Throughput and Complexity Tradeoffs in the Multi-antenna Downlink

by

Charles Swannack
B.S., Computer Engineering (2003)
Clemson University

Submitted to the Department of Electrical Engineering and Computer Science in partial fulfillment of the requirements for the degree of

Masters of Science in Electrical Engineering and Computer Science

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 2005

Author ....
Department of Electrical Engineering and Computer Science
May 18, 2005

Certified by ....
Gregory W. Wornell
Professor, Department of Electrical Engineering and Computer Science
Thesis Supervisor

Certified by ....
Dr. Elif Uysal-Biyikoglu
Postdoctoral Lecturer of Electrical Engineering and Computer Science
Thesis Supervisor

Accepted by ....
Arthur C. Smith
Chairman, Department Committee on Graduate Students

Throughput and Complexity Tradeoffs
in the Multi-antenna Downlink

by

Charles Swannack
B.S., Computer Engineering (2003)
Clemson University

Submitted to the Department of Electrical Engineering and Computer Science
on May 18, 2005, in partial fulfillment of the
requirements for the degree of
Masters of Science in Electrical Engineering and Computer Science

Abstract

This thesis considers the joint design of the multiplexing and scheduling of independent data streams for the Gaussian multiple-input multiple-output (MIMO) broadcast channel. It is well known that the use of multiple transmit antennas can greatly increase the capacity of the broadcast channel. However, the complexity of a capacity-achieving strategy is dominated by the underlying search for the best user subset to multiplex across the transmitter array, which can be prohibitively high if the number of users is much greater than the transmit dimension. To reduce this complexity, one can limit the search to a smaller set of users while ensuring that this restricted pool contains a set that is close to optimal with high probability. To this end, we define sets with guaranteed signal to interference ratio (SIR) and signal to noise ratio (SNR) values.

We provide bounds on the probability that such a set exists. These bounds are derived through an interpretation of the multi-user multi-antenna channel as a random packing of the unit sphere. As such, we provide refined estimates on the content of spherical caps so that they can be applied as a model for interference. We then employ recent developments in the area of random geometric graph theory to characterize the probability of existence. We further show there is a phase transition phenomenon in channel geometry that can be used in the design of efficient algorithms for scheduling in the MIMO broadcast channel. Further, we use this transition to provide novel lower bounds for the expected rate of any multiplexer in a channel with choice.

Thesis Supervisor: Gregory W. Wornell
Title: Professor, Department of Electrical Engineering and Computer Science

Thesis Supervisor: Dr. Elif Uysal-Biyikoglu
Title: Postdoctoral Lecturer of Electrical Engineering and Computer Science

Acknowledgments

This work could have not have been completed if not for the wonderful discussions I have had while at MIT. I am particularly thankful to Greg Wornell and Elif Uysal for introducing me to this problem and helping me as I struggled through it early on. Additional thanks to Uri Erez, who always provided good conversations about coding on our daily trips to building 4 for coffee. Also, I owe a great deal to my office mate Hiroyuki Ishii who always kept me upbeat and on occasion made sure that I remembered to stop working and head home before the T stopped running. Additionally, I am indebted to the rest of the Signals, Information and Algorithms Lab, particularly, Giovanni Aliberti, Albert Chan, Vijay Divi, Everest Huang, Ying-zong Huang, Emin Martinian, Tricia Mulcahy and Maryam Shanechi for always allowing me to ask questions and for often motivating new thoughts and ideas.

I am also greatful for the wonderful community at MIT. Especially, Alan Edelman, Vivek Goyal, Al Oppenheim and Moe Win for their time, advice and discussions. Also, I owe a great deal to the greater "36-6XX" community (DSPG, LIDS and STIR) for always allowing me to drop by for a conversation. I am especially greatful to the STIR group for allowing me access to their espresso machine which helped polish off many papers and helped me meet many deadlines.

Also, this work would have not have been possible with the early advising and mentoring of Neil Calkin, Chris Cox, Kevin James and Dan Stanzione. All of whom spent far more time with me than was required. I am greatful to them for introducing me to research and the excitement of academic discovery. I am additionally greatful to them for providing me with a challenging environment in which I could learn and and pursue new ideas. Additional thanks to all those in the Algebra and Discrete Mathematics Group, the Center for Advanced Engineering Fibers and Films (CAEFF) and the Parallel Architecture Research Lab (PARL) at Clemson University for past conversations and continued support.

Most importantly, I would like to thank my friends and family. This work could not have been completed without their unending support.

This material is based upon work supported under a National Science Foundation Graduate Research Fellowship and by the HP-MIT Alliance.

Contents

List of Figures

Chapter 1

Introduction

Wireless data traffic has seen remarkable growth over the past several years and is expected to have continued growth in the future. The increased demand has posed a particularly challenging problem for system designers who have power, bandwidth and complexity constraints at transmitters and receivers. A proposed solution to this problem is to take advantage of spatial diversity by using multiple antennas at the transmitter. This can improve spectral efficiency in environments with high scattering at no cost in power or bandwidth [13]. In systems with multiple users, multiple-antenna systems can additionally be used to multiplex independent data streams to different users in some geographical area.

Developing efficient wireless multiuser communication systems is a problem of substantial interest. An example for such a system is the wireless downlink (as depicted in Figure 1-1) where independent data streams need to be transmitted to users that are geographically

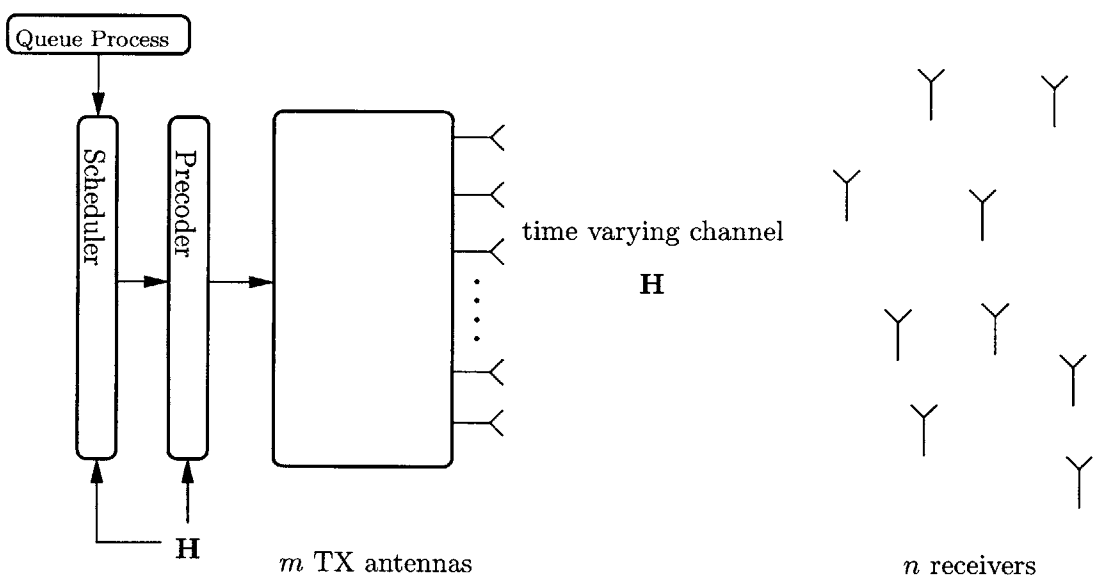

Figure 1-1. The MIMO downlink system overview

distributed. It is well known that using multiple antennas can greatly increase the capacity of the broadcast channel [29]. Multiplexing users (multiplexing multiple users' data at the same time) can potentially further increase the throughput of the downlink system. Time variation in the channel states of users leads to the question of which users to choose to encode at a given time to satisfy some overall time-averaged performance criterion. When the number of users, n, is larger than the number of antennas, m, the multiple-input multiple-output (MIMO) broadcast channel contains a quite rich joint scheduling and multiplexing problem. The richness of this system design problem stems from the fact that it is one of *spatio-temporal scheduling* , *i.e.*, both temporal scheduling and spatial multiplexing aspects of the design must be considered. This problem has been considered in [20], where a variety of scheduling algorithms were proposed. Additionally, a number of mostly heuristic approaches have been proposed [3,12,24,2] to explore this multiplexing/scheduling problem space.

This thesis has been motivated by the seemingly prohibitive complexity of this joint scheduling/multiplexing problem. In a MIMO channel with choice over users, one expects to improve a particular performance criterion as a growing user pool is searched. This could be maximizing total throughput (or sum rate), for example. The complexity of such an optimization is dominated by the underlying search for the best (possibly ordered) user subset to multiplex across the transmitter array, which must be performed each time the system changes state. To reduce this complexity, one may limit the search to a smaller pool of users while ensuring that a user set that will be found in this restricted pool is close to optimal with high probability.

In order to provide a solution to the scheduling problem that yields a low complexity selection algorithms while additionally providing close to optimal performance, we shall define sets with certain signal-to-interference ratio (SIR) and signal-to-noise ratio (SNR) guarantees. These guarantees are defined by placing restrictions on the norm of users channels and the magnitudes of the inner products. The central goal of the thesis is to characterize the probability of finding such a set in a pool of n users. We will show that the probability exhibits a quite sharp transition from 0 to 1 with increasing n. We will obtain upper bounds on the rate at which the SNR and SIR guarantees can be increased while maintaining a non-zero probability that a set with those guarantees exists. More specifically, we show that as a function of the number of users that have been examined, k, the probability of finding a near-orthogonal set passes through a threshold, after which it behaves like $\Theta(k^{-m})$.

The remainder of this thesis is organized as follows. In Chapter 2 we begin by providing an overview of the MIMO downlink and channel model. We examine various multiplexing strategies for the multiple antenna downlink with an emphasis on multiuser scheduling. In particular we review the multiplexing techniques of superposition, zero-forcing and dirty-paper multiplexing.

In Chapter 3 we propose a simple selection procedure provide analysis of its performance. In particular, we define sets that meet a particular SIR and SNR guarantees. We provide an interpretation of the existence of such sets through random packing in Section 3.1 and show that the problem of existence is equivalent to that of finding a complete graph in a larger random graph in Section 3.2. In order to provide good bounds on the probability of near orthogonal sets in small dimensions we refine the estimates of Shannon [25] on the fraction of area covered by a cap on the unit sphere. In particular we provide estimates that are non-trivial in a neighborhood around $\pi/2$. In Section 3.4 we uses the bounds on the probability of existence to lower bound the expected rate of zero-forcing multiplexing in a channel in which there is a choice of the users to select.

