be solved analytically only in certain special cases.

The performance of our scheme is then characterized in form of a total power-TMSE tradeoff function, and compared with an information-theoretic outer bound. Our CDMA based uncoded transmission scheme is found to be always suboptimal compared to the outer bound at all SNR's. This is in contrast to a recent result for the scalar version of the problem [15], where it is shown that uncoded transmission is optimal below a threshold SNR. A separation based scheme is also considered and it is observed that at low SNRs, our scheme's performance is slightly better than an upper bound on the performance of the separation based scheme. However, a fair comparison between the schemes should take into account the implementation complexity as well. While the separation based scheme requires quantizers, sophisticated source and channel coders and decoders, our scheme is a essentially an un-coded transmission scheme and therefore should be much simpler to implement in practice.

3.6 Appendix

3.6.1 Majorization: Definitions and Some Key Results

This section outlines certain mathematical relationships that are needed in obtaining the results of this chapter. A detailed survey of these inequalities and their properties may be found in [42]. A brief but comprehensive tutorial is provided in [18]. In this section we reproduce some of their definitions and results for convenience.

Definition 1 For any $\mathbf{x} = (x_1, x_2, \cdots x_n) \in \Re^n$, let

$$x_{[1]} \ge x_{[2]} \ge \dots \ge x_{[n]}$$

denote the components of \mathbf{x} in decreasing order, called the order statistics of \mathbf{x}

Definition 2 Let $\mathbf{x}, \mathbf{y} \in \Re^n$. Then, \mathbf{x} is majorized by \mathbf{y} (denoted by $\mathbf{x} \prec \mathbf{y}$) if

$$\sum_{i=1}^{k} x_{[i]} \leq \sum_{i=1}^{k} y_{[i]}, \quad k = 1, 2, \dots, n-1$$

and,
$$\sum_{i=1}^{n} x_{[i]} = \sum_{i=1}^{n} y_{[i]}.$$

Thus, majorization of x by y suggests that the components of x are "less spread out" or "more nearly equal" than the components of y.

An important example of majorization between two vectors is the following:

Example: For every $\mathbf{A} \in \Re^n$ such that $\sum_{i=1}^n a_i = 1$,

$$(a_1, a_2, \dots, a_n) \succ \left(\frac{1}{n}, \frac{1}{n}, \dots, \frac{1}{n}\right)$$

Definition 3 A real-valued function $\phi : \Re^n \to \Re$, defined on a set $\mathcal{A} \subset \Re^n$, is Schur-convex on \mathcal{A} if

$$\mathbf{x} \prec \mathbf{y} \text{ on } \mathcal{A} \Rightarrow \phi(\mathbf{x}) \leq \phi(\mathbf{y}).$$

The function ϕ is strictly Schur-convex if $\mathbf{x} \prec \mathbf{y}$ and $\mathbf{x} \neq \mathbf{y}$ implies that $\phi(\mathbf{x}) < \phi(\mathbf{y})$. Also, the function ϕ is Schur-concave if $-\phi$ is Schur-convex.

An important class of Schur-convex functions is the following:

Lemma 6 If $g : \Re \to \Re$ is convex, then the symmetric convex function $\phi(\mathbf{x}) = \sum_{i=1}^{n} g(x_i)$ is Schur-convex.

Proof: See [42].

Example: Trace (A^{-1}) is Schur-convex in the eigenvalues of matrix A.

Example: $\log \det \mathbf{A}$ is Schur-concave in the eigenvalues of matrix A.

3.6.2 Cooperative Gaussian Multiacess Vector Channel Capacity

The cooperative capacity of a Gaussian multiaccess vector channel is given by

$$\max_{p_{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_M}(.)} I(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_M; \mathbf{y})$$
(3.52)

under the constraint

$$E\left[\sum_{i=1}^{M} |\mathbf{x}_{i}|^{2}\right] \le P_{tot}$$
(3.53)

The steps below follow from standard information-theoretic arguments [2, Chapter 2, 10].

$$I(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_M; \mathbf{y}) \tag{3.54}$$

$$\leq h(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_M, \mathbf{y}) - h(\mathbf{y} | \mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_M)$$
(3.55)

$$=h(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_M, \mathbf{y}) - \frac{L}{2}\log\sigma^2$$
(3.56)

$$\leq \frac{1}{2} \log \det \left[E \left[(\mathbf{x}_1 + \ldots + \mathbf{x}_M) (\mathbf{x}_1 + \ldots + \mathbf{x}_M)^\top \right] + \sigma^2 \mathbf{I}_L \right] - \frac{L}{2} \log \sigma^2$$
(3.57)

(3.58)

To find the capacity, we need to maximize

$$\det \left[E\{ (\mathbf{x}_1 + \ldots + \mathbf{x}_M) (\mathbf{x}_1 + \ldots + \mathbf{x}_M)^\top \} + \sigma^2 \mathbf{I}_L \right]$$
(3.59)

under the constraint $E\left[\sum_{i=1}^{M} |\mathbf{x}_i|^2\right] \leq P_{tot}$. It is easy to verify that the determinant is maximized when $x_{ij} = \sqrt{P_{tot}/(ML)}v_j$ for all i, j, such that v_1, v_2, \ldots, v_L are zero mean i.i.d. Gaussian random variables with unit variances. The capacity expression given in (5) follows after substituting the maximizing x_{ij} 's.

3.6.3 Bivariate Gaussian Rate-Distortion with Independent Encoders

Reference [49] shows that the minimum rate pair (R_1, R_2) to independently encode the components of a bivariate Gaussian source (with covariance $B = [1, \rho; \rho, 1]$) with average distortions (D_1, D_2) is given by

$$R_1 \ge \frac{1}{2} \log^+ \left[\frac{1}{D_1} \left(1 - \rho^2 + \rho^2 2^{-2R_2} \right) \right]$$
(3.60)

$$R_2 \ge \frac{1}{2} \log^+ \left[\frac{1}{D_2} \left(1 - \rho^2 + \rho^2 2^{-2R_1} \right) \right]$$
(3.61)

$$R_1 + R_2 \ge \frac{1}{2} \log^+ \left[\frac{(1 - \rho^2)\beta(D_1, D_2)}{2D_1 D_2} \right]$$
(3.62)

where $\log^+(x) = \max(\log(x), 0)$ and

$$\beta(D_1, D_2) = 1 + \sqrt{\left(1 + \frac{4\rho^2 D_1 D_2}{(1 - \rho^2)^2}\right)}.$$
(3.63)

For a given total distortion D, we are interested in finding the sum rate $R(D) = (R_1(D) + R_2(D))$ required to acheieve that distortion. However, clearly the value of R(D) depends on how D is split into D_1 and D_2 . For our CDMA based scheme, if we substitute the optimal value of $\mathbf{SP}^{1/2}$ into the MSE matrix (3.8) and evaluate the diagonal elements for M = 2 case, we can show that for all values of P_{tot} , $D_1 = D_2$. For a given D and ρ , the value of R(D) can now be determined based on which of the three inequalities among (3.60), (3.61) or (3.62) is active. For $\mathbf{B} = [1, \rho; \rho, 1]$ and $D_1 = D_2$, one can argue based on symmetry that either (3.62) is active and/or both (3.60) and (3.61) are active simultaneously. Equating $R_1 = R_2$ in the first two inequalities and solving gives $R_1(D) + R_2(D) = \log^+ \left[\frac{(1-\rho^2) + \sqrt{(1-\rho^2)^2 + 2\rho^2 D}}{D} \right]$. Comparing with the R.H.S. of (3.62) gives

$$R(D) \ge \max\left[\frac{1}{2}\log^{+}\left[\frac{1}{D_{1}}\left(1-\rho^{2}+\rho^{2}2^{-2R_{2}}\right)\right],$$
(3.64)

$$\frac{1}{2}\log^{+}\left(\frac{2(1-\rho^2)}{D^2}\left[1+\sqrt{1+\frac{D^2\rho^2}{(1-\rho^2)^2}}\right]\right)\right]$$
(3.65)

We need to show that the second argument of the max(.,.) function in (3.64) always dominates the first term if $\rho \in [0,1]$ and $D \in [0,2]$. Straightforward algebra shows that this is equivalent to saying that $h_1(\rho, D) \ge h_2(\rho, D)$ for $\rho \in [0,1], D \in [0,2]$, where

$$h_1(\rho, D) = (1 - \rho^2) + \sqrt{(1 - \rho^2)^2 + \rho^2 D^2}$$
(3.66)

$$h_2(\rho, D) = (1 - \rho^2)^2 + \rho^2 D + (1 - \rho^2)\sqrt{(1 - \rho^2)^2 + 2\rho^2 D}$$
(3.67)

Lemma 7 For $\rho \in [0, 1]$ and $D \in [0, 2]$, $h_1(\rho, D) \ge h_2(\rho, D)$.

Proof: It is easy to see that for $\rho \in [0, 1]$, $h_1(\rho, D)$ and $h_2(\rho, D)$ are convex and concave in D respectively. Now consider the following function

$$h_3(\rho, D) = \frac{2(1-\rho)^2}{1+\rho^2} + \frac{2\rho^2}{1+\rho^2}D$$
(3.68)

For a fixed $\rho \in [0, 1]$, $h_3(\rho, D)$ is a line in D which is a common tangent to both $h_1(\rho, D)$ and $h_2(\rho, D)$ at point (2, 2). Now a convex/concave function always lies above/below its tangent, and the result follows.