Chapter 2

Multiplexing Strategies for MIMO Broadcasting

We begin by providing an overview of our formulation of the MIMO downlink. We will focus on the problem where there are independent data streams arriving at a transmit base equipped with multiple antennas and destined for a set of uncoordinated receivers (as depicted in Figure 1-1). In particular, we will address the problem when the number of receivers, n, is much greater than the number of transmit antennas, m. This scenario leads to a choice over which user to select to multiplex across the transmit array in order to satisfy some time-averaged performance criterion. In such a *channel with choice*, we can choose a channel from among some dependent set of channel realizations over which to communicate.

In order to make our discussion precise, we begin by presenting our model for the MIMO broadcast channel with choice. After we have stated our model and assumptions, we present the standard multiplexing techniques for the MIMO broadcast. When possible, we will present a geometric approach to these multiplexing techniques that will help not only aid in the development of selection algorithms, but also aid in the analysis.

2.1 System Model

We will consider the additive white Gaussian noise (AWGN) broadcast channel as depicted in Figure 2-1. We consider the general case with an m-antenna transmitter and n uncoordinated receivers each having a single receive antenna. We will denote the set of n users by $\mathcal{U} = \{1, 2, \dots, n\}$ and let $\mathcal{A} \subset \mathcal{U}$ be an arbitrary subset of users. If $\mathcal{A}$ is the subset of users chosen to receive non-zero rate we will call $\mathcal{A}$ the *active* set of users or call $\mathcal{A}$ the *activation set.* We will assume the standard input-output model for the channel.

Let $\mathbf{u} \in \mathbb{C}^n$ be the vector of message symbols for the n users and let $\mathbf{x} \in \mathbb{C}^m$ be

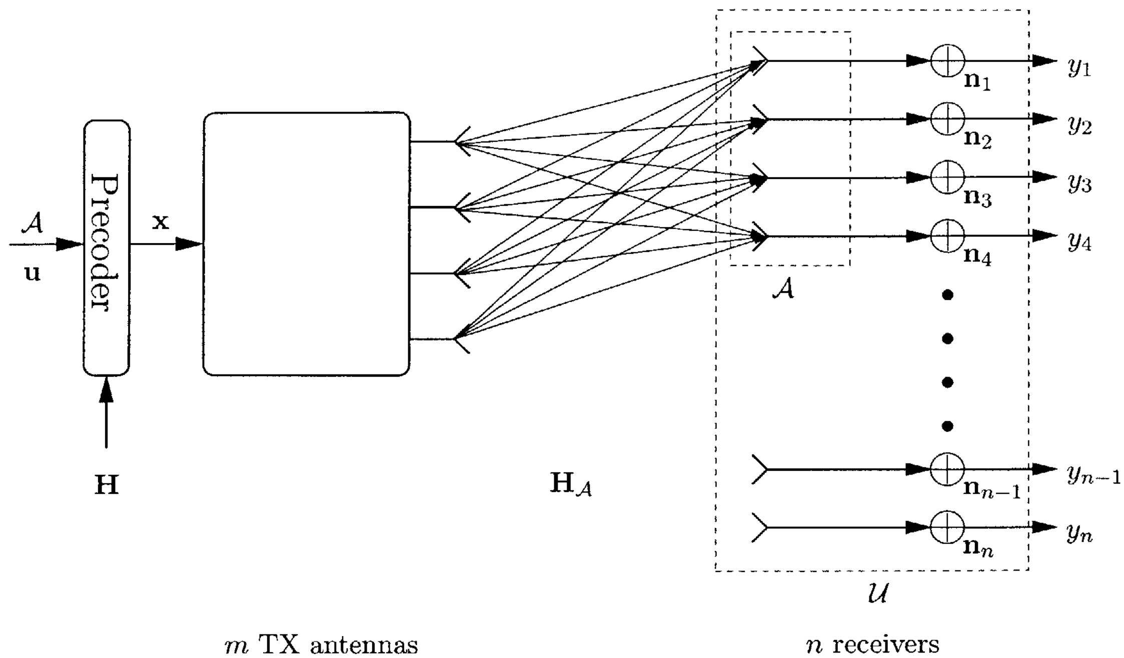

Figure 2-1. The AWGN Gaussian MIMO broadcast channel model

the transmitted signal vector and $\mathbf{h}_i \in \mathbb{C}^{1\times m}$ be the channel of the ith user. We will let $\mathbf{H} \in \mathbb{C}^{n\times m}$ be the channel matrix of all users and let $\mathbf{H}_{\mathcal{A}}$ be the channel matrix of the set of users $\mathcal{A}$. We further assume the channel vectors $\mathbf{h}_i$ are distributed as *i.i.d.* complex circularly symmetric Gaussian m-vectors such that $\|\mathbf{h}_i\| = 1$. Under the assumption of complex circularly symmetric Gaussian noise we have, for any chosen channel matrix $\mathbf{H}_{\mathcal{A}}$,

$$\mathbf{y} = \begin{bmatrix} y_{a_1} \\ \vdots \\ y_{a_{|\mathcal{A}|}} \end{bmatrix} = \mathbf{H}_{\mathcal{A}}\mathbf{x} + \mathbf{n} \quad \text{where } \mathbf{H}_{\mathcal{A}} = \begin{bmatrix} \mathbf{h}_{a_1} \\ \vdots \\ \mathbf{h}_{a_{|\mathcal{A}|}} \end{bmatrix}$$

and $\mathbf{n} \sim \mathcal{N}_{\mathbb{C}}(0, \mathbf{I}_n)$. An upper bound on the input covariance is assumed, which corresponds to an total input power constraint of P which takes the form:

$$\operatorname{Tr}(\mathbf{K}_{xx}) \leq P \tag{2.1}$$

where $\mathbf{K}_{xx} = \mathbb{E}\left[\mathbf{x}\mathbf{x}^{\dagger}\right]$. Further, we make the assumption that the channel is slow *block-fading*, *i.e.* the channel will be assumed to vary slow enough so that we can use a sufficiently long code for error probabilities to be close to zero. As depicted in Figure 2-1, we can imagine

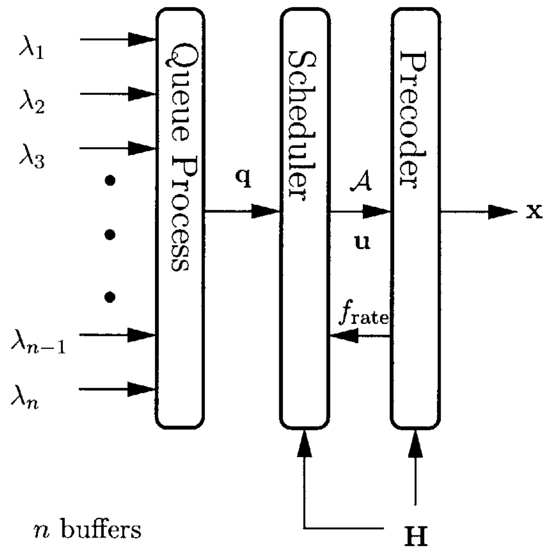

Figure 2-2. The queuing and scheduling model for the MIMO downlink

that the precoder is given a set of messages, $\mathbf{u}$, and user indices $\mathcal{A}$. Then, then precoder multiplexes the signal only considering the sub-channel $\mathbf{H}_{\mathcal{A}}$. We now consider how one may select the set $\mathcal{A}$ to insure throughput optimality.

The queueing model is depicted in Figure 2-2. We will assume that arrivals occur at the beginning of every scheduling interval and that each arriving packet is destined for a single user. The arriving packet is placed into a buffer corresponding to that user. The scheduler is then informed of both the queue state, $\mathbf{q}$, and the current channel realization, $\mathbf{H}$. We will assume that the scheduler has been given some objective function in which to maximize, corresponding to some precoder. In general we will denote the objective function to be maximized as $f_{\text{rate}}(\mathbf{H}, \mathbf{q})$ and replace the subscript when we wish to discuss specific functions. At each interval the scheduler selects the activation set $\mathcal{A}$ and message symbols $\mathbf{u}$ and forwards this set and vector to the precoder which forms the signal to be multiplexed across the array.

We have presented a throughput optimal policy based on selecting a group of users with *maximum weight* at any time [27]. That is, choosing the set $\mathcal{A}$ and power allocation that solves

$$\max_{\mathbf{r} \in \mathcal{C}(\mathbf{H}, P)} \mathbf{q} \cdot \mathbf{r} \tag{2.2}$$

where $\mathcal{C}(\mathbf{H}, P)$ is the region of achievable rates under a power constraint of P. The optimality of this policy was originally proven under a simpler communication model where the coding scheme was restricted to a choice of rates from a finite set [28]. Using the methods of [28], a throughput optimal resource allocation policy for a downlink channel, again under simplifying assumptions on data transmission, was shown in [22]. Further advances were made within a similar framework, but incorporating information-theoretically optimal coding and power adaptation, in [32, 39]. In [39] it was shown that a certain maxweight policy achieves absolute upper bounds on throughput for a single-antenna fading broadcast channel.

In many applications one is interested in finding a set $\mathcal{A}$ that maximizes some objective function that additionally depends strongly on $\mathbf{H}_{\mathcal{A}}$. This optimization in general may be very complex since one may have to search over all subsets of $\mathcal{U}$. Additionally, since as the systems changes states, the channel matrices $\mathbf{H}_{\mathcal{A}}$ vary independently from the current state, a solution may need to be recomputed at each step. Thus, it is important to develop algorithms that efficiently solve or approximate the solution to (2.2). For example, we can restrict our search to that of finding a set of near orthogonal channel vectors. If there exists a set of channel vectors, $\mathcal{A} \subset \mathcal{U}$, that are orthogonal, multiplexing is trivial. More generally, a near-orthogonal set allows zero-forcing, beamforming or other multiplexers to perform well. In the following sections we examine the effects of such channel geometry on various precoders. Then in Chapter 3, we examine near orthogonal sets as an approximation to any optimal multiplexing set.

■ 2.2 Multiplexing for the MIMO Broadcast

As noted previously, the use of multiple antennas can greatly increase the capacity of the broadcast channel. However, when independent messages are transmitted to uncoordinated users mutual interference can occur between users. This mutual interference is a *fundamental* limitation to the capacity of multi-users systems. However, in the case where the channel is slow fading so that the channel may be accurately estimated and feedback to the transmitter, there are a variety of ways the channel state information may be used. For example, the transmitter may adapt power allocation so that users receive proportional rate or the transmitter may employ intelligent multiplexing techniques to combat the known interference.