3.6.4 Bounding the Capacity of the Gaussian Multiple Access Vector Channel with Correlated Inputs

The capacity in this case will be given by

$$\max_{p_{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_M}(.)} I(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_M; \mathbf{y})$$
(3.69)

under the constraints

$$E\left[\sum_{i=1}^{M} |\mathbf{x}_{i}|^{2}\right] \le P_{tot},\tag{3.70}$$

$$\mathbf{x}_i = \mathbf{f}_i(b_i^n), \forall i \in [0, M].$$
(3.71)

Proceeding in the same manner as the cooperative case in Appendix3.6.2, the problem boils down to maximizing the determinant given in (3.59), however with the additional constraint that \mathbf{x}_i 's are outputs of independent encoders ($\mathbf{x}_i = \mathbf{f}_i(b_i^n)$). This constraint puts a limit on the maximum correlation between any x_{ij} and x_{mn} . Proceeding as in [15] or [50], one can use Witsenhausen's lemma [51] and a well known result about the maximum correlation between functions of random variables [52] to derive the following bound:

$$E[x_{ij}x_{mn}] \leq E[b_ib_m], \qquad (3.72)$$

for all $i, j, m, n \in [1, M].$

Using Hadamard's inequality [2, Theorem 16.8.2], the determinant value for M = L = 2 can be bounded by $E[(x_{11} + x_{21})^2]E[(x_{12} + x_{22})^2]$, which can further be bounded by $[P_{tot}/2(1+\rho)]^2$ using (3.72).

Chapter 4

Random Access for Variable Rate Links

Design of efficient multiple access schemes has been an active area of research for more than four decades. Recently, advances in radio technology and spectrum policies have driven research to build interference aware systems like "cognitive radios" [53]. Earlier work [5] studies the role of "spectrum servers" as centralized schedulers in devising fair and efficient schedule for interfering links that are capable of varying their rates of transmission. Reference [54] investigates the role of the spectrum server to schedule end-to-end flows in a network of interfering links. The above mentioned schemes involved centralized scheduling that requires the scheduler to know complete global information about the links. The information could be all interference gains between each pair of links in the network. More often, the availability of such global information requires a lot of overhead processing by the central entity. Hence, perfect centralized scheduling schemes act as a benchmark for imperfect scheduling schemes [55] and decentralized or distributed multiple access schemes.

Distributed random access schemes, e.g., ALOHA have been widely used in practical multiple access systems. The CSMA/CA schemes used in the IEEE 802.11 networks are very popular, thanks to the ease of implementation and decentralized control of these random access techniques. Of late, research effort has been directed towards analyzing the performance of these random access schemes. Stability properties of random access schemes have been studied in [56, 57]. In [58, 59], the authors propose distributed approaches for fair random access. The throughput characteristics of random access schemes have been studied in in [60, 61]. A recent work [62] characterizes the Pareto boundary of the network throughput region as the family of solutions optimizing a weighted proportional fairness objective, parametrized by weights chosen by the links. The authors also propose a distributed random access scheme to achieve a desired point within the Pareto optimal boundary.

In this chapter, we consider a model in which links turn ON and OFF in each slot. The rate obtained in a link depends on the interference from other active links. We characterize

and compare the achievable throughput region of a centralized scheduling scheme with a probabilistic random access scheme. In the centralized scheduling scheme, the scheduler provides the fraction of time a set of links are on, in order to maximize an objective function. In the probabilistic random access scheme, each link turns ON or OFF with a fixed probability chosen independently of the other links in each slot. Section 4.2 defines the throughput region of both schemes. A natural question arises: *Are the set of rates that can be achieved in both cases the same*?. In section 4.3 we characterize the throughput regions of both schemes and identify conditions under which they are the same. We derive analytic expressions for the rate regions for a network with two links and provide an intuitive geometric explanation. In section 4.4 we propose a distributed algorithm in which each link updates its probability of transmission based on its current rate. This memoryless policy achieves any feasible point in the rate region of the probabilistic random access scheme. The distributed algorithm is shown to converge under a very loose PHY layer assumption on the devices that just requires that the rate achieved at any link goes down strictly with increased interference.

4.1 System Model

Consider a wireless network with N nodes forming L logical links sharing a common spectrum. The network can be represented as a directed graph $\mathcal{G}(\mathcal{V}, \mathcal{E})$, where the nodes in the network are represented by the set of vertices \mathcal{V} of the graph and the links are represented by a set of directed edges \mathcal{E} . Therefore, the cardinalities $|\mathcal{V}| = N$ and $|\mathcal{E}| = L$. A directed edge from a node m to node n implies that m wishes to communicate data to node n.

Define the set of *transmission modes* $T = \{0, 1, ..., M - 1\}$, where $M = 2^L$ denotes the number of possible transmission modes. Then the *mode activity vector* t_j of mode j is a binary vector, indicating the on-off activity of the links. If $t_j = (t_{1j}, t_{2j}, ..., t_{Lj})$ is a mode activity vector, then

$$t_{lj} = \begin{cases} 1, & \text{link } l \text{ is active under transmission mode } j, \\ 0, & \text{otherwise.} \end{cases}$$
(4.1)

Figure 4.1 shows a representative network and Figure 4.2 shows particular transmission mode for the set of links. Note that there are M possible transmission modes including the mode in which all links are off. Let $\mathbf{T} = [t_{lj}]$ be the transmission mode matrix. Similarly, we



Figure 4.1: Graph of network showing the nodes and directed links

can construct the $L \times M$ rate matrix $\mathbf{C}_L = [c_{lj}]$, where c_{lj} is the rate obtained by link l in mode j. By construction, $t_{lj} = 0 \Rightarrow c_{lj} = 0$. We impose an additional constraint on the entries of \mathbf{C}_L : any additional interference reduces the rate of an active link. In other words, if \mathcal{L}_j is the set of active links in mode j and $\mathcal{L}_{j'} \subset \mathcal{L}_j$ is the set of active links in mode $j \in \mathcal{L}_{j'} \cap \mathcal{L}_j$.

Many systems with interfering links can be modeled using the C_L described above, e.g., [5, 63]. The following two examples show the rate matrices for a network with two and three links respectively. For simplicity, in Sections 4.2 and 4.3, we will assume that each link gets a normalized rate of 1 unit, when it transmits in isolation. However, this assumption is not necessary in Section 4.4.

Example 1:
$$\mathbf{C}_2 = \begin{bmatrix} 0 & 1 & 0 & \alpha \\ 0 & 0 & 1 & \beta \end{bmatrix}$$
,
Example 2: $\mathbf{C}_3 = \begin{bmatrix} 0 & 1 & 0 & a & 0 & c & 0 & g \\ 0 & 0 & 1 & b & 0 & 0 & e & h \\ 0 & 0 & 0 & 0 & 1 & d & f & i \end{bmatrix}$

The conditions for C_2 are

$$\alpha, \ \beta < 1, \tag{4.2}$$



Figure 4.2: Graph of network showing transmission mode corresponding to (1 0 1 0)

and the conditions for \mathbf{C}_3

$$a, b, c, d, e, f < 1,$$
 (4.3)

$$g < a, c, \tag{4.4}$$

$$h < b, e, \tag{4.5}$$

$$i < d, f. \tag{4.6}$$

4.2 Rate Regions

We define the *rate region* as the set of rate vectors that can be achieved by a multiple access scheme. In this chapter, we compare the rate regions of a centralized scheduling scheme with a probabilistic random access scheme.

4.2.1 Centralized scheduling

In this scheme, a schedule is the specified by fractions of time each transmission mode is active. A centralized scheduler can be used to compute the the optimum time fractions of activity, to maximize a certain utility function [5]. Let x_j be the fraction of time that transmission mode j is active and r_l be the average data rate of link l. The average data rate in link l is the time average of the data rates of all the transmission modes that include link l. Thus,

$$r_l = \sum_j c_{lj} x_j, \tag{4.7}$$

or in vector form,

$$\mathbf{r} = \mathbf{C}_L \mathbf{x}.\tag{4.8}$$

Thus the rate region for the centralized scheduling scheme is given by

$$\mathcal{R}_L^S := \{ (r_1, \dots, r_L) : \mathbf{r} = \mathbf{C}_L \mathbf{x}, \mathbf{x} \ge 0, \mathbf{x}^T \mathbf{1} = 1 \}.$$
(4.9)

Clearly, the region \mathcal{R}_L^S is a polytope defined by its 2^L vertices which are given by the column vectors of \mathbf{C}_L .

4.2.2 Random Access Scheme

In this scheme, link l transmits with a probability p_l chosen independent of the other links in the network. The rate region for the random access scheme is given by

$$\mathcal{R}_L^P := \{ (r_1, \dots, r_L) : \mathbf{r} = \mathbf{C}_L \mathbf{x}, \mathbf{x} = \mathbf{f}(\mathbf{p}), \ \mathbf{0} \le \mathbf{p} \le \mathbf{1} \}$$
(4.10)

where $\mathbf{f}: \mathcal{R}^L \to \mathcal{R}^{2^L}$ is given by

$$\mathbf{f}(\mathbf{p}) = \begin{bmatrix} (1-p_1)(1-p_2)\dots(1-p_L) \\ p_1(1-p_2)\dots(1-p_L) \\ \vdots \\ (1-p_1)p_2\dots p_L \\ p_1\dots p_L \end{bmatrix}.$$
 (4.11)

It is easy to see that $\mathcal{R}_L^P \subseteq \mathcal{R}_L^S$. Also, since $\mathbf{f}(\cdot)$ is a continuous mapping, the set $\{\mathbf{x} : \mathbf{x} = \mathbf{f}(\mathbf{p}), 0 \leq \mathbf{p} \leq 1\}$ must be a closed and continuous region and therefore \mathcal{R}_L^P must also be closed and continuous. Our aim will be to characterize the Pareto boundary of \mathcal{R}_L^P and identifying the conditions, if any, under which $\mathcal{R}_L^P \equiv \mathcal{R}_L^S$. We first consider the following simple cases to obtain insight about the shape of the rate regions.

4.3 Characterization of \mathcal{R}_L^P

4.3.1 *L* = 2

Using (4.10) and definition of C_2 from Example 1, the rates on two links are

$$r_1 = p_1(1 - p_2) + \alpha p_1 p_2, \tag{4.12}$$

$$r_2 = (1 - p_1)p_2 + \beta p_1 p_2. \tag{4.13}$$

The above equations can be rewritten as

$$r_1 = p_2(p_1\alpha + (1 - p_1).0) + (1 - p_2)(p_1.1 + (1 - p_1).0),$$
(4.14)

$$r_2 = p_2(p_1\beta + (1-p_1).1) + (1-p_2)(p_1.0 + (1-p_1).0).$$
(4.15)

In vector form,

$$\begin{bmatrix} r_1 \\ r_2 \end{bmatrix} = p_2 \left(p_1 \begin{bmatrix} \alpha \\ \beta \end{bmatrix} + (1-p_1) \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right) + (1-p_2) \left(p_1 \begin{bmatrix} 1 \\ 0 \end{bmatrix} + (1-p_1) \begin{bmatrix} 0 \\ 0 \end{bmatrix} \right)$$
(4.16)

The above representation of the rate vector, as a nested convex combination of the polytope vertices, is useful in visualizing the rate region \mathcal{R}_2^P . We now consider two different cases.