The fact that the transmitter has a choice for the power allocation for different users leads to a notion of capacity different from that of the single user ergodic capacity. That is we have a maximal achievable rate region as opposed to a maximal achievable rate. Recall that the capacity region for the broadcast channel is the set of all rate vectors $\mathbf{r} = (r_1 r_2 \ldots r_n)$ that are simultaneously achievable [11]. We will denote the capacity region for an arbitrary precoder under power constraint P by $\mathcal{C}(\mathbf{H}, P)$ (as in (2.2)). It should be clear that this region is convex since we may time share between any two pairs of rates in $\mathcal{C}(\mathbf{H}, P)$.

The problem of finding the point(s) in $\mathcal{C}(\mathbf{H}, P)$ that maximizes the sum rate for the MIMO broadcast channel has been well studied. This point was simultaneously characterized by many authors through the use of convex duality theory [33,34,35,40]. It was long believed that the region known as Marton's achievable region [21] is the capacity region for the Gaussian broadcast channel. Recently, it has been reported that this region is indeed the capacity region for the Gaussian broadcast channel [37]. We will now focus on some of these multiplexing techniques and characterize the capacity regions for some simpler multiplexing techniques. We will focus on the techniques of beamforming, zero-forcing and the multiplexing technique, dirty paper coding, that achieves Marton's region.

■ 2.2.1 Superposition Coding and Beamforming

We begin by focusing on beamforming as a method to combat interference in multi-user systems. In order to provide a desired SIR between users we may use transmit and receive beamforming. We focus on the case where the instantaneous signal $\mathbf{x}$, can be represented as the linear combination

$$\mathbf{x} = \sum_{i \in \mathcal{A}} u_i \mathbf{w}_i = \mathbf{W}\mathbf{u}$$

where u_i is the message symbol for users i and $\mathbf{w}_i$ is the *beamforming* vector. This vector in general may be optimized for each transmission but may also come from some finite code book. If the receiver employs an MMSE receiver to maximize the receive SINR the resulting SINR for the ith user is [36]

$$\text{SINR}_i(\mathcal{A}) = \frac{P_i |\mathbf{h}_i \mathbf{w}_i^\dagger|^2}{1 + \sum_{i \neq j} P_j |\mathbf{h}_i \mathbf{w}_j^\dagger|^2} \tag{2.3}$$

where P_i is the power allocated to user i.

It will be shown that in the case of maximizing $\mathbf{q} \cdot \mathbf{r}$, finding $\zeta_{\text{sp}} = \zeta_{\text{sp}}(\mathcal{A})$ that satisfies, for a fixed choice of $\mathcal{A}$,

$$\sum_{i \in \mathcal{A}} \left(q_i \zeta_{\text{sp}} - \frac{1}{\text{SINR}_i(\mathcal{A})} \right)_+ = P \tag{2.4}$$

where $(x)_+ = \max\{0, x\}$ yields the optimal power allocation. We will let $\mathcal{C}_{\text{sp}}(\mathbf{H}, P)$ be the set of all rate vectors achievable by beamforming under a power constraint P.

Note that in this scenario a user set may contain users that receive zero power and thus zero rate. We will call an arbitrary set of users *valid* if all users of the set are assigned a non-zero rate with regard to a specific rate allocation policy. This distinction is important since in general if an activation set contains a user receiving zero rate we will incur no penalty for considering the subset of users receiving a strictly positive rate. It should be clear from (2.4) that a set $\mathcal{A}$ is valid under beamforming if and only if

$$\frac{1}{\text{SINR}_{i_{\min}}(\mathcal{A})} \leq q_{i_{\min}} \zeta_{\text{sp}}(\mathcal{A}) \tag{2.5}$$

where $i_{\min}(\mathcal{A}) \triangleq \arg\min_{i \in \mathcal{A}} q_i \, \text{SINR}_i(\mathcal{A})$. Thus, under beamforming the collection of valid user sets is precisely

$$\mathcal{S}_{\text{sp}} \triangleq \left\{ \mathcal{A} \;\middle|\; \frac{1}{\text{SINR}_{i_{\min}}(\mathcal{A})} \leq q_{i_{\min}} \zeta_{\text{sp}}(\mathcal{A}) \right\} \tag{2.6}$$

Further, it is clear from (2.11) that for every valid set $\mathcal{A}$ we have

$$\zeta_{\text{sp}}(\mathcal{A}) \triangleq \frac{P + \sum_{i \in \mathcal{A}} \frac{1}{\text{SINR}_i(\mathcal{A})}}{\sum_{i \in \mathcal{A}} q_i} \tag{2.7}$$

The collection $\mathcal{S}_{\text{sp}}$ is sufficiently large so that we can still find a set that achieves the optimal value of (2.2). Additionally, if $\mathcal{A} \in \mathcal{S}_{\text{sp}}$, then the positivity constraint in the water filling equation (2.4) is always satisfied. This yields the following characterization of the weighted sum rate.

Theorem 2.2.1. *Under a total power constraint P, the maximum weighted beamforming sum rate is*

$$\max_{\substack{\mathcal{A} \subset \mathcal{U} \\ \mathbf{r} \in \mathcal{C}_{\text{sp}}(\mathbf{H},\mathbf{q})}} \mathbf{q} \cdot \mathbf{r} = \max_{\mathcal{A} \in \mathcal{S}_{\text{sp}}} \left(\sum_{i \in \mathcal{A}} q_i \right) f_{\text{sp}}(\mathbf{H}_{\mathcal{A}}, \mathbf{q}(\mathcal{A}))$$

where

$$f_{\text{sp}}(\mathbf{H}_{\mathcal{A}}, \mathbf{q}(\mathcal{A})) = \log \left(1 + \frac{P}{|\mathcal{A}|} \mathcal{H}(\{\text{SINR}_i(\mathcal{A})\}) \right) + D(\, q(\mathcal{A}) \parallel \text{SINR}(\mathcal{A})) \qquad (2.8)$$

where $\mathcal{H}$ is the harmonic mean, $D(\cdot \| \cdot)$ is the Kullback Leibler distance and $\text{SINR}(\mathcal{A})$ *and $q(\mathcal{A})$ are the empirical distributions of the* $\text{SINR}_i(\mathcal{A})$ *and q_i restricted to the set $\mathcal{A}$.*

Proof. See Appendix A.1 □

■ 2.2.2 Zero-Forcing Multiplexing

We now examine a special case of beamforming, zero-forcing multiplexing, in which we take the beamforming matrix to be the inverse of the channel matrix. Zero-forcing multiplexing exhibits low complexity user selection since it does not incur a problem of user ordering. This makes it an attractive choice if we have a complexity constraint at the transmitter. Zero-forcing multiplexing is less complex, since it simply inverts the channel at the transmitter by choosing a transmit vector $\mathbf{x} = \mathbf{H}_{\mathcal{A}}^{+}\mathbf{u}$, where $\mathbf{H}_{\mathcal{A}}^{+}$ is the pseudo-inverse of the channel matrix for the active user set $\mathcal{A}$ and $\mathbf{u}$ is the vector of message symbols. We assume throughout that $\mathbf{H}_{\mathcal{A}}$ is non-singular since this occurs with probability 1.

Note that the sub-optimality of zero-forcing appears in the power price paid in inverting the channel. It can be shown that the power constraint becomes,

$$\sum_i P_i / b_i \leq P \qquad (2.9)$$

where

$$b_i(\mathcal{A}) = \frac{1}{\left(\mathbf{W}_{\mathcal{A}}^{-1}\right)_{i,i}} \quad \text{and} \quad \mathbf{W}_{\mathcal{A}} = \mathbf{H}_{\mathcal{A}} \mathbf{H}_{\mathcal{A}}^{\dagger} \qquad (2.10)$$

The b_i's have an important geometric interpretation as noted in [9] as the squared norm of user i's channel when projected away from every other users channel in the activation set $\mathcal{A}$. This suggests that we pay a large price in power if we have users who are nearly collinear.

It is well known (for example see [9]) that water pouring with respect to $1/b_i(\mathcal{A})$, yields the optimal power allocation for the sum rate in the case of single receive antennas. A similar expression can be derived in the case that we consider. That is, in the case of maximizing $\mathbf{q} \cdot \mathbf{r}$ we will show that finding $\zeta_{\text{zf}} = \zeta_{\text{zf}}(\mathcal{A})$ that satisfies, for a fixed set $\mathcal{A}$,

$$\sum_{i \in \mathcal{A}} \left(q_i \zeta_{\text{zf}} - \frac{1}{b_i(\mathcal{A})} \right)_+ = P \tag{2.11}$$

where $(x)_+ = \max\{0, x\}$ yields the optimal power allocation. We will let $\mathcal{C}_{\text{zf}}(\mathbf{H}, P)$ be the set of all rate vectors achievable by zero-forcing multiplexing under a power constraint P. As in the case of beamforming, under zero forcing multiplexing the collection of valid user sets is precisely

$$\mathcal{S}_{\text{zf}} \triangleq \left\{ \mathcal{A} \;\middle|\; \frac{1}{b_{i_{\min}}(\mathcal{A})} \leq q_{i_{\min}} \zeta_{\text{zf}}(\mathcal{A}) \right\} \tag{2.12}$$

where $i_{\min}(\mathcal{A}) \triangleq \arg\min_{i \in \mathcal{A}} q_i \, b_i(\mathcal{A})$. Thus, for every valid set $\mathcal{A}$ we have

$$\zeta_{\text{zf}}(\mathcal{A}) \triangleq \frac{P + \sum_{i \in \mathcal{A}} \frac{1}{b_i(\mathcal{A})}}{\sum_{i \in \mathcal{A}} q_i} \tag{2.13}$$

The collection $\mathcal{S}_{\text{zf}}$ is sufficiently large so that we can still find a set that achieves the optimal value of (2.2). To see this note, in the case $\mathbf{q}$ is the all one vector, the optimal sum rate can be given as [9]

$$R_{\text{sum}}^{\text{zf}}(\mathcal{A}) \triangleq \sum_{i \in \mathcal{A}} \log \left(1 + b_i(\mathcal{A}) \left(\zeta_{\text{zf}}(\mathcal{A}) - \frac{1}{b_i(\mathcal{A})} \right)_+ \right) \tag{2.14}$$

Note that $b_i(\mathcal{A})$ can only increase by removing any user from the set $\mathcal{A}$. That is, for any i and j, $b_i(\mathcal{A}) \leq b_i(\mathcal{A} \setminus \{j\})$ with equality if and only if $\mathbf{h}_j \mathbf{h}_i^\dagger = 0$. Since, $\mathbf{h}_j$ and $\mathbf{h}_i$ are orthogonal with probability zero removing any user receiving zero rate, say user j, will with probability 1 increase the sum rate. This yields the following characterization of the sum rate.