Low Interference Case : $\alpha + \beta \ge 1$

Figure 4.3 shows \mathcal{R}_2^S . Any point in the quadrilateral OABC can be achieved using centralized scheduling. Notice that the vertices of the polytope OABC are the columns of \mathbf{C}_2 . For a given probability vector $\mathbf{p} = [p_1 \ p_2]^T$, the rate vector \mathbf{r} given by (4.16) is shown as point F in Figure 4.3. As p_1 varies between 0 and 1, points D and E completely trace the line segments AB and OC respectively. As p_2 varies between 0 and 1, the point F traverses the line segment ED completely. Hence, it can be seen that by varying \mathbf{p} , the achieved rate region \mathcal{R}_L^P is the same as \mathcal{R}_L^S .



Figure 4.3: \mathcal{R}_L^S and \mathcal{R}_L^P for the case $\alpha + \beta \ge 1$. $\mathcal{R}_L^P \equiv \mathcal{R}_L^S$ and is given by the area enclosed by OABC. B represents (α, β) .

Analytically, we can write (derivation given in the Appendix 4.7)

$$\mathcal{R}_{L}^{P} = \left\{ \begin{array}{cc} (r_{1}, r_{2}) : \\ 0 \leq r_{1} \leq \alpha \quad \Rightarrow \quad 0 \leq r_{2} \leq \frac{\alpha - (1 - \beta)r_{1}}{\alpha}, \\ \alpha \leq r_{1} \leq 1 \quad \Rightarrow \quad 0 \leq r_{2} \leq \frac{\beta(1 - r_{1})}{1 - \alpha}. \end{array} \right\}.$$

$$(4.17)$$

High Interference Case : $\alpha + \beta < 1$

In this case, \mathcal{R}_L^S is given by the triangle formed by points O, A and C in Figure 4.4. As in the previous case, point F in Figure 4.4 corresponds to the rate vector \mathbf{r} achieved for a given $\mathbf{p} = [p_1 \ p_2]^T$. If $p_1 = 1$, the line segment DE coincides with BC. As p_1 varies from 1 to 0, DE moves from BC to an intermediate position HG to finally AO (for $p_1 = 0$) tracing out the region \mathcal{R}_L^P as the area enclosed by OAHIC. Note that the boundary AHIC of the region is convex (verified from the analytical expression for \mathcal{R}_L^P in the appendix) and contains two linear components AH and IC. The presence of linear component AH can be geometrically understood by observing that as DE moves from HG to AO, endpoint D always lies on the linear segment AH. In order to intuitively understand the presence of IC, it helps to notice that as p_1 varies from 1 to 0, J, the point of intersection of DE and BC initially moves from B towards C, goes up to a certain point I, and then moves back towards B.

Note that we could also have expressed the rate equations as



Figure 4.4: \mathcal{R}_L^S and \mathcal{R}_L^P for the case: $\alpha + \beta < 1$. \mathcal{R}_L^S is given by the area enclosed by OAC and \mathcal{R}_L^P is given by the area enclosed by OAHIC. $B = (\alpha, \beta)$. Note that the dotted and dashed lines are just auxiliary constructions used for understanding the evolution of the concave curve AHIC (as elaborated in the previous paragraph).

$$\begin{bmatrix} r_1 \\ r_2 \end{bmatrix} = p_1 \left(p_2 \begin{bmatrix} \alpha \\ \beta \end{bmatrix} + (1 - p_2) \begin{bmatrix} 1 \\ 0 \end{bmatrix} \right) + (1 - p_1) \left(p_2 \begin{bmatrix} 0 \\ 1 \end{bmatrix} + (1 - p_2) \begin{bmatrix} 0 \\ 0 \end{bmatrix} \right)$$
(4.18)

The above equations give an alternate equivalent way of looking at the region, where now instead of lines AB and OC, we consider lines BC and AO.

The analytical characterization of the above region is given below (derivation given in the Appendix 4.7)

$$\mathcal{R}_{L}^{P} = \left\{ \begin{array}{ccc} (r_{1}, r_{2}) : \\ 0 \leq r_{1} \leq \frac{\alpha^{2}}{1-\beta} \Rightarrow & 0 \leq r_{2} \leq \frac{\alpha - (1-\beta)r_{1}}{\alpha}, \\ \frac{\alpha^{2}}{1-\beta} < r_{1} < 1-\beta \Rightarrow & 0 \leq r_{2} \leq \frac{(\sqrt{(1-\beta)r_{1}-1)^{2}}}{1-\alpha}, \\ 1-\beta \leq r_{1} \leq 1 \Rightarrow & 0 \leq r_{2} \leq \frac{\beta(1-r_{1})}{1-\alpha}. \end{array} \right\}.$$
(4.19)
4.3.2 *L* = 3

The analytical characterization of the rate region is cumbersome for the three dimensional case because of the number of sub-cases that need to be considered. However, the geometric intuition that we developed for the two link case can easily be extended to this case. Using the definition of C_3 , we can write the rate vector $\mathbf{r}(\mathbf{p})$ in the following form:

$$\begin{bmatrix} r_{1} \\ r_{2} \\ r_{3} \end{bmatrix} = p_{3} \left\{ p_{2} \left(p_{1} \begin{bmatrix} g \\ h \\ i \end{bmatrix} + (1 - p_{1}) \begin{bmatrix} 0 \\ e \\ f \end{bmatrix} \right) + (1 - p_{2}) \left(p_{1} \begin{bmatrix} c \\ 0 \\ d \end{bmatrix} + (1 - p_{1}) \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right) \right\} + (1 - p_{3}) \left\{ p_{2} \left(p_{1} \begin{bmatrix} a \\ b \\ 0 \end{bmatrix} + (1 - p_{1}) \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \right) + (1 - p_{2}) \left(p_{1} \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + (1 - p_{1}) \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \right) \right\}$$
(4.20)

Figure 4.5 illustrates the nested convex combination structure given above, where point N corresponds to the rate vector $\mathbf{r}(\mathbf{p})$. Working with our geometric intuition, we make the following claim (without proof):

Claim 1: $\mathcal{R}_L^P \equiv \mathcal{R}_L^S \Leftrightarrow$ following conditions are satisfied:

- $\bullet \ a+b \geq 1, c+d \geq 1, e+f \geq 1.$
- Points $\{(0, 0, 1), (c, 0, d), (g, h, i), (0, e, f)\}$ are coplanar.
- Points $\{(1,0,0), (a,b,0), (g,h,i), (c,0,d)\}$ are coplanar.
- Points $\{(0, 1, 0), (0, e, f), (g, h, i), (a, b, 0)\}$ are coplanar.



Figure 4.5: Visualizing the rate region for the L = 3 case

In this section, we present a distributed random access algorithm to achieve a feasible point in the rate region \mathcal{R}_L^P for a network with L links. Each link updates its probability of transmission based on the rate it achieves in the previous slot. We start by identifying a property of the function $r_i(\mathbf{p})$ (Lemma 8) that is the key for proving the convergence of our distributed algorithm.

The rate $r_i(\mathbf{p})$ achieved by link *i* in the random access scheme can be written as

$$r_i(\mathbf{p}) = \sum_{j=1}^{M} c_{ij} \prod_{l=1}^{L} \left[t_{lj} p_l + (1 - t_{lj})(1 - p_l) \right]$$
(4.21)

$$= p_i \sum_{j:t_{ij}=1} c_{ij} \prod_{l \neq i} \left[t_{lj} p_l + (1 - t_{lj})(1 - p_l) \right]$$
(4.22)

(4.23)

Let us define

$$g_i(\mathbf{p}_{-i}) = \sum_{j:t_{ij}=1} c_{ij} \prod_{l \neq i} \left[t_{lj} p_l + (1 - t_{lj})(1 - p_l) \right]$$
(4.24)

where

$$\mathbf{p}_{-i} = [p_1, \dots, p_{i-1}, p_{i+1}, \dots, p_L]^T$$
(4.25)

Then $r_i(\mathbf{p})$ can be written as

$$r_i(\mathbf{p}) = p_i g_i(\mathbf{p}_{-i}) \tag{4.26}$$

Lemma 8 $g_i(\cdot)$ is a positive and strictly decreasing function of p_j for all $j \neq i$. Therefore, $r_i(.)$ is a strictly increasing function of p_i and a strictly decreasing function of p_j for all $j \neq i$.

Proof See Appendix 4.8.

Now for each link *i*, consider the following iterative update of $p_i(n)$ based on the current rate $r_i(n)$ and the desired rate r_i^d . In practice the current rate $r_i(n)$ is measured by averaging the rates obtained over many slots.

$$p_i(n+1) = \frac{r_i^d}{r_i(n)} p_i(n)$$
(4.27)

Now we are ready to present the main result of this chapter that proves the convergence of our distributed algorithm given by iteration (4.27).

Theorem 2 Given a feasible rate vector $\mathbf{r}^d \in \mathcal{R}_L^P$, if all the links perform the above iteration independently starting with $\mathbf{p}(0) = \mathbf{0}$, then their iterations converge to a fixed point $(\mathbf{p}^*, \mathbf{r}^*)$ such that $\mathbf{r}^* = \mathbf{r}^d$ and $\mathbf{p}(n) \leq \mathbf{1}$ for all n.

Proof Using (4.26), we can rewrite (4.27) as

$$p_i(n+1) = \frac{r_i^d}{g_i(\mathbf{p}_{-i}(n))}$$
(4.28)

Substituting $\mathbf{p}(0) = \mathbf{0}$ in the iteration, we get $\mathbf{p}(1) = \mathbf{r}^d$ and therefore $\mathbf{p}(1) \ge \mathbf{p}(0)$. Using lemma 8 with the above fact, it follows that $\mathbf{p}(2) \ge \mathbf{p}(1)$ and in general $\mathbf{p}(n+1) \ge \mathbf{p}(n)$ for all *n*. Therefore, if $\mathbf{p}(n)$ is bounded from above by 1, as *n* increases, it must converge to a fixed point \mathbf{p}^* and the corresponding \mathbf{r}^* is then equal to \mathbf{r}^d .

Now we prove that if \mathbf{r}^d is feasible, then $\mathbf{p}(n)$ remains bounded below 1. Feasibility of \mathbf{r}^d means that there exists $\mathbf{0} \leq \mathbf{p}^d \leq \mathbf{1}$ such that

$$p_i^d = \frac{r_i^d}{g_i(\mathbf{p}_{-i}^d(n))} \tag{4.29}$$

By definition, $\mathbf{p}^d \ge \mathbf{p}(0)$. Using (4.28) and (4.29), we can see that $\mathbf{p}^d \ge \mathbf{p}(1)$ and in general $\mathbf{p}^d \ge \mathbf{p}(n)$ for all *n*. Therefore $\mathbf{p}(n)$ must also remain bounded below 1.

In case the users choose an infeasible r^d , the above iteration will lead to a situation where some $p_i(n)$'s exceed 1. To avoid such infeasible conditions, we can modify the iteration to the one given below.

$$p_i(n+1) = \min\left\{\frac{r_i^d}{r_i(n)}p_i(n), 1\right\}$$
(4.30)

The above iteration converges to the desired rate vector \mathbf{r}^d if \mathbf{r}^d is feasible.

In case the users start with an infeasible \mathbf{r}^d , we make some simple observations that are stated below as Lemma 9. Let \mathbf{r}^1 denote the rate vector corresponding to $\mathbf{p} = \mathbf{1}$ and $(\mathbf{p}^*, \mathbf{r}^*)$ denote the probability and rate vectors obtained when the above iteration converges for all *i*.

Lemma 9 Assume \mathbf{r}^d is infeasible. Then for any link *i*, the following conditions hold true: (a) $r_i^d > r_i^1 \Rightarrow r_i^* \ge r_i^1$, (b) $r_i^d \le r_i^1 \Rightarrow r_i^* = r_i^d$. (c) If L = 2, then $\mathbf{r}^d > \mathbf{r}^1 \Rightarrow \mathbf{r}^* = \mathbf{r}^1$. *Proof* (a) Clearly, $\mathbf{p}^* \leq \mathbf{1}$. If $p_i^* = 1$, then $r_i^* = 1.g_i(\mathbf{p}_{-i}^*) \geq 1.g_i(\mathbf{1}) = r_i^1$ (using Lemma 8). If $p_i^* < 1$, then using the convergence condition we get, $p_i^* = \min\{p_i^* r_i^d / r_i^*, 1\} = p_i^* r_i^d / r_i^*$ which gives $r_i^* = r_i^d > r_i^1$.

(b) We prove this by showing that for all n, $\min\{p_i(n)r_i^d/r_i(n), 1\} = p_i(n)r_i^d/r_i(n)$ which implies (using proof of Theorem 2) that $r_i^* = r_i^d$. To see that $p_i(n)r_i^d/r_i(n) \le 1$ for all n, consider the following chain of inequalities using Lemma 8:

$$\frac{r_i^d}{r_i(n)}p_i(n) = \frac{r_i^d}{g_i(\mathbf{p}_{-i}(n))} \le \frac{r_i^d}{g_i(\mathbf{1})} \le \frac{r_i^1}{g_i(\mathbf{1})} = 1$$
(4.31)

(c) We want to show that if \mathbf{r}^d is infeasible and $\mathbf{r}^d > \mathbf{r}^1$, then $p_1^* = p_2^* = 1$. We prove this by contradiction. It is not possible that both $p_1^* < 1$ and $p_2^* < 1$ because this would then imply that \mathbf{r}^d was feasible. Without loss of generality assume that $p_1^* < 1$ and $p_2^* = 1$. Then we can write $r_1^d = r_1(p_1^*, 1) \le r_1(1, 1) = r_1^1 < r_1^d$ (using Lemma 8) which is a contradiction.

4.5 Conclusion and Discussion

In this chapter, we compared the achievable throughput region of a probabilistic transmission scheme with that of centralized scheduling. We also presented a distributed algorithm to achieve any feasible rate vector in the throughput region of the probabilistic transmission scheme and proved its convergence. The algorithm is guaranteed to converge for any underlying PHY layer that ensures that rate on a link goes down strictly as the interference increases. Also, our distributed algorithm does not require any level of coordination or information sharing between the links. The only information exchange happens between the receiver of a link and the corresponding transmitter, where the receiver computes the instantaneous link throughput and feeds it back to the transmitter. The above properties of the algorithm make it attractive for adoption in unlicensed band scenarios where centralized control is difficult to implement. Also, since our algorithm allows the participating links to achieve any point in the feasible rate region, it could be used for guarantying different QoS levels for different users. This feature of our algorithm could be seen as a distinct advantage over conventional 802.11 based multiaccess schemes that don't have any such provision for guarantying differing QoS levels across users.

4.6 Appendix

4.7 Derivation of the pareto-boundary of the rate region for L = 2 case

Let $\gamma_1 = 1 - \alpha$ and $\gamma_2 = 1 - \beta$. Then we can write (4.12) and (4.13) as,

$$r_1 = p_1(1 - \gamma_1 p_2), \tag{4.32}$$

$$r_2 = p_2(1 - \gamma_2 p_1). \tag{4.33}$$

The Pareto boundary of the rate region can be defined as the set of rate pairs (r_1, r_2) such that at least one of them is non-zero and none of them can be increased without decreasing the other component. This boundary can be obtained by maximizing r_2 (or r_1) for each value of r_1 (or r_2). The constraints are $0 \le p_1, p_2 \le 1$. If we substitute $p_1 = r_1/(1 - \gamma_1 p_2)$ in the expression for r_2 , then constraints $0 \le p_1 \le 1$ imply that

$$0 \le p_2 \le \min\left\{\frac{1-r_1}{\gamma_1}, 1\right\}.$$

Now, for a given $r_1 \in [0, 1]$, we can find r_2 that lies on the Pareto boundary by solving the following optimization problem:

$$r_{2} = \max_{p_{2}} \left(p_{2} - \frac{\gamma_{2}p_{2}r_{1}}{1 - \gamma_{1}p_{2}} \right)$$

subject to
$$0 \le p_{2} \le \min \left\{ \frac{1 - r_{1}}{\gamma_{1}}, 1 \right\}.$$
 (4.34)

Since $r_2 = 0$ at $p_2 = 0$, the maximum occurs either at the boundary point $p_2 = \min\{(1 - r_1)/\gamma_1, 1\}$ or at a point where the derivative of the above function w.r.t. p_2 is zero. Setting the derivative $r'_2(p_2) = 0$ gives $p_2 = (1 \pm \sqrt{\gamma_2 r_1})/\gamma_1$. One of these values is greater than 1 and can be discarded. For the other value of p_2 to be valid, we need

$$0 \leq \frac{1 - \sqrt{\gamma_2 r_1}}{\gamma_1} \leq \min\left\{\frac{1 - r_1}{\gamma_1}, 1\right\},\,$$

which is satisfied only if $(1 - \gamma_1)^2 / \gamma_2 \leq r_1 \leq \gamma_2$. Rewriting in terms of α and β ,

$$\frac{\alpha^2}{1-\beta} \le r_1 \le 1-\beta. \tag{4.35}$$

We now consider the following two cases:

4.7.1 Low Interference Case : $\alpha + \beta \ge 1$

In this case, the maximum value of r_2 always occurs at the boundary point $p_2 = \min\{(1 - r_1)/(1 - \alpha), 1\}$. The optimal value of p_2 that maximizes r_2 is either $(1 - r_1)/(1 - \alpha)$ or 1 depending on whether $r_1 \ge \alpha$ or not. Substituting the maximizing value of p_2 in (4.34), we obtain the rate region given by (4.17).

4.7.2 High Interference Case : $\alpha + \beta < 1$

In this case, we have to consider separately the following ranges of r_1 :

$$0 \leq r_{1} \leq \frac{\alpha^{2}}{1-\beta},$$

$$\frac{\alpha^{2}}{1-\beta} < r_{1} < 1-\beta,$$

$$1-\beta \leq r_{1} \leq 1.$$
(4.36)

The optimal values of p_2 corresponding to the above three ranges of r_1 are

$$p_2 = \left[1, \frac{1 - \sqrt{(1 - \beta)r_1}}{1 - \alpha}, \frac{1 - r_1}{1 - \alpha}\right]$$
(4.37)

Substituting in (4.34), we obtain the rate region given by (4.19).

4.8 Proof of lemma 8

Positivity of $g_i(.)$ is evident from its definition. We must now show that $g_i(.)$ is a strictly decreasing function of p_k for all $k \neq i$. Computing the partial derivative of g_i w.r.t. p_k we get

$$\frac{\partial g_i}{\partial p_k} = \sum_{j:t_{ij}=1} c_{ij} (2t_{kj} - 1) \prod_{l \neq i, l \neq k} \left[t_{lj} p_l + (1 - t_{lj})(1 - p_l) \right]$$
(4.38)

In the above expression, the index j counts all the transmission modes in which $t_{ij} = 1$. Lets denote this set by $\mathcal{T}_{i=1}$. The set $\mathcal{T}_{i=1}$ can be partitioned into two disjoint sets $\mathcal{T}_{i=1,k=1}$ and $\mathcal{T}_{i=1,k=0}$ depending on whether the link k is active or not. Then for each $j \in \mathcal{T}_{i=1,k=1}$, there exists a unique $j' \in \mathcal{T}_{i=1,k=0}$ such that $t_{lj} = t_{lj'}$ for all $l \neq k$. Now noting that $j \in \mathcal{T}_{i=1,k=1}$ implies by definition that $(2t_{kj} - 1) = 1$ and similarly $j' \in \mathcal{T}_{i=1,k=0}$ implies that $(2t_{kj'} - 1) = -1$, we can write [4.38] as

$$\frac{\partial g_i}{\partial p_k} = \sum_{j \in \mathcal{T}_{i=1,k=1}} (c_{ij} - c_{ij'}) \prod_{l \neq il \neq k} [t_{lj} p_l + (1 - t_{lj})(1 - p_l)]$$
(4.39)

Using the properties of C_L , we know that $c_{ij} < c_{ij'}$ for all $j \in \mathcal{T}_{i=1,k=1}$. Therefore the above expression is negative and the result follows.