Theorem 2.2.2. *Under a total power constraint P, the maximum weighted zero-forcing sum rate is*

$$\max_{\substack{\mathcal{A}\subset\mathcal{U}\\ \mathbf{r}\in\mathcal{C}_{\mathrm{zf}}(\mathbf{H},\mathbf{q})}} \mathbf{q}\cdot\mathbf{r} = \max_{\mathcal{A}\in\mathcal{S}_{\mathrm{zf}}} \left(\sum_{i\in\mathcal{A}} q_i\right) f_{\mathrm{zf}}(\mathbf{H}_{\mathcal{A}}, \mathbf{q}(\mathcal{A}))$$

where

$$\begin{aligned} f_{\mathrm{zf}}(\mathbf{H}_{\mathcal{A}}, \mathbf{q}(\mathcal{A})) &= \log\left(1 + \frac{P}{|\mathcal{A}|}\mathcal{H}\left(\{b_i(\mathcal{A})\}\right)\right) + D(q(\mathcal{A})\|b_{\mathcal{A}}) \\ &= \log\left(1 + \frac{P}{\mathrm{Tr}(W_{\mathcal{A}}^{-1})}\right) + D(q(\mathcal{A})\|b_{\mathcal{A}}) \end{aligned} \tag{2.15}$$

where $\mathcal{H}$ is the harmonic mean, $D(\cdot\|\cdot)$ is the Kullback Leibler distance and $b_{\mathcal{A}}$ and $q(\mathcal{A})$ are the empirical distributions of the b_i and q_i restricted to the set $\mathcal{A}$.

Proof. See Appendix A.1 □

Theorems 2.2.1 and 2.2.2 provide significant insight into user selection for beamforming and zero-forcing multiplexing. In particular, (2.15) may be broken up into three main components (see Figure 2-3). First, examining the objective function (2.15) one can see that it is almost linear in the sum of the queue states of the active users. The only non-linearity comes from the divergence term. We call this leading linear term the *queueing gain* since it captures the significance of the magnitude of the queue state on the objective function. Further, we define the logarithmic term as the *geometry gain* since it captures the effects the geometry of the channels have on the optimal solution (recall that the b_i can be interpreted as the square of the distance of each user's channel vector from the subspace spanned by the other users' channels [9]). We note that these two terms are "bulk" properties of the vectors of queue and channel states and are decoupled from one another. The inter-dependence of these two vectors is captured in the *pairing gain*. This term reflects the reward one gets for choosing a set of users whose queue state is matched well with its channel.

$$\underbrace{\left(\sum_{i\in\mathcal{A}} q_i\right)}_{\text{queueing gain}} \left(\underbrace{\log\left(1 + \frac{P}{\mathrm{Tr}(W_{\mathcal{A}}^{-1})}\right)}_{\text{geometry gain}} + \underbrace{D(q_{\mathcal{A}}\|b_{\mathcal{A}})}_{\text{pairing gain}} \right)$$

Figure 2-3. The decomposition of the zero-forcing rate expression (2.15) into its various components

Now, provided that we are given the set $\mathcal{S}_{\text{zf}}$, we have a good understanding on how the objective function f_{zf} depends on both the geometry of the channel matrix $\mathbf{H}_{\mathcal{A}}$ and the queue state $\mathbf{q}$. However, $\mathcal{S}_{\text{zf}}$ is a random variable. For the sake of analysis we may assume that a genie has given us the collection $\mathcal{S}_{\text{zf}}$. Then, given that we know the probabilistic structure of $\mathcal{S}_{\text{zf}}$ we may use this to find the expected sum rate. First note that the collection $\mathcal{S}_{\text{zf}}$ is non-empty since all single user sets are valid. Also, we can see that as $P \to 0$ that $\mathcal{S}_{\text{zf}}$ is with probability 1 is the set of single users. Similarly, if $P \to \infty$, and $\mathbf{q}_i < \infty$ for all i, then with probability 1, $\mathcal{S}_{\text{zf}}$ is the collection of all subset with cardinality less than or equal to m. We note this observation in the following:

Theorem 2.2.3. *Let $\mathcal{S}_{\text{zf}}$ be the collection of valid user sets under zero-forcing multiplexing with peak power constraint P. Then, with probability 1*

$$\lim_{P \to 0} \mathcal{S}_{\text{zf}} = [n]^i_<$$

and

$$\lim_{P \to \infty} \mathcal{S}_{\text{zf}} = \bigcup_{i=1}^{m} [n]^i_<$$

where $[n]^m_<$ is the collection of sets of the form $\{a_1, a_2, \ldots, a_m\}$ such that $1 \leq a_1 < a_2 < \cdots < a_m \leq n$.

Determining the optimal set from $\mathcal{S}_{\text{zf}}$ may still be very complex, since the sets in $\mathcal{S}_{\text{zf}}$ do not have an obvious structure. Recall that from (2.10) that $b_i(\mathcal{A})$ is the square of the norm of user i's channel vector $\mathbf{h}_i$ from the span of every other channel in the activation set $\mathcal{A}$. Thus, selecting users that are nearly orthogonal may yield a set whose sum rate is a reasonable approximation to the optimal sum rate obtained by a search over the sets in $\mathcal{S}_{\text{zf}}$. In Chapter 3 we examine a structure that we can characterize and serves as an approximation to the optimal zero-forcing set. In addition, we can exploit this new structure to reduce the complexity of user set selection.

■ 2.2.3 Dirty Paper Precoding

It is well known that knowledge of interference at the transmitter can lead to no loss in capacity [10]. That is, if a component of the noise is known at the transmitter then the channel can have a capacity equivalent to that of the channel where the noise is not present.

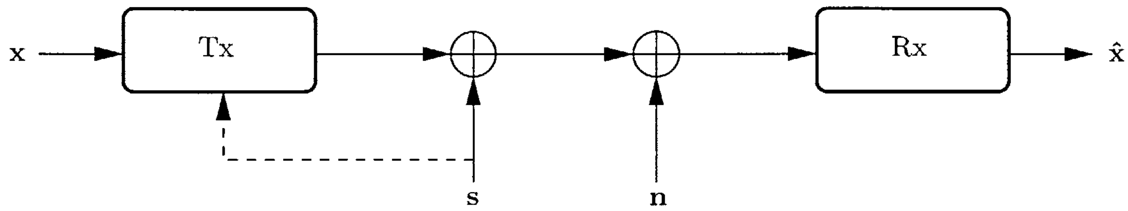

Figure 2-4. The Costa Channel

For example, consider the scalar channel

$$\mathbf{y} = \mathbf{x} + \mathbf{s} + \mathbf{n}$$

where $\mathbf{x}$ is the transmitted signal and $\mathbf{s}$ and $\mathbf{n}$ are independent Gaussian noise. Further, suppose that the noise $\mathbf{s}$ is known only at the transmitter. A depiction of this channel can be seen in Figure 2-4. Then, the capacity of this channel is equivalent to that of the AWGN channel $\mathbf{y} = \mathbf{x} + \mathbf{n}$ [10].

This fact is a bit surprising. We have seen in the broadcast channel that by expending some of the transmit power to cancel part of the known interference we can see some gain in rate (*i.e.* zero forcing multiplexing). However, the result that there is no gap in capacity between the Costa channel and that of the AWGN channel without the noise $\mathbf{s}$ is non-trivial to see. Using our intuition from analogous methods that have been used for inter-symbol interference (ISI) cancellation in the ISI channel and in successive interference cancellation in multi-access channels we can develop a similar method for the broadcast channel. The natural dual of these techniques in the broadcast channel is dirty paper (or Costa) precoding [31].

Recall from (2.1), that we are considering the channel of the form, $\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{n}$ and consider again a transmit signal of the form

$$\mathbf{x} = \sum_{i \in \mathcal{A}} u_i \mathbf{w}_i$$

Then, the output of the channel for the ith user is

$$y_i = u_i \mathbf{h}_i \mathbf{w}_i^\dagger + \sum_{j<i} u_j \mathbf{h}_i \mathbf{w}_j^\dagger + \sum_{j>i} u_j \mathbf{h}_i \mathbf{w}_j^\dagger$$

Now, if we regard the term $\sum_{j>i} u_j \mathbf{h_i w}_j^\dagger$ as interference in the channel of user i, then using Costa's result each user may receives a rate [30],

$$R_i = \log\left(1 + \frac{P_i |\mathbf{h}_i \mathbf{w}_i^\dagger|^2}{1 + \sum_{j>i} P_j |\mathbf{h}_i \mathbf{w}_j^\dagger|^2}\right)$$

However, this expression is sensitive to the *order* in which we encode the users. Thus, it is much more difficult to derive a closed form expression for the boundary of the capacity region as was done in Theorems 2.2.1 and 2.2.2. Furthermore, there is no clear characteristic for a "best" set of users under this multiplexing scheme and it is not clear how one would wish to approximate a solution to the optimization $\mathbf{q} \cdot \mathbf{r}$. However, we do have the following description of the region of achievable rates under dirty paper precoding [34]. Let, $R_i(\pi, \mathbf{P})$ be the rate that user i receives if the encoding is done in the order π using a power allocation $\mathbf{P}$. That is,

$$R_i(\pi, \mathbf{P}) \triangleq \log\left(1 + \frac{P_i |\mathbf{h}_i \mathbf{w}_i^\dagger|^2}{1 + \sum_{\pi(j) > \pi(i)} P_{\pi(j)} |\mathbf{h}_i \mathbf{w}_{\pi(j)}^\dagger|^2}\right)$$

Additionally, let $\mathbf{R}(\pi, \mathbf{P})$ be the vector of all user rates under power allocation $\mathbf{P}$ and ordering π. Then the set of all achievable rates $\mathcal{C}_{\text{dpc}}(\mathbf{H}, P)$, is

$$\mathcal{C}_{\text{dpc}}(\mathbf{H}, P) \triangleq \text{co}\left(\bigcup_{\substack{\mathbf{P}: \sum P_i \leq P \\ \pi}} \mathbf{R}(\pi, \mathbf{P})\right)$$

Optimization over the $\mathcal{C}_{\text{dpc}}(\mathbf{H}, P)$ region can be quite complex, since in general we need to consider all permutations of users. This complexity coupled with the underlying search over user sets can result in an intolerable level of complexity. Due to this inherent complexity of optimal selection using dirty paper coding, we derive a suboptimal selection procedure and characterize its performance with respect to dirty paper coding, zero-forcing and beamforming in the following chapter.