Chapter 5

Greedy Users and Resource Allocation Advisory Services

In the previous chapters we considered a variety of resource allocation scenarios from fully distributed independent users in chapter 2 to users with correlated information in chapter 3 to distributed but not particularly greedy in chapter 4.

The work in chapter 4 in particular suggests a new twist to the greedy/strategic distributed resource allocation problem. That is, the resource allocation process of chapter 4 could be carried out by a centralized mediating authority called the *spectrum server* [5, 6, 13, 54], that receives local interference reports from the links, and instructs them on spectrum usage. The actual performance figures and how users affect one another is considered public (and true) information.

However, in a distributed system aided by a spectrum server but without a provision for "spectrum police" – a facility that measures mutual interference in an objective fashion and reports it to the spectrum server – there could be motivation for users to lie about the levels of interference they experience. For example, how much user A interferes with user B is something that can only be measured by user B, it is not clear why user B will or will not report a higher or lower interference value if there is some personal some benefit to doing so. Thus, we could not resist at least formulating the problem and offering some preliminary analysis on when greedy users might choose to report interference conditions truthfully or lie through their teeth as part of a strategy to maximize their utilities.

5.1 System Model and Problem Statement

Consider a directed interference graph G(V, E) formed by a set of M interfering links (sourcedestination pairs). $V = \{1, 2, ..., M\}$ is the set of vertices, denoting the links, and E denotes the set of edges among the vertices (|E| = M(M - 1)). Existence of an edge from vertex ito vertex j indicates that transmission on link i renders any simultaneous transmission on link j unsuccessful. Let $e_{ij} = 0/1$ be a binary variable indicating the presence or absence of edge from link *i* to link *j*. A set of links that can be simultaneously operated in a single time slot (or channel) without causing interference between any two links constitute an independent set (denoted by $I = \{i_1, i_2, \dots, i_{|I|}\}$). For a given graph G(V, E), a *maximal* independent set *I* is a set such that its cardinality |I| is as large as possible.

5.1.1 Server Optimization Problem

Lets assume that the graph topology G(V, E) is known at the server, and consider the following optimization function at the server

$$\max_{\mathbf{x}} \sum_{i=1}^{M} w_i(x_i) \tag{5.1}$$

where $w_i()$'s are concave, increasing functions, and x_i is the time fraction for which link *i* is ON. The feasible set of time fractions, **x**, can be characterized in terms of the time fractions for which different independent sets are scheduled to be ON. Let $\{I_1, \ldots, I_L\}$ denote the collection of all independent sets, and $\mathbf{y} = [y_1, \ldots, y_L]$ be the corresponding time fractions for which they are scheduled, then we can write the following constraints for the above optimization problem.

$$\mathbf{x} = \mathbf{A}\mathbf{y} \tag{5.2}$$

$$\mathbf{y}^{\top}\mathbf{1} = 1, \mathbf{y} \ge 0 \tag{5.3}$$

where **A** is an $M \times L$ binary matrix such that $\mathbf{A}(i, j) = 1$ if link $i \in I_j$, else $\mathbf{A}(i, j) = 0$.

Characterization of the Optimal Solution

The optimization problem given by (5.1) is a convex optimization problem and its optimal solution can be found by first writing the corresponding Lagrangian function and then looking at the K.K.T. conditions for optimality. Let us first rewrite $\mathbf{A} = [\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_M]^{\top}$ so that \mathbf{a}_i is the i^{th} row of matrix \mathbf{A} . The time fraction for link i can now be written as, $x_i = \mathbf{a}_i^{\top} \mathbf{y}$. The Lagrangian function corresponding to problem (5.1) is then given by

$$L(\mathbf{y},\lambda,\mu) = \sum_{i=1}^{M} w_i \left(\mathbf{a}_i^{\top} \mathbf{y}\right) + \lambda \left(\mathbf{y}^{\top} \mathbf{1} - 1\right) + \mu^{\top} \mathbf{y}$$
(5.4)

The K.K.T. conditions at the optimal point $(\mathbf{y}^*, \lambda^*, \mu^*)$ can be found by taking derivatives of the Lagrangian function w.r.t. \mathbf{y}, λ, μ

$$\sum_{i \in \mathbf{I}_j} \left[\frac{\partial w_i \left(\mathbf{a}_i^\top \mathbf{y}^* \right)}{\partial y_j} \right] + \lambda^* + \mu_j^* = 0, \forall j$$
(5.5)

$$\mu_j^* y_j^* = 0, \forall j \tag{5.6}$$

$$\lambda^* \left[(\mathbf{y}^*)^\top \mathbf{1} - 1 \right] = 0 \tag{5.7}$$

The above condition states that, for any independent set I_j , either $y_j^* = 0$, i.e., no time fraction is alloted to independent set I_j , or

$$\sum_{i \in \mathbf{I}_j} \left[\frac{\partial w_i \left(\mathbf{a}_i^\top \mathbf{y}^* \right)}{\partial y_j} \right] = -\lambda^*$$
(5.8)

i.e., the sum of derivatives of weight functions corresponding to all links $i \in \mathbf{I}_j$ w.r.t. the time fraction y_j^* alloted to mode j is a constant. This makes intuitive sense once one realizes that the L.H.S. of equation (5.8) when multiplied by Δy_j corresponds to the change in server's objective function value if the allocation y_j^* is increased by a small value Δy_j . Now, for any deviation $\Delta \mathbf{y} = [\Delta y_1, \Delta y_2, \dots, \Delta y_L]$ from \mathbf{y}^* s.t. $\Delta \mathbf{y}^\top \mathbf{1} = 0$, the value of the server optimization function should not change. This leads to the above optimality conditions.

5.1.2 User Optimization Problem

We assume that each link is aware of the presence of the other (M - 1) links and can figure out which all links are interfering with it, i.e., link *i* can determine the length (M - 1) binary neighbor list, $\mathbf{e}_i = \{e_{1i}, e_{2i}, \dots, e_{(i-1)i}, e_{(i+1)i}, \dots, e_{Mi}\}$. Let $\hat{\mathbf{e}}_i$ denote the neighbor list reported by link *i* to the spectrum server. Based on the reports sent by the links, the server constructs a graph topology $\hat{G}(V, \hat{E})$, finds out the maximal independent sets $\{\hat{I}_1, \dots, \hat{I}_L\}$, solves some global optimization problem, and conveys the optimal time fractions and schedule to the links. Lets assume that each link derives zero utility when it is OFF or is being interfered by some other link, and a positive utility when it is ON and not being interfered by anyone.

Lets partition and rearrange the independent sets as

$$\{\hat{I}_1, \dots, \hat{I}_L\} = \{\hat{I}_1, \dots, \hat{I}_l\} \cup \{\hat{I}_{l+1}, \dots, \hat{I}_m\} \cup \{\hat{I}_{m+1}, \dots, \hat{I}_L\}$$
(5.9)

$$\left\{\begin{array}{ll}
i \in \widehat{I}_{k}, \text{ and } e_{ji} = 0, \forall j \in \widehat{I}_{k} \quad k \leq l \\
i \in \widehat{I}_{k}, \text{ and } \exists j \in \widehat{I}_{k} \text{ s.t. } e_{ji} = 1 \quad l < k \leq m \\
i \notin \widehat{I}_{k}, \quad k > m
\end{array}\right\}$$
(5.10)

Let (x^*, y^*) be the solution of (5.1), then link *i*'s utility is given by

$$u_i = \sum_{j=1}^{l} y_j^*$$
(5.11)

It can be seen from above that link *i*'s utility depends on y^* which is computed by the spectrum server based on the reported graph topology and its optimization objective. The only way link *i* can influence y^* is through the choice of \hat{e}_i . The main question explored in this chapter is given below.

Truthful Reporting Problem Statement

Q: What is the optimal reporting strategy, $\hat{\mathbf{e}}_i$, for link *i* for the server (5.1) and user (5.11) objective functions given above, assuming that link *i* has access to its own incoming edge vector \mathbf{e}_i and the reports sent by other links, $\hat{\mathbf{e}}_{V_{-i}} = (\hat{\mathbf{e}}_1, \dots, \hat{\mathbf{e}}_{i-1}, \hat{\mathbf{e}}_{i+1}, \dots, \hat{\mathbf{e}}_M)$? Are there some conditions under which $\hat{\mathbf{e}}_i = \mathbf{e}_i$?

5.2 Optimal Reporting Strategy for Some Special Server Objective Functions

5.2.1 Scheduling for maximizing weighted sum of time allocations of users

Lets consider a special case when the weight functions given in (5.1) are linear, i.e.,

$$w_i(x_i) = w_i x_i, \forall i \tag{5.12}$$

Let $\sum_{i \in \mathbf{I}_j} (w_i)$ be the total weight associated with an independent set \mathbf{I}_j . Lets assume for simplicity that each independent set has a unique value of total weight. Then condition (5.5) implies that the optimal solution to the server optimization problem is to schedule only the independent set that has the maximum total weight. Lets call that independent set, the solution independent set. Then we can make the following claim:

Theorem 3 Truthful Reporting $(\widehat{\mathbf{e}}_i = \mathbf{e}_i)$ is the dominant strategy for link $i, \forall i = 1, 2, ..., M$, i.e. $u_i(\mathbf{e}_i, \widehat{\mathbf{e}}_{V_{-i}}) \ge u_i(\widehat{\mathbf{e}}_i, \widehat{\mathbf{e}}_{V_{-i}})$ for all $\widehat{\mathbf{e}}_i$, for the server optimization problem given by (5.1) and (5.12).

Proof: Given a truthful report $\hat{\mathbf{e}}_i = \mathbf{e}_i$ and a set $F \subseteq V_{-i}$, we have to show that flipping bit \hat{e}_{ji} for all $j \in F$ will not increase u_i .

First note that the independent set computation only considers the presence or absence of an edge between two vertices, and ignores the edge directions. Therefore, if $\hat{e}_{ij} = 1$ for some $j \in V_{-i}$, i.e. link j has reported an incoming edge from link i, then the value of \hat{e}_{ji} has no influence on link i's utility. So, we need to consider only those sets F for which $\hat{e}_{ij} = 0$ for all $j \in F$.

Lets now partition the set F as $F = F_0 \cup F_1$, where $e_{ji} = 0$ if $j \in F_0$ and $e_{ji} = 1$ if $j \in F_1$. Flipping \hat{e}_{ji} 's for $j \in F_0$ will result in addition of false incoming edges to vertex i. This can only push link i out of the solution independent set and will clearly reduce u_i . On the other hand, flipping \hat{e}_{ji} 's for $j \in F_1$ might push i into the solution independent set (say I), but in that case, there must exist a link $j \in F_1$ that also becomes a part of the same independent set, because if I contains i but excludes all $j \in F_1$, then it must have been possible to construct I earlier itself. Since all links $j \in F_1$ interfere with link i, scheduling any one of them in the same independent set as link i will force u_i to zero.

5.2.2 Scheduling for ensuring weighted proportional fairness across user time allocations

Maximizing the weighted sum of time allocations of users can lead to an unfair time schedule where links outside the solution independent set are never scheduled and get zero utilities. To incorporate fairness in the time schedule produced by the spectrum server, one can choose the following set of weight functions in the server optimization problem (5.1):

$$w_i(x_i) = w_i \log\left(x_i\right), \forall i \tag{5.13}$$

This particular choice of weight functions corresponds to a notion of fairness called the *weighted proportional fairness* [64]. We can now make the following claim for a network of three (M = 3) interfering links:



Figure 5.1: All possible topologies for a three link network.

Theorem 4 Truthful Reporting ($\widehat{\mathbf{e}}_i = \mathbf{e}_i$) is the dominant strategy for link *i*, $\forall i = 1, 2, 3$, *i.e.* $u_i(\mathbf{e}_i, \widehat{\mathbf{e}}_{V_{-i}}) \geq u_i(\widehat{\mathbf{e}}_i, \widehat{\mathbf{e}}_{V_{-i}})$ for all $\widehat{\mathbf{e}}_i$, for the server optimization problem given by (5.1) and (5.13).

Proof: We will prove the above fact in a brute force manner by considering all possible topologies of a three link network. W.L.O.G. we will only look at the truthful reporting problem for link 1. Figure 5.1 lists all distinct topologies for a three link network. Note that since the construction of independent sets ignores the directed nature of the links, therefore figure 5.1 shows only undirected graphs resulting from the directed graphs. There are $2^3 = 8$ possible topologies that can be divided into groups A and B depending on the presence or absence of an edge between link 2 and link 3. If links 2 and 3 have reported no edge between them, then varying link 1's report will result in one of the topologies from group A to be constructed at the server. Similarly if link 2 and/or 3 have reported an edge between them, then changing 1's report will result in one of the topologies from group B to be constructed at the server. The server optimization problem given by (5.1) and (5.13) is easy to solve for the topologies given

in Figure 5.1, and the optimal time fractions for link 1 (x_1^*) are shown in the same figure. The figure also shows the maximal independent sets for each topology.

Lets first consider the topologies in group A. It can be verified that if link 1 falsely reports any present edges as missing, then he gets a zero utility. For example, if $e_{21} = 0$, $e_{31} = 1$ and $\hat{e}_{21} = 1$, $\hat{e}_{31} = 0$, then link 1 will be scheduled along with link 3 and will get zero utility. Also, if link 1 reports any additional edges which actually are not present, then he will get a lower utility. For example, if $e_{21} = 1$, $e_{31} = 0$ and $\hat{e}_{21} = 1$, $\hat{e}_{31} = 1$, then link 1's utility decreases from $w_1/(w_1 + w_2)$ to $w_1/(w_1 + w_2 + w_3)$.

Now consider the topologies in group B. In this case, if link 1 falsely reports any present edges as missing, then he will get a lower utility, which is not necessaily zero. For example, if $e_{21} = 1$, $e_{31} = 0$ and $\hat{e}_{21} = 0$, $\hat{e}_{31} = 0$, then link 1 will be scheduled with link 2 for a time fraction of $w_2/(w_2 + w_3)$ and with link 3 for a time fraction of $w_3/(w_2 + w_3)$. However since link 2 actually interfers with link 1, link 1's net utility will just be $w_3/(w_2 + w_3)$. Therefore $e_{21} = 1$, $e_{31} = 0$ and $\hat{e}_{21} = 0$, $\hat{e}_{31} = 0$ reduces link 1's utility from $(w_1 + w_3)/(w_1 + w_2 + w_3)$ to $w_3/(w_2 + w_3)$. Also, if link 1 reports any additional edges which actually are not present, then he will get a lower utility. For example, if $e_{21} = 1$, $e_{31} = 0$ and $\hat{e}_{21} = 1$, $\hat{e}_{31} = 1$, then link 1's utility decreases from $(w_1 + w_3)/(w_1 + w_2 + w_3)$ to $w_1/(w_1 + w_2 + w_3)$.

It is possible to verify the truthful reporting property for a network with four links but the exponential increase in the number of possible topologies makes the analysis more tedious. We have verified the truthful reporting property for larger number of links using numerical simulations and present the following conjecture:

Conjecture 5 Truthful Reporting ($\widehat{\mathbf{e}}_i = \mathbf{e}_i$) is the dominant strategy for link $i, \forall i = 1, 2, ..., M$, i.e. $u_i(\mathbf{e}_i, \widehat{\mathbf{e}}_{V_{-i}}) \ge u_i(\widehat{\mathbf{e}}_i, \widehat{\mathbf{e}}_{V_{-i}})$ for all $\widehat{\mathbf{e}}_i$, for the server optimization problem given by (5.1) and (5.13).

5.3 Conclusion and Discussion

In this chapter, we considered the a model of greedy interfering links that coordinate with each other through a mediating authority called the "spectrum server". Each link reports the set of links that interfere with it, based on which the spectrum server constructs the interference graph

and finds an optimal schedule for the links (maximizing a certain global objective). Since the links are greedy, they choose their reports to maximize their individual utilities. For certain user utility functions and global objective functions considered at the spectrum server, we showed that the links will truthfully report their sets of interfering links. Finding out a more general set of conditions under which the links report truthfully is part of future work. Another interesting area of future work would be designing mechanisms [65, Chapter 2] for enforcing truthfulness when links have an incentive to lie.

Chapter 6

Conclusions and Future Work

The sheer number and diversity of wireless applications being developed and deployed in unlicensed bands is increasing at a prolific rate. Due to the nature of wireless medium and the lack of any enforcing body pushing for resource segregation between the unlicensed band technologies, development of resource management algorithms suitable for unlicensed bands is of utmost importance. Interference avoidance is one promising direction along which many such resource management algorithms are geared, and one that has seen numerous interesting research efforts by the community in the last decade or so [21,22,27,28]. This thesis is a result of our belief that the *interference avoidance* approach is somewhat narrow in scope when trying to deal with plethora of resource management problems in unlicensed bands. This is because the term *interference avoidance* presumes that the interferer is someone who has to be avoided because he can't be expected to cooperate or coordinate, and one who is using the spectrum resource in a way that maximizes his greedy objective, which is assumed to at complete odds with the objective of the device being interfered. This infact is the true model for some resource management problem settings in the unlicensed bands, but in no way gives the complete picture. For example, at least in today's world, most wireless devices are mass manufactured, and their spectrum usage intelligence is hard-coded by those manufacturers in the form of etiquette protocols that the devices must follow while operating in the unlicensed bands. Also, it is not completely unimaginable for town municipalities to set up some coordination infrastructure to aid spectrum management between mutually interested parties, say for doing centralized channel allocation for 802.11 links operated by residents in an apartment complex. Therefore the wireless devices operating in unlicensed networks need not always be modeled as greedy, and the condition of complete lack of coordination between them could be too restrictive in certain scenarios. A useful theoretical model for studying wireless resource allocation in unlicensed bands would be one that allows the possibilities of the devices being greedy or non-greedy, and coordination being present or absent, and helps study the overall performance as a function

of say, the ratio of number of greedy versus non-greedy devices, and the degree of coordination available. This is quite a formidable task and one that we don't undertake in this thesis. Instead, in this thesis, as a precursor to this task, we have identified and analyzed four unlicensed band scenarios where the underlying resource management problem can be modeled with the four possible assumptions on the nature of devices (greedy/non-greedy) and level of coordination (present/ absent or minimal).

In chapter 2, we analyzed and simulated distributed interference avoidance (IA) based on covariance feedback broadcast from the receiver and incremental codeword changes by each user. The feedback could be a covariance matrix estimate from the receiver, or a sequence of received vectors $\mathbf{r}(n)$ to allow estimates to be constructed by each transmitter. The receiver tracks codeword changes by adapting the associated filters under a symbol error criterion. With perfect covariance feedback, the distributed method is equivalent in terms of codeword ensemble performance to centralized methods where codewords are computed by the receiver and conveyed to the transmitters.

In this scenario, each user acts on a greedy basis and the only coordination that is available between them is the common knowledge of the received covariance matrix which is broadcasted to them by the receiver. Note that we can call the periodic receiver broadcast as only a *minimal* coordination requirement because it does not scale with the number of users and only depends on the dimension of the signal space that they span.