Chapter 3

Near Orthogonal Sets for MIMO Broadcasting and Scheduling

For a given multiplexer it may be prohibitively hard to find the set of users that achieves the maximum throughput if the number of users is much larger than the number of transmit antennas. It is natural to think that this search need not include all possible subsets of users. In this chapter we show that the search can be limited to a smaller subset of users that is likely to obtain a sum rate close to that of the optimal set. Intuitively, every user in the optimal set should have high SNR and low SIR. In this direction, we examine the probability of existence of sets of users that have specified SNR and SIR values.

Note the receive SNR for any user is a function of the transmit covariance, noise variance and norm of the user's channel vector. Thus, selecting a set of users that have large channel gains, to first order, appears to be a good approximation to the optimal set. Indeed, in the case of a single transmit antenna or in the case the transmitter is constrained to choose a single user (as in [15]), not choosing the user with the largest channel gain would result in a loss in the expected rate. However, in the general case of multiple transmit antennas and $|\mathcal{A}| > 1$, a user whose channel norm is much greater than the other users in the set can cause large interference on the other users in that set. For this reason we impose both an upper bound and lower bound on the norm of channels of selected users. To this end, let the set

$$\mathcal{S}_\rho = \left\{ i \in \mathcal{U} \;\middle|\; \rho^- \leq \|\mathbf{h}_i\|^2 \leq \rho^+ \right\} \tag{3.1}$$

be the set of all user who meet this constraint.

We would additionally like to provide guarantees on the SIR experienced by each user. The guarantee on SIR can be achieved by bounding the magnitude of the inner product

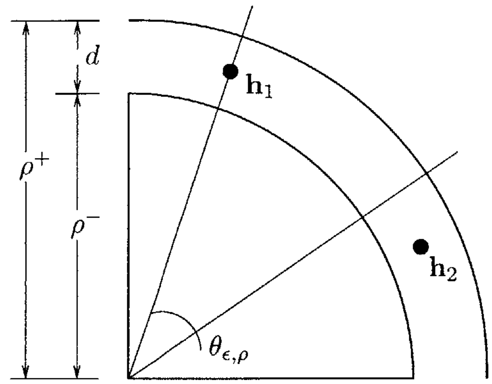

Figure 3-1. The geometry of the near-optimal selection procedure. Any pair of selected points must fall outside a cone of half angle $\theta_{\epsilon,\rho}$ centered at each point and lie inside a spherical shell defined by ρ^- and ρ^+

between any two users. We alternatively could achieve a guarantee on SIR by bounding the average inner product between users in a set. We only examine the question of bounding the pairwise inner product since it yields a simple graph structure that can be used to develop low complexity algorithms. Additionally, this graph structure is useful in the analysis of the existence probability. Now, consider the collection of sets

$$\mathcal{S}_{\epsilon,\rho}^{(l)} = \left\{ \mathcal{A} \,\middle|\, |\mathcal{A}| = l \text{ and } \mathcal{A} \subset \mathcal{S}_\rho \text{ and } |\mathbf{h}_i \mathbf{h}_j^\dagger| \leq \epsilon \,, \forall\, i \neq j \in \mathcal{A} \right\} \tag{3.2}$$

Any set $\mathcal{A} \in \mathcal{S}_{\epsilon,\rho}^{(l)}$ we will call an ϵ-*orthogonal* set. A geometric interpretation is depicted in Figure 3-1. One can think of any ϵ-orthogonal set as a collection of points that lie in the spherical shell between radii ρ^- and ρ^+ such that any two points form an angle no smaller than $\theta_{\epsilon,\rho}$, where

$$\theta_{\epsilon,\rho} \triangleq \cos^{-1}\left(\frac{\epsilon}{\rho^+}\right) \tag{3.3}$$

It is important to understand the probability that $\mathcal{S}_{\epsilon,\rho}^{(l)}$ is non-empty. In this direction, let N_ρ be the number of users that fall in the spherical shell defined by ρ^- and ρ^+. That is,

$$N_\rho \triangleq |\mathcal{S}_\rho| = \sum_{i=1}^{n} 1_{\{i \in \mathcal{S}_\rho\}}$$

Now, since the norms of the channels are independent by assumption, we have that N_ρ is a binomial random variable with parameter p_s, where p_s is the probability that a point falls in the spherical shell defined by the parameters ρ^- and ρ^+. Since our vectors are *i.i.d* complex circularly symmetric Gaussian with variance $1/2m$, then

$$p_s = \Gamma(2m, m\rho^-) - \Gamma(2m, m\rho^+)$$

where $\Gamma(m, x)$ is the incomplete gamma function [5]

$$\Gamma(m, x) = \int_x^{\infty} t^{m-1} e^{-t}\, dt$$

Recall, that the magnitude of the channel vector is closely related to the SNR guarantee that we provide on any selected set. Thus, it is natural to consider the question on the rate we may vary ρ^+ and ρ^- as a function of n and still have $N_\rho \neq 0$ with non-zero probability. This is the content of the following lemma and theorem.

Lemma 3.0.4. *Let* $m\rho^+(n) = c\log(n)$ *and* $m\rho^-(n) = \log(n) - d$ *where* $c \geq 1$ *and* $d > 0$. *Then,*

$$2m\left(e^d - 1\right) \leq \lim_{n\to\infty} np_s \leq 2me^d$$

Further, if $\log n = o(\rho^-(n))$ *then*

$$\lim_{n\to\infty} np_s = 0$$

Proof. See appendix B.1. □

We now use this lemma to characterize the rates at which the channel norms, and thus the SNR, can grow while still having a non-zero probability that the set $\mathcal{S}_\rho$ is non-empty.

Theorem 3.0.5. *We can achieve* $\Pr(N_\rho = 0) = \Theta(n^{-m})$ *with*

$$m\rho^-(n) = \log(n) - \log(\log(n) + 1)$$

Further, if $\rho^-(n) = o(\log(n))$ *then* $\Pr(N_\rho = 0) \to 1$ *as* n *tends to infinity.*

Proof. First note that N_ρ is a Binomial random variable with mean np_s. Thus, by using a Chernoff bound we have that [18]

$$\Pr(N \geq l) \geq 1 - \exp\left(-\frac{(np_s - l)^2}{2np_s}\right)$$

if $l \leq np_s = \mathbb{E}N$. Thus,

$$\log\left(\frac{1}{\Pr(N_\rho = 0)}\right) \geq \frac{(np_s)^2}{2np_s} \tag{3.4}$$

Now applying Lemma 3.0.4, for sufficiently large n

$$\log\left(\frac{1}{\Pr(N_\rho = 0)}\right) \geq m\left(e^d - 1\right)$$

Substituting $d = \log(\log(n) + 1)$ proves that with this choice of ρ^- we have $\Pr(N_\rho = 0) \leq n^{-m} = O(n^{-m})$. We additionally have

$$\Pr(N_\rho = 0) = (1 - p_s)^n \geq \exp\left(-\frac{np_s}{1 - p_s}\right)$$

using the inequality $1 - x \geq \exp \frac{-x}{1-x}$ [18]. Repeating the argument above, $\Pr(N_\rho = 0)$ is easily shown to be $\Omega(n^{-m})$.

$\square$

We wish to similarly characterize the probability that the set $\mathcal{S}^{(l)}_{\epsilon,\rho}$ is non-empty. Recall that ϵ is simply the constraint that we place on the pairwise SIR. Define $N^{(l)}_\epsilon$ to be the number of sets in $\mathcal{S}^{(l)}_{\epsilon,\rho}$. That is,

$$N^{(l)}_\epsilon \triangleq |\mathcal{S}^{(l)}_{\epsilon,\rho}| = \sum_{\substack{\mathcal{A} \subset \mathcal{S}_\rho \\ |\mathcal{A}| = l}} 1_{\{\mathcal{A} \in \mathcal{S}_{\epsilon,\rho}\}} \tag{3.5}$$

Note that the indicator random variables in the sum are not independent. Indeed, if $\mathcal{A} \cap \mathcal{B} \neq \emptyset$ the random variables $1_{\{\mathcal{A} \in \mathcal{S}_{\epsilon,\rho}\}}$ and $1_{\{\mathcal{B} \in \mathcal{S}_{\epsilon,\rho}\}}$ are dependent. We will provide good bounds on the probability that $N^{(l)}_\epsilon \neq 0$ given $N_\rho = j$ in Section 3.1, but for now note that it is non-zero if $k \geq l$ and $\epsilon > 0$. Thus we may, by conditioning on the number of points that fall in the spherical shell, write

$$\Pr(N^{(l)}_\epsilon > 0) = \sum_{j=l}^{n} \Pr\left(N_\rho = j\right) \Pr\left(N^{(l)}_\epsilon > 0 \mid N_\rho = j\right) \tag{3.6}$$

We now turn to the question of bounding the conditional probability that there exists an ϵ-orthogonal set, given that $N_\rho = j$.

■ 3.1 Random Packing and Interference

In what follows, we derive bounds on the probability that there exists a set that is ϵ-orthogonal given that $N_\rho = j$ via an interpretation of a random packing of the complex unit m-sphere. We begin by first examining the probability that any set is ϵ-orthogonal using this interpretation. Then, we use this to further bound the probability of existence of an ϵ-orthogonal set in a pool of n independent channel vectors. We continue to use the geometry that is depicted in Figure 3-1.

Begin by defining for any set, say $\mathcal{A}_l$, of size l

$$p_{\epsilon,l} = \Pr\left(\mathcal{A}_l \in \mathcal{S}_{\epsilon,\rho}\right) \tag{3.7}$$

to be the probability that the set is ϵ-orthogonal. We will omit the subscript l when it is clear in the context it is used. Note that we can rewrite p_ϵ as $p_\epsilon = p_s^l \, p_\perp$ where $p_\perp$ is the conditional probability that the set is ϵ-orthogonal given that all the points in the set are inside the spherical shell. We can lower bound $p_\perp$ by pessimistically assuming all channels to have norm ρ^+, yielding

$$p_\perp \geq \Pr\left(|\tilde{\mathbf{h}}_j \tilde{\mathbf{h}}_i^\dagger| \leq \frac{\epsilon}{\rho^+} \, \forall \, i,j \; i \neq j \,\middle|\, \mathcal{A} \subset \mathcal{S}_\rho \right) \tag{3.8}$$

where $\tilde{\mathbf{h}}_i = \mathbf{h}_i / \|\mathbf{h}_i\|^2$.

In our approach to further bound (3.8) it will be useful to consider the set of points which violate the constraint $|\tilde{\mathbf{h}}_j \tilde{\mathbf{h}}_i^\dagger| \leq \frac{\epsilon}{\rho^+}$ of the lower bound on $p_\perp$ for both a pair of users and for a set of users (*i.e.* sets that contains all points which violate the SIR guarantee). Will denote these sets of points $c_m(\tilde{\mathbf{h}};\theta)$ and $c_m(\mathcal{A};\theta)$ respectively. More formally, we may write these sets as

$$c_m(\tilde{\mathbf{h}};\theta) \triangleq \left\{ \tilde{\mathbf{x}} \in \mathbb{C}^m \,\middle|\, \|\tilde{\mathbf{x}}\|^2 = 1 \quad \text{and} \quad |\tilde{\mathbf{x}}\tilde{\mathbf{h}}^\dagger| > \cos\theta \right\} \tag{3.9}$$

and

$$c_m(\mathcal{A};\theta) \triangleq \bigcup_{i \in \mathcal{A}} c_m(\tilde{\mathbf{h}}_i;\theta) \tag{3.10}$$

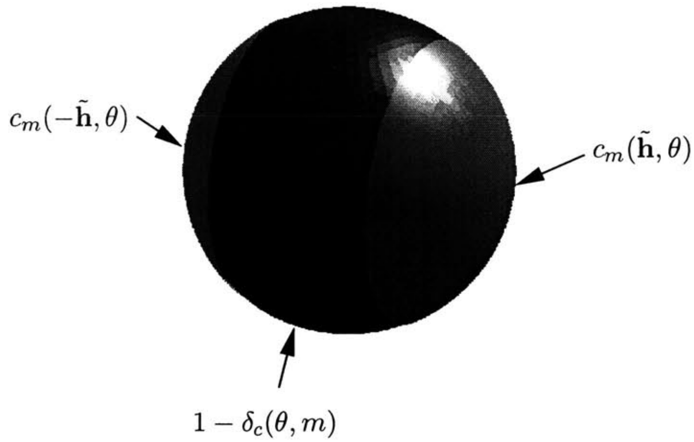

Figure 3-2. An interpretation of the probability that any two users meet a SIR constraint. In this example any user is ϵ-orthogonal with user 1 if either channel vector falls in the region between the two caps $c_m(-\tilde{\mathbf{h}}, \theta)$ and $c_m(\tilde{\mathbf{h}}, \theta)$

Now, note that if $\tilde{\mathbf{h}}_j \notin c_m(\tilde{\mathbf{h}}_i; \theta_{\epsilon,\rho})$ then, $|\tilde{\mathbf{h}}_j \tilde{\mathbf{h}}_i^\dagger| \leq \cos\theta = \frac{\epsilon}{\rho^+}$. So that,

$$|\mathbf{h}_j \mathbf{h}_i^\dagger| \leq \epsilon \, \frac{\|\mathbf{h}_i\| \|\mathbf{h}_j\|}{\rho^+} \leq \epsilon$$

Thus, the RHS of (3.8) is a lower bound to $p_\perp$.

We will let $\delta_c(\theta, 2m)$ be the probability that a point chosen at random uniformly on the surface of the complex unit m-sphere lies in the set $c_m(\tilde{\mathbf{h}}; \theta)$. That is,

$$\delta_c(\theta, 2m) \triangleq \mu_c\left(c_m(\tilde{\mathbf{h}}; \theta)\right)$$

where μ_c is the uniform measure on the complex unit m-sphere. It is well known that this measure can be obtained from the uniform measure on the real unit $2m$-sphere. This can be done by regarding the complex sphere as a surface in $\mathbb{R}^{2m}$ and noting that the function $|\tilde{\mathbf{x}}\tilde{\mathbf{y}}^T|$ is unmodified if either $\tilde{\mathbf{x}}$ or $\tilde{\mathbf{y}}$ is multiplied by -1 [19]. Thus,

$$\delta_c(\theta, 2m) = 2\frac{\Omega_{2m}(\theta)}{\Omega_{2m}(\pi)} \tag{3.11}$$

where $\Omega_{2m}(\theta)$ is the area of the real unit $2m$-sphere cut out by a cone of half angle θ. A geometric interpretation of $\delta_c(\theta, 2m)$ can be seen in Figure 3-2. It is simply the fraction of the real $2m$-sphere covered by antipodal caps.

Obtaining the probability that any point falls in the set $c_m(\mathcal{A}; \theta)$ requires a bit more care.

We may, in general, still interpret the probability as the fraction of the the area covered by the caps centered at the points in $\mathcal{A}$. However, we must take care in appropriately accounting for the regions of intersection between caps. In addition, if the set of points $\mathcal{A}$ is chosen at random, the fraction of area covered by the set $\mathcal{A}$ is a random variable. We will let the fractional area covered by any set $\mathcal{A}$ as,

$$F_c(\mathcal{A}, \theta) \triangleq \mu_c\left(c_m(\mathcal{A}; \theta)\right) \tag{3.12}$$

Note, that this problem is similar to the problem of random packing studied in [38] and is an example of the method of random codes due to Shannon [25]. We will require a stronger result for our pairwise SIR guarantee. We briefly review these results since they will motivate a set of more general questions that provide valuable insights into the problem of user selection and serve as a useful example of the method that we will follow. Begin by considering the random variable,

$$g_i(\mathcal{A}, \theta) = \begin{cases} 1 \quad \text{if} \quad |\tilde{\mathbf{h}}_i \tilde{\mathbf{h}}_j^\dagger| < \cos\theta \quad \forall\, j \in \mathcal{A}\; j \neq i \\ 0 \qquad\qquad\qquad \text{otherwise} \end{cases} \tag{3.13}$$

That is, $g_i(\mathcal{A}, \theta)$ is the indicator random variable of the event that all points in $\mathcal{A}$ fall outside the cap of half angle θ centered at $\tilde{\mathbf{h}}_i$ on the complex unit m-sphere. Now, consider the random variable

$$F_p(\mathcal{A}, \theta) = \frac{1}{|\mathcal{A}|} \sum_{j \in \mathcal{A}} g_i(\mathcal{A})$$

More concretely, F_p is the fraction of caps of half angle $\frac{\theta}{2}$ centered at the points in the set $\mathcal{A}$ that do not intersect any cap in the set. It is important to note the distinction between $F_p(\mathcal{A}, \theta)$ and $F_c(\mathcal{A}, \theta)$. Note that the event $F_p(\mathcal{A}, \theta) = 1$ can be used to lower bound the probability the set $\mathcal{A}$ meets the SIR guarantee. That is,

$$\Pr(\mathcal{A} \in \mathcal{S}_{\epsilon,\rho}) \leq \Pr\left(F_p(\mathcal{A}, \theta) = 1\right)$$

Now, when $\mathcal{A} \in \mathcal{S}_{\epsilon,\rho}$, $F_c(\mathcal{A}, \theta)$ can be interpreted as the probability that any point chosen uniformly on the sphere will lead to a packing. More precisely, for any j,

$$F_c(\mathcal{A} \in \mathcal{S}_{\epsilon,\rho}, \theta) = \Pr(F_p(\mathcal{A} \cup \{j\}, \theta) = 1 \mid \mathcal{A} \in \mathcal{S}_{\epsilon,\rho})$$

In particular we have that $F_c(\{i\},\theta) = \delta(\theta, 2m)$ which is just the probability that any pair meets the SIR guarantee. In this direction we have the following result as a direct consequence of [38].

Theorem 3.1.1. *Let θ be fixed and consider an experiment in which the l members of a set, say $\mathcal{A}$, are chosen at random with a uniform distribution from the complex unit m-sphere. Then,*

$$\mathbb{E}\ F_p(\mathcal{A},\theta) = (1-\delta_c(\theta,2m))^{l-1}$$

Furthermore, let $l(m)$ be the cardinality of a sequence of sets, say $\mathcal{A}(m)$. Then, as $m \to \infty$

$$\begin{array}{llll} \text{if} & l(m)\delta_c(\theta,2m) \to \infty & \text{then} & \mathbb{E}\ F_p(\mathcal{A}(m),\theta) \to 0 \\ \text{if} & l(m)\delta_c(\theta,2m) \to 0 & \text{then} & \mathbb{E}\ F_p(\mathcal{A}(m),\theta) \to 1 \end{array}$$

Proof. We follow the proof given in [38]. First, note that

$$\mathbb{E}g_i(\mathcal{A},\theta) = \Pr(g_i = 1) = \Pr(\tilde{\mathbf{h}}_i\tilde{\mathbf{h}}_j^\dagger < \cos\theta\ \forall i \neq j \in \mathcal{A})$$

Now, since the points were drawn independently this probability is just the product of the pairwise probabilities. Thus,

$$\mathbb{E}g_i(\mathcal{A},\theta) = (1-\delta_c(\theta,2m))^{l-1}$$

By symmetry

$$\mathbb{E}F_p(\mathcal{A},\theta) = (1-\delta_c(\theta,2m))^{l-1} = \left(1 - \frac{1}{\delta_c(\theta,2m)^{-1}}\right)^{\frac{(l-1)\delta_c(\theta,2m)^{-1}}{\delta_c(\theta,2m)^{-1}}}$$

Now since $\lim_{m\to\infty} \delta_c(\theta,2m)^{-1} \to \infty$

$$\left(1 - \frac{1}{\delta_c(\theta,2m)^{-1}}\right)^{\delta_c(\theta,2m)^{-1}} \sim \exp(-1)$$

and

$$\mathbb{E}F_p(\mathcal{A},\theta) \sim \exp\left(-(l-1)\delta_c(\theta,2m)\right)$$

□

Note that the previous theorem answers the question that any set is ϵ-orthogonal asymptotically as a function of the complex dimension m. We wish to use a random packing argument as a way to characterize the probability that there are sets with a prescribed SIR guarantee in finite dimensions. However, we are not primarily concerned in the probability that any particular set meets the SIR guarantee as a function of the dimension, but rather we would like to find the probability that there exists a small set in some larger collection of sets that meets this guarantee. More precisely, we are interested in the random variable

$$1_{\{\mathcal{A}\in\mathcal{S}_{\epsilon,\rho}\}} = \prod_{i\in\mathcal{A}} g_i(\mathcal{A},\theta) \tag{3.14}$$

and the existence of a threshold in the number of points we place on the sphere. This yields the more general question.

Question 1. *Given a set of size n, say $\mathcal{U}$, what is the probability that there is a subset, $\mathcal{A} \subset \mathcal{U}$ of size l such that $F_p(\mathcal{A},\theta) = 1$ as a function of n, l and θ? Further, is there a threshold in the probability that there exists a set $\mathcal{A}$ such that $F_p(\mathcal{A},\theta) = 1$ in n?*

We will answer this question through our discussion in Section 3.2. We will show that for any $p_s > 0$ and $p_\perp > 0$ the probability $\Pr(\exists\ \mathcal{A}$ s.t. $F_p(\mathcal{A},\theta) = 1)$ tends to one as n tends to infinity. In order to find a lower bound to this probability let,

$$E(p,l) \triangleq \max\left\{\frac{2p^2}{l}, \frac{8p}{25l}\left(\frac{l-1}{el}\right)^{l-1}\right\} \tag{3.15}$$

Recall that N_ρ is defined to be the random variable that counts the number of channel vectors that fall in the spherical shell defined by the radii ρ^+ and ρ^- and is binomially distributed with parameter p_s. Then we provide the following theorem in Section 3.3.

Theorem 3.2.3 *Let $\epsilon, \rho^-, \rho^+ \in \mathbb{R}^+$. Then if $\epsilon \le \rho^+$*

$$\Pr(N_\epsilon^{(l)} > 0) \ge \Pr\left(N_\rho \ge l\right) - c_1\left(1 + p_s\left(e^{-E(p_\perp,l)} - 1\right)\right)^n \tag{3.16}$$

where $c_1 = c_1(p_\perp, l) = \exp\frac{2p_\perp^2(l-1)}{l}$ if $E(p_\perp,l) = \frac{2p_\perp^2}{l}$ and $c_1 = 1$ otherwise.

Note, that Theorem 3.2.3 provides bounds to the probability that an ϵ-orthogonal set exists given that we can characterize the probabilities p_s and $p_\perp$. Thus, we must still find good bounds on $p_\perp$ in order to completely characterize the probability that there is a set of users that meet or SIR constraint. We use the method of exhaustion used in [25] to lower bound $p_\perp = \Pr(\mathcal{A} \in \mathcal{S}_{\epsilon,\rho})$. That is, we lower bound this probability by successively placing the points in $\mathcal{A}$ on the sphere and marking the cap about each point as invalid.

Begin by placing any point, say $\tilde{\mathbf{h}}_{a_1} \in \mathcal{A}$, on the sphere. We can do this with probability 1. Now, pick a second point from $\mathcal{A}$, say $\tilde{\mathbf{h}}_{a_2}$, with $a_2 \neq a_1$. These two points are ϵ-orthogonal with probability $\delta_c(\theta, 2m)$. Continuing this procedure the ith point can be successful placed with probability

$$\mathbb{E}\left[F_c(\{\tilde{\mathbf{h}}_{a_j}\}_{j<i}, \theta) \;\middle|\; \{\tilde{\mathbf{h}}_{a_j}\}_{j<i} \in \mathcal{S}_{\epsilon,\rho}^{(i-1)} \right]$$

Now, the set is ϵ-orthogonal if every point falls outside the spherical caps about all previously placed points. This leads to the last question we must answer in order to bound the the probability $\Pr(N_\epsilon^{(l)} \neq 0)$.

Question 2. *Given that $\mathcal{A} \in \mathcal{S}_{\epsilon,\rho}$ how does $\mathbb{E}F_c(\mathcal{A},\theta)$ vary as a function of θ and $|\mathcal{A}|$?*

Through the application of the union bound and new estimates on the surface area of spherical caps provided in Theorem 3.3.2, we have the following lower bound.

Theorem 3.1.2. *Let $\epsilon < \rho^- < \rho^+$ be given. Then, for any set $\mathcal{A}$, we have*

$$\mathbb{E}F_c(\mathcal{A},\theta) \geq 1 - |\mathcal{A}|(1-\cos\theta_{\epsilon,\rho})^{\psi_{m-2}^{-1}}(\sin\theta_{\epsilon,\rho})^{\psi_{m-2}^{-2}-\psi_{m-2}^{-1}}$$

Furthermore,

$$p_\perp \geq (\mathbb{E}F_c(\mathcal{A}\setminus\{1\},\theta))^{|\mathcal{A}|-1}$$

where $\psi_m \triangleq \frac{\sqrt{\pi}}{m}\frac{\Gamma(\frac{m+1}{2})}{\Gamma(\frac{m}{2})}$

The remainder of this chapter is organized as follows. We first provide a framework to answer Question 1 in Section 3.2 using the theory of random geometric graphs. Then, we provide answers to Question 2 by first providing a new bound on the content of spherical caps in Section 3.3. These bounds are then used to derive lower bounds on the multiplexing rate in a channel with choice in Section 3.4.

■ 3.2 Existence Probabilities and Dependent Sums of Random Variables

We now consider a general method to provide and answer to our first question. More precisely, we consider the problem of the sum of dependent indicator random variables that appear in the sum of (3.5). This is so that we may bound the probability of existence of an ϵ-orthogonal set. We will make use of the theory of a *random geometric graphs* and the idea of dependency graphs in order to bound the probability of existence. An introduction to graph theory and the theory of random graphs can be found in [7, 14] and [8, 18, 23], respectively.

Recall that our SIR guarantee was defined in terms of a *pairwise* guarantee ϵ. Thus, for any two points in our user pool $\mathcal{U}$, say i and j, it is natural to consider the indicator random variable $1_{(i,j)}$ that is 1 if users i and j meet the SIR specification. Recall that geometrically this occurs if the point $\tilde{\mathbf{h}}_j$ falls outside the cap of half angle $\theta_{\epsilon,\rho}$ centered at $\tilde{\mathbf{h}}_i$. That is,

$$1_{(i,j)} = \begin{cases} 1 & \text{if } |\tilde{\mathbf{h}}_j \tilde{\mathbf{h}}_i^\dagger| \leq \cos\theta_{\epsilon,\rho} \\ 0 & \text{if } |\tilde{\mathbf{h}}_j \tilde{\mathbf{h}}_i^\dagger| \geq \cos\theta_{\epsilon,\rho} \end{cases} \tag{3.17}$$

Since the random variable $1_{(i,j)}$ depends only on the users i and j we can think of the random variables $1_{(i,j)}$ as simply the edges in some random graph. More formally, we consider the random graph $G_\epsilon(\mathbf{HH}^\dagger)$ defined with vertex set $\mathcal{U}$ and edge set $\{(i,j) : 1_{(i,j)} = 1\}$. This relation is geometrically depicted in Figure 3-3.

It is instructive to consider how one may form this graph in practice. We start by eliminating the users who do not meet the restriction on channel norms. We can do this by first forming the Wishart matrix $\mathbf{W} = \mathbf{HH}^\dagger$ and eliminate the *ith* row and column if $\mathbf{w}_{ii}$ does not satisfy $\rho^- \leq \mathbf{w}_{ii} \leq \rho^+$. We will call this reduced matrix, scaled by $\frac{1}{\rho^+}$, $\widehat{\mathbf{W}}$. We can then check the inner product constraint by examining the entries of the matrix $\widehat{\mathbf{W}}$ that appear above the diagonal and replace the entry $\widehat{\mathbf{w}}_{ij}$ with 1 if the inequality $|\widehat{\mathbf{w}}_{ij}| < \cos\theta_{\epsilon,\rho}$ is satisfied and 0 otherwise. Upon completion we have a $(0,1)$-matrix that is simply the adjacency matrix, $A(G_\epsilon(\mathbf{HH}^\dagger))$, for our graph $G_\epsilon(\mathbf{HH}^\dagger)$. An example of this process can be seen in Figure 3-4.

We now discuss how this construction of $G_\epsilon(\mathbf{HH}^\dagger)$ relates to the problem as we have studied it thus far. Note that the first stage of the construction simply reduce the problem to the set of users in $\mathcal{S}_\rho$. That is $\widehat{\mathbf{W}} = \frac{1}{\rho^+}\mathbf{W}_{\mathcal{S}_\rho}$. We also have a simple interpretation of the

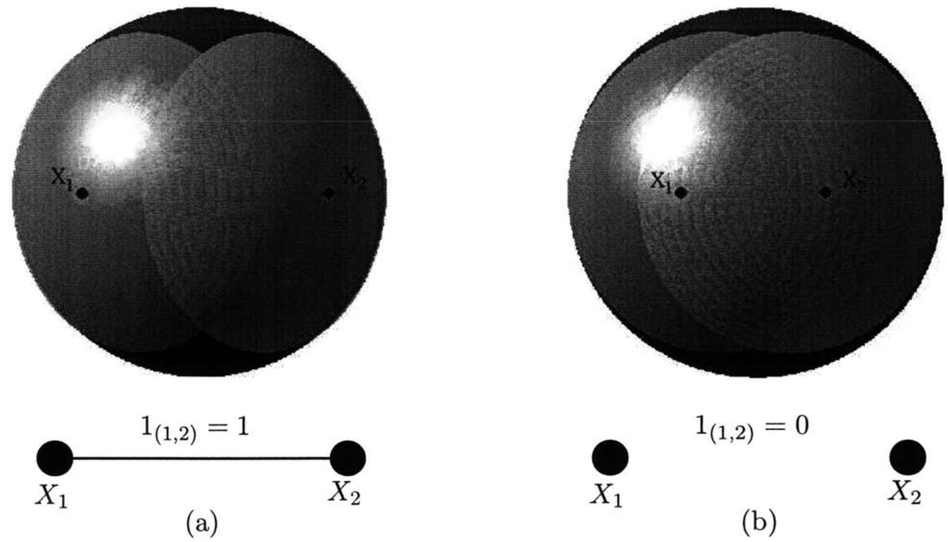

Figure 3-3. The geometric interpretation of the existence of an edge in the graph $G_\epsilon(\mathbf{H}\mathbf{H}^\dagger)$. An edge exists between any two points if each point falls outside the area covered by the cap of half angle $\theta_{\epsilon,\rho}$

random variables $g_i(\mathcal{A})$ and $1_{\{\mathcal{A}\in\mathcal{S}_\rho\}}$. Note, that we can write $g_i(\mathcal{A})$ (and thus $1_{\{\mathcal{A}\in\mathcal{S}_{\epsilon,\rho}\}}$ by (3.14)) in terms of the indicator random variables $1_{(i,j)}$ for $i, j \in \mathcal{A}$. That is,

$$g_i(\mathcal{A}) = \prod_{\substack{j\in\mathcal{A}\\ j\neq i}} 1_{(i,j)} \quad \text{and} \quad 1_{\{\mathcal{A}\in\mathcal{S}_{\epsilon,\rho}\}} = \prod_{\substack{i,j\in\mathcal{A}\\ i<j}} 1_{(i,j)}$$

Thus, $g_i(\mathcal{A})$ is one if the row corresponding to i in $A(G_\epsilon(\mathbf{H}_\mathcal{A}\mathbf{H}_\mathcal{A}^\dagger))$ is all ones off the diagonal. Similarly, $1_{\{\mathcal{A}\in\mathcal{S}_{\epsilon,\rho}\}}$ is one if $A(G_\epsilon(\mathbf{H}_\mathcal{A}\mathbf{H}_\mathcal{A}^\dagger))$ is all ones off the diagonal. Recall a graph is called *complete* if every pair of vertices are adjacent. We will denote the compete graph on l vertices as K_l and write $K_l \subset G$ if there is a copy of K_l in the graph G. Thus, $N_\epsilon^{(l)}$ simply counts the number of copies of K_l. This interpretation can be seen in Figure 3-5. This yields the following lemma.

Lemma 3.2.1. *Let* $\mathbf{H}$ *be given. Then,* $N_\epsilon^{(l)} \neq 0$ *if* $K_l \subset G_\epsilon(\mathbf{H}\mathbf{H}^\dagger)$.

The problem of determining the probability that a Bernoulli random graph contains a copy of some given graph has been well studied [17,18,8]. Note however, the graph $G_\epsilon(\mathbf{H}\mathbf{H}^\dagger)$ is not a Bernoulli random graph since the edges are not independent due to the geometric nature of the problem. For example, note that any two edges are independent, however the edges (i, j) and (j, k) are not independent of (i, k). This dependence structure can be

$$\underset{\mathbf{H}}{\begin{bmatrix} 0.07 & 0.01 & 0.21 \\ 0.00 & -0.40 & -0.38 \\ -0.33 & -0.01 & -0.15 \\ -0.31 & -0.37 & -0.08 \\ -0.12 & -0.44 & -0.40 \\ -0.39 & -0.08 & -0.43 \\ -0.35 & 0.31 & 0.31 \\ 0.49 & 0.04 & 0.00 \end{bmatrix}} \Rightarrow \underset{\mathbf{W}}{\begin{bmatrix} 0.05 & -0.09 & -0.06 & -0.05 & -0.10 & -0.12 & 0.04 & 0.03 \\ -0.09 & 0.31 & 0.06 & 0.18 & 0.33 & 0.20 & -0.25 & -0.01 \\ -0.06 & 0.06 & 0.13 & 0.12 & 0.11 & 0.20 & 0.06 & -0.16 \\ -0.05 & 0.18 & 0.12 & 0.24 & 0.24 & 0.19 & -0.03 & -0.17 \\ -0.10 & 0.33 & 0.11 & 0.24 & 0.38 & 0.26 & -0.22 & -0.08 \\ -0.12 & 0.20 & 0.20 & 0.19 & 0.26 & 0.35 & -0.02 & -0.19 \\ 0.04 & -0.25 & 0.06 & -0.03 & -0.22 & -0.02 & 0.32 & -0.15 \\ 0.03 & -0.01 & -0.16 & -0.17 & -0.08 & -0.19 & -0.15 & 0.24 \end{bmatrix}}$$

$$\Downarrow$$

$$\underset{A(G_\epsilon(\mathbf{HH}^\dagger))}{\begin{bmatrix} 0 & 1 & 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 & 0 & 1 \\ 1 & 1 & 1 & 1 & 1 & 0 \end{bmatrix}} \Leftarrow \underset{\widehat{\mathbf{W}}}{\begin{bmatrix} 0.31 & 0.18 & 0.33 & 0.20 & 0.25 & 0.01 \\ 0.18 & 0.24 & 0.24 & 0.19 & 0.03 & 0.17 \\ 0.33 & 0.24 & 0.38 & 0.26 & 0.22 & 0.08 \\ 0.20 & 0.19 & 0.26 & 0.35 & 0.02 & 0.19 \\ 0.25 & 0.03 & 0.22 & 0.02 & 0.32 & 0.15 \\ 0.01 & 0.17 & 0.08 & 0.19 & 0.15 & 0.24 \end{bmatrix}}$$

Figure 3-4. A example construction of the graph $G_\epsilon(\mathbf{HH}^\dagger)$ for $\rho^- = 0.2, \rho^+ = 1$ and $\epsilon = 0.2$

analyzed using the theory of dependency graphs. We say that a graph G with vertex set $V = V(G)$ is a *dependency graph* of the family of random variables $\{1_\mathcal{A}\}_{\mathcal{A}\in V}$ if for any two disjoint subsets of V, say $A, B \subset V$, the two sub-families $\{1_\mathcal{A}\}_{\mathcal{A}\in A}$ and $\{1_\mathcal{A}\}_{\mathcal{A}\in B}$ are independent. Since our channel vectors are *i.i.d.* it should be clear that any two collections of random variables formed from two disjoint subsets of users are independent. Clearly our independent Gaussian vectors fit this criterion. We will require the following proposition.

Proposition 3.2.2. *[16, Thrm. 2.1] Let $\mathcal{P}_l(\mathcal{U})$ be the collection of all unordered sets of size l on n items and let*

$$X = \sum_{\mathcal{A}\in\mathcal{P}_l(\mathcal{U})} 1_\mathcal{A} \tag{3.18}$$

where $\{1_\mathcal{A}\}$ is a family of Bernoulli random variables with $\Pr(1_\mathcal{A} = 1) = p$*, which are independent if $\mathcal{A} \cap \mathcal{B} = \emptyset$. Then,*

$$\Pr(X = 0) \le \exp\left(-\max\left\{2p^2\left\lfloor\frac{n}{l}\right\rfloor, \frac{8p}{25}\frac{\binom{n}{l}}{\binom{n}{l-1}}\right\}\right) \tag{3.19}$$

Using this theory we can address the probability of existence of an ϵ-orthogonal set. Indeed, if we are given that N_ρ users fall in the spherical shell defined by ρ^- and ρ^+ we can use the previous theorem to bound the probability that there is an ϵ-orthogonal set among these users. First, we take a slightly different exponent than the one in Theorem 3.2.2 so

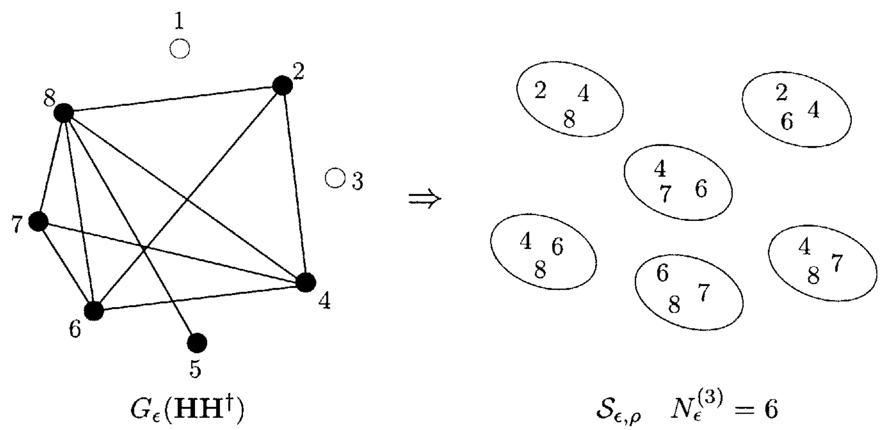

Figure 3-5. The graph $G_\epsilon(\mathbf{HH}^\dagger)$ of Figure 3-4 and the corresponding collection of sets $\mathcal{S}_{\epsilon,\rho}$

that the bound is continuous. Let,

$$E(p,l) \triangleq \max\left\{\frac{2p^2}{l}, \frac{8p}{25l}\left(\frac{l-1}{el}\right)^{l-1}\right\} \tag{3.20}$$

Then as a simple consequence of (3.6) we have the following theorem.

Theorem 3.2.3. *Let $\epsilon, \rho^-, \rho^+ \in \mathbb{R}^+$. Then if $\epsilon \le \rho^+$*

$$\Pr(N_\epsilon^{(l)} > 0) \ge \Pr\left(N_\rho \ge l\right) - c_1\left(1 + p_s\left(e^{-E(p_\perp,l)} - 1\right)\right)^n \tag{3.21}$$

where $c_1 = c_1(p_\perp, l) = \exp\frac{2p_\perp^2(l-1)}{l}$ if $E(p_\perp, l) = \frac{2p_\perp^2}{l}$ and $c_1 = 1$ otherwise.

Proof. Note by conditioning on the number of users that fall in the spherical shell defined by ρ^- and ρ^+ we have,

$$\begin{aligned}
\Pr(|\mathcal{S}_\epsilon| > 0) &= \sum_{j=l}^{n} \Pr\left(N_\rho = j\right)\Pr\left(X_l > 0 | N_\rho = j\right) \\
&> \sum_{j=l}^{n} \binom{n}{j} p_s^j (1-p_s)^{n-j}\left(1 - c_1 e^{-jE(p_\perp,l)}\right) \\
&= \Pr\left(N_\rho \ge l\right) - c_1 \sum_{j=l}^{n} \binom{n}{j}\left(p_s e^{-E(p_\perp,l)}\right)^j (1-p_s)^{n-j} \\
&> \Pr\left(N_\rho \ge l\right) - c_1\left(p_s e^{-E(p_\perp,l)} + (1-p_s)\right)^n
\end{aligned}$$

where $c_1 = c_1(p_\perp, \theta)$ appears by bounding $\lfloor \frac{n}{l} \rfloor$ by $\frac{n}{l} - \frac{l-1}{l}$ and is constant in n.

□

Corollary 3.2.4. *Let* $\rho^-(n) = \log n$ *and* $\rho^+(n) = \log n + \log(\log(n) + 1)$. *Then for any fixed* $p_\perp$ *we have that* $\Pr(|\mathcal{S}_{\epsilon,\rho}| = 0) = O(n^{-\gamma m})$ *where* $\gamma = 1 - e^{-E(p_\perp, m