Next, in chapter 3, we considered a model that is similar to the model of chapter 2 in that there are multiple transmitters and a single receiver, and the transmitters employ CDMA type codewords for transmitting their data. However, the big difference now is that the target application is that of sensors measuring a physical process in a field and transmitting to a central repository. The symbols that need to be transmitted by the sensors therefore could be correlated and since all the sensors are deployed by a single authority, it could be possible for the sensors to operate in a coordinated way. We show that if the common objective of the sensors is to jointly minimize the total mean square error, then the optimal codewords here are no longer the same as those corresponding to the scenario of chapter 2. This illustrates the importance of making the right assumptions on the device objectives and level of coordination available, when considering problems in unlicensed bands.

We further characterized the performance of our CDMA based transmission scheme in the form of a total power-TMSE tradeoff function, and compared with an information-theoretic outer bound. Our scheme, which can be viewed an uncoded transmission scheme with optimal power allocations ovr parallel channels, is found to be always suboptimal compared to the outer bound at all SNR's. This is in contrast to a recent result for the scalar version of this scheme [15], where it is shown that uncoded transmission is optimal below a threshold SNR. We also considered a separation based scheme and showed that at low SNRs, our scheme's performance is slightly better than an upper bound on the performance of the separation based scheme. However, a fair comparison between the schemes should take into account the implementation complexity as well. While the separation based scheme requires quantizers, so-phisticated source and channel coders and decoders, our scheme is a essentially an un-coded transmission scheme and therefore should be much simpler to implement in practice.

In chapter 4, we considered the problem of scheduling interfering links and came up with a distributed algorithm which requires absolutely no coordination between the links and allows them to achieve their desired rates. It is assumed however that the links follow the algorithm honestly and their actions are not dictated by pure greed. We first compared the achievable throughput region of a simple probabilistic transmission scheme with that of centralized scheduling. We then presented a distributed implementation of the probabilistic transmission scheme which can achieve any feasible rate vector in the throughput region of the probabilistic transmission scheme. We showed that the algorithm is guaranteed to converge for any underlying PHY layer that ensures that rate on a link goes down strictly as the interference increases. Our distributed algorithm does not require any level of coordination or information sharing between the links. The only information exchange happens between the receiver of a link and the corresponding transmitter, where the receiver computes the instantaneous link throughput and feeds it back to the transmitter. The above properties of the algorithm make it attractive for adoption in unlicensed band scenarios where centralized control is difficult to implement. Also, since our algorithm allows the participating links to achieve any point in the feasible rate region, it could be used for guarantying different QoS levels for different users. This feature of our algorithm could be seen as a distinct advantage over conventional 802.11 based multiaccess schemes that don't have any such provision for guarantying differing QoS levels across

users. There are several possible directions for future work here:-

- We proved that the iteration defined by equation (4.27) converges to the desired operating point if we start from the all zero probability vector. Complete characterization of the convergence/divergence conditions of the iteration for all starting points is part of future work.
- 2. Finding out the running time complexity of the iteration (4.27) will be useful from a practical point of view and is part of future work.
- 3. Iteration (4.27) assumes that each link can measure its instantaneous rate at the receiver and use it to update its transmission probability. However, in practice the receiver has to estimate this rate by measuring and averaging the throughput over a certain time interval. For practical implementation of our distributed procedure, effects of the rate estimation error on the iteration must be studied numerically and possibily a stochastic convergence result along the lines of [34] could be developed.
- 4. The probabilistic transmission scheme considered in this chapter has the nice property that absolutely no coordination is required among different links. However the flip side of this benefit is that the achievable rate region could be strictly smaller than that achievable through centralized scheduling. Going from chapter 2 to chapter 3, we observed that allowing the possibility of correlation between the transmitted signature waveforms helps in achieving a higher value of performance objective. Similarily in this case, if we view the transmissions of any link over a sequence of time slots as its random *codeword*, then maybe allowing correlation between these codewords can help us in enlarging the achievable rate region. This observation is true and in fact if we can impose an arbitrary joint probability distribution over the link transmissions, then we can achieve the complete rate region corresponding to the centralized scheduling scheme. How to achieve correlation between the link transmissions without loosing the benefit of requiring no coordination among the links is then the main question. One promising approach in this direction (that is not explored in this thesis) is the introduction of *common randomness* in the link transmissions, maybe through the broadcast of the result of a random experiment

(say coin toss) by a coordinating authority. Each links combines this broadcast information with result of its own independent random experiment decide whether to transmit or not in a given time slot. This would increase the coordination requirements only by a minimal amount. Infact, the same amount of coordination was required in the model in chapter 2 where the receiver periodically broadcasts the covariance matrix or the received vector.

The work in chapter 4 suggests a new twist to the greedy/strategic distributed resource allocation problem. That is, the resource allocation process of chapter 4 could be carried out by a centralized mediating authority called the *spectrum server* [5,6,13,54], that receives local interference reports from the links, and instructs them on spectrum usage. The actual performance figures and how users affect one another is considered public (and true) information.

However, in a distributed system aided by a spectrum server but without a provision for "spectrum police" – a facility that measures mutual interference in an objective fashion and reports it to the spectrum server – there could be motivation for users to lie about the levels of interference they experience.

We explored this issue in chapter 5 by formulating a centralized scheduling problem with greedy or strategic links. Each link reports the set of links that interfere with it, based on which the spectrum server constructs the interference graph and finds an optimal schedule for the links (maximizing a certain global objective). Since the links are greedy, they choose their reports to maximize their individual utilities. For certain user utility functions and global objective functions considered at the spectrum server, we showed that the links will truthfully report their sets of interfering links. Finding out a more general set of conditions under which the links report truthfully is part of future work. Another interesting area of future work would be designing mechanisms [65] for enforcing truthfulness when links have an incentive to lie.

References

- [1] I. Akyildiz, Y. Sankarasubramaniam, and E. Cayirci, "A Survey on Sensor Networks," *IEEE Communications Magazine*, vol. 40, no. 8, pp. 102–114, August 2002.
- [2] T. Cover and J. Thomas, *Elements of Information Theory*. Wiley-Interscience, 1991.
- [3] J. Singh and C. Rose, "Channel probing under a power budget," in *Proc. CISS*, 2006, princeton, NJ.
- [4] F. C. Commission, "FCC Report and Order 97-5: Amendment of the commission's rules to provide for operation of unlicensed NII devices in the 5 GHz frequency range," 1997, eT Docket No. 96-102.
- [5] C. Raman, R. Yates, and N. Mandayam, "Scheduling variable rate links via a spectrum server," in *Proc. IEEE DySPAN*, 2005, baltimore, MD.
- [6] R. Yates, C. Raman, and N. Mandayam, "Fair and efficient scheduling of variable rate links via a spectrum server," in *Proc. IEEE ICC*, Jun 2006, istanbul, Turkey.
- [7] W. Yu, W. Rhee, S. Boyd, and J. M. Cioffi, "Iterative Water-filling for Gaussian Vector Multiple Access Channels," *IEEE Transactions on Information Theory*, vol. 50, no. 1, pp. 145–151, January 2004.
- [8] N. Clemens and C. Rose, "Intelligent power allocation strategies in an unlicensed spectrum," in *Proc. IEEE DySPAN*, 2005, baltimore, MD.
- [9] R. Etkin, A. Parekh, and D. Tse, "Spectrum sharing for unlicensed bands," in *Proc. IEEE DySPAN*, 2005, baltimore, MD.
- [10] X. Jing and D. Raychaudhuri, "A spectrum etiquette protocol for efficient coordination of radio devices in unlicensed bands," *Proceedings of PIMRC, Beijing, China*, September 2003.
- [11] J. Lee, M. Chiang, and R. Calderbank, "Utility-optimal random access control," *IEEE Trans. on Wireless Communications*, vol. 6, no. 7, pp. 2741–2751, July 2007.
- [12] N. Mandayam, "Cognitive algorithms and architectures for open access to spectrum," *The Conference on the Economics, Technology and Policy of Unlicensed Spectrum, East Lansing, MI*, available online at: http://quello.msu.edu/conferences/spectrum/papers/ man-dayam.pdf.
- [13] O. Ileri, D. Samardzija, T. Sizer, and N. Mandayam, "Demand responsive pricing and competitive spectrum allocation via a spectrum policy server," in *Proc. IEEE DySPAN*, Nov 2005, baltimore, MD.
- [14] V. Rodriguez, K. Moessner, and R. Tafazolli, "Auction driven dynamic spectrum allocation: optimal bidding, pricing and service priorities for multi-rate, multi-class CDMA," *IEEE International Symposium on Personal, Indoor and Mobile Radio Communications PIMRC*, Sept 2005.

- [16] S. Verdu, "Capacity Region of Gaussian CDMA Channels: The Symbol–Synchronous Case," Proceedings of the 24th Allerton Conference on Communication, Control and Computing, 1989.
- [17] M. Rupf and J. Massey, "Optimum sequence multisets for synchronous code-division multiple-access channels," *IEEE Transactions on Information Theory*, vol. 40, no. 4, pp. 1226–1266, July 1994.
- [18] P. Viswanath and V. Anantharam, "Optimal Sequences and Sum Capacity of Synchronous CDMA Systems," *IEEE Transactions on Information Theory*, vol. 45, no. 6, pp. 1984– 1991, September 1999.
- [19] —, "Optimal Sequences for CDMA with colored noise: A Schur-Saddle Function Property," *IEEE Transactions on Information Theory*, vol. 48, no. 6, pp. 1295–1318, June 2002.
- [20] P. Rapajic and B. Vucetic, "Linear adaptive transmitter-receiver structures for asynchronous CDMA," *European Trans. on Telecomm.*, vol. 6, no. 1, pp. 21–28, January 1995.
- [21] S. Ulukus and R. D. Yates, "Iterative construction of optimum signature sequence sets in synchronous CDMA systems," *IEEE Trans. Info. Theory*, vol. 47, no. 5, pp. 1989–1998, July 2001.
- [22] C. Rose, "CDMA codeword optimization interference avoidance and convergence via class warfare," *IEEE Transactions on Information Theory*, vol. 47, no. 7, pp. 2368–2382, September 2001.
- [23] D. C. Popescu and C. Rose, "Codeword Quantization for Interference Avoidance," in Proc. 2000 Int. Conf. on Acoustics, Speech, and Signal Processing - ICASSP 2000, vol. 6, June 2000, pp. 3670–3673, istanbul.
- [24] D. Tabora, "An Analysis of Covariance Estimation, Codeword Feedback, and Multiple Base Performance of Interference Avoidance," Master's thesis, Rutgers University, Dept. of Electrical and Computer Engineering, 2001, thesis Director: Prof. C. Rose.
- [25] W. Santipach and M. L. Honig, "Signature Optimization for DS-CDMA with Limited Feedback," in 2002 International Symposium on Spread Spectrum Systems and Applications, Prague, Czech Republic, September 2002.
- [26] "Special issue on software radio," *IEEE Personal Communications Magazine*, vol. 6, no. 4, August 1999, editors: K-C. Chen and R. Prasad and H.V. Poor.
- [27] D. C. Popescu and C. Rose, "Interference Avoidance for Wireless Systems," in *Wiley Encyclopedia of Telecommunications*, December 2002.
- [28] C. Rose, S. Ulukus, and R. Yates, "Interference Avoidance in Wireless Systems," *IEEE Transactions on Wireless Communications*, vol. 1, no. 3, pp. 415–428, July 2002, (also available at www.winlab.rutgers.edu/~crose/papers/avoid17.ps).

- [29] G. Strang, *Linear Algebra and Its Applications*, 2nd ed. Academic Press, 1992.
- [30] S. Ulukus and A. Yener, "Iterative Transmitter and Receiver Optimization for CDMA Networks," *IEEE Transactions on Wireless Communications*, vol. 3, no. 6, pp. 1879– 1884, November.
- [31] S. Haykin, Communication Systems. John Wiley & Sons, 1994.
- [32] S. Orfanidis, Optimum Signal Processing. Newyork: Mc-Graw Hill, 1988.
- [33] D. Popescu and C. Rose, Interference Avoidance for Wireless Systems. Kluwer Academic Publishers, Information Technology: Transmission, Processing, and Storage series, March 2004.
- [34] S. Ulukus and R. Yates, "Stochastic power control for cellular radio systems," *IEEE Trans. Comm.*, vol. 46, no. 6, pp. 784–798, June 1998.
- [35] F. Rellich, *Perturbation Theory of Eigenvalue Problems*. Newyork: Gordon and Breach, 1992.
- [36] B. Champagne, "Adaptive eigendecomposition of data covariance matrices based on firstorder perturbations," *IEEE Transactions on Signal Processing*, vol. 42, no. 10, pp. 2758– 2770, Oct 1994.
- [37] P. Viswanath, V. Anantharam, and D. Tse, "Optimal sequences, power control and capacity of spread-spectrum systems with multiuser receivers," *IEEE Transactions on Information Theory*, vol. 45, no. 6, pp. 1968–1983, 1999.
- [38] G. Woodward, R. Rasatuk, M. L. Honig, and P. B. Rapajic, "Minimum mean-squared error multiuser decision-feedback detectors for DS-CDMA," *IEEE Trans. on Communications*, vol. 50, no. 12, December 2002.
- [39] O. Kaya and S. Ulukus, "Jointly Optimal Power and Signature Sequence Allocation for Fading CDMA," December 2003.
- [40] H. Viswanathan and T. Berger, "The quadratic gaussian CEO problem," *IEEE Transac*tions on Information Theory, vol. 42, no. 10, pp. 2758–2770, Oct 1994.
- [41] V. Sergio, *Multiuser Detection*. Cambridge, UK: Cambridge University Press, 1998.
- [42] A. Marshall and I. Olkin, *Inequalities: Theory of Majorization and Its Applications*. New York, NY: Academic Press, 1979.
- [43] O. Popescu, C. Rose, and D. C. Popescu, "Maximizing the Determinant for a Special Class of Block-Partitioned Matrices," *Mathematical Problems in Engineering*, vol. 2004, no. 1, pp. 49–62, May 2004.
- [44] S. Boyd and L. Vandenberghe, "Convex optimization," *Cambridge University Press*, 2004.
- [45] D. C. Popescu and C. Rose, "Codeword optimization for uplink CDMA dispersive channels," *IEEE Transactions on Wireless Communications*, vol. 4, no. 4, pp. 1563–1574, July 2005.

- [46] T. Cover, A. A. El-Gamal, and M. Salehi, "Multiple Access Channels with Arbitrarily Correlated Sources," *IEEE Transactions on Information Theory*, vol. 26, no. 6, pp. 648– 657, September 1980.
- [47] L. Vandenberghe, S. Boyd, and S.-P. Wu, "Determinant maximization with linear matrix inequality constraints," *SIAM Journal on Matrix Analysis and Applications*, vol. 19, no. 2, pp. 499–533, 1998.
- [48] M. Gastpar and M. Vetterli, "Power, spatio-temporal bandwidth, and distortion in large sensor networks," *IEEE Journal on Selected Areas in Communications*, vol. 23, no. 4, pp. 745–754, April 2005.
- [49] A. B. Wagner, S. Tavildar, and P. Viswanath, "Rate region of the quadratic gaussian twoterminal source-coding problem," *IEEE Transactions on Information Theory (submitted)*, 2005. [Online]. Available: http://www.citebase.org/abstract?id=oai:arXiv.org:cs/0510095
- [50] M. Gastpar, "Uncoded transmission is exactly optimal for a simple gaussian sensor network," *Proc. 2007 Information Theory and Applications Workshop, San Diego, CA, USA*, February 2007.
- [51] H. S. Witsenhausen, "On sequence of pairs of dependent random variables," *SIAM Journal Applied Math*, vol. 28, pp. 100–113, January 1975.
- [52] H. O. Lancaster, "Some properties of the bivariate normal distribution considered in the form of a contingency table," *Biometrica*, vol. 44, pp. 289–292, 1957.
- [53] S. Haykin, "Cognitive radio: brain-empowered wireless communications," *IEEE Journal on Selected Areas in Communications*, vol. 23, no. 2, pp. 201–220, Feb 2005.
- [54] C. Raman, R. Yates, and N. Mandayam, "Cross layer scheduling of end-to-end flows using a spectrum server," in *Proc. CISS*, 2006, princeton, NJ.
- [55] X. Lin and N. B. Shroff, "The impact of imperfect scheduling on cross-layer rate control in multihop wireless networks," in *IEEE Infocom*, 2005.
- [56] B. Hajek and T. V. Loon, "Decentralized dynamic control of a multiaccess broadcast channel," *IEEE Trans. Auto. Control*, vol. 27, pp. 559–569, Jun 1982.
- [57] S. Ghez, S. Verdu, and S. C. Schwartz, "Stability properties of slotted ALOHA with multipacket reception capability," *IEEE Trans. Auto. Control*, vol. 33, no. 7, pp. 640–649, Jul 1988.
- [58] K. Kar, S. Sarkar, and L. Tassiulas, "Achieving proportional fairness using local information in aloha networks," *IEEE Trans. Auto. Control*, vol. 49, no. 10, pp. 1858–1862, Oct 2004.
- [59] X. Wang and K. Kar, "Distributed approaches for proportional and max-min fairness in random access ad hoc networks," in *Proc. CISS*, 2006, princeton, NJ.
- [60] J. Massey and P. Mathys, "The collision channel without feedback," *IEEE Trans. Info. Theory*, vol. 31, no. 2, pp. 192–204, Mar 1985.

- [61] L. Kleinrock and F. Tobagi, "Packet switching in radio channels: Part I: Carrier sense multiple access and their throughput delay characteristics," *IEEE Trans. Communications*, vol. 23, no. 12, pp. 1400–1412, Dec 1975.
- [62] P. Gupta and A. Stolyar, "Optimal throughput allocation in general random-access networks," in *Proc. CISS*, 2006, princeton, NJ.
- [63] B. Radunovic and J. Y. L. Boudec, "Rate performance objectives of multihop wireless networks," *IEEE Trans. Mobile Computing*, vol. 3, no. 4, pp. 334–349, Oct.-Dec 2004.
- [64] F. Kelly, "Networks of Queues," Advances in Applied Probability, pp. 416–423, 1976.
- [65] D. C. Parkes, "Iterative combinatorial auctions," Ph.D. dissertation, University of Pennsylvania, May 2001.

Curriculum Vitae Jasvinder Singh

- 2007 PhD in Electrical and Computer Engineering, Rutgers The State University of New Jersey
- 2003 M.S. in Electrical and Computer Engineering, Rutgers The State University of New Jersey
- 2000 B.Tech. in Electrical Engineering from Indian Institute of Technology, Kharagpur, India
- 2002-2006 Graduate Research Assistant, WINLAB, Rutgers The State University of New Jersey
- 2007 Jasvinder Singh, Joydeep Acharya, Christopher Rose, "A CDMA Based Scheme for Transmission of Correlated Gaussian Symbols over a Gaussian Multiple Access Channel," to be submitted to IEEE Transactions on Wireless Communications.

Jasvinder Singh, Christopher Rose, "Distributed Incremental Interference Avoidance," submitted to IEEE Transactions on Wireless Communications.

Chandru Raman, Jasvinder Singh, Roy Yates, Narayan Mandayam, "Scheduling in Cognitive Networks: Scheduling variable rate links - centralized and decentralized approaches," In Cognitive Wireless Networks, Fitzek, Frank H.P.; Katz, Marcos D. (Eds.) 2007, Springer, Signals and Communication series.

2006 Jasvinder Singh, Chandru Raman, Roy Yates, Narayan Mandayam, "Random Access for Variable Rate Links," in *Proc. IEEE Military Communications Conference (MILCOM)*, October 2006.

Jasvinder Singh, Christopher Rose, "Channel Probing under a Power Budget," in *Proc. 40th Annual Conference on Information Sciences and Systems (CISS)*, March 2006.

- **2004** Joydeep Acharya, Ritabrata Roy, Jasvinder Singh, Christopher Rose, "Optimal Signature Sets for Transmission of Correlated Information over a Multiple Access Channel," in *IEEE Global Telecommunications Conference (GLOBECOM)*, November 2004.
- **2003** Jasvinder Singh, Christopher Rose, "Distributed Incremental Interference Avoidance," in *IEEE Global Telecommunications Conference (GLOBECOM)*, November 2003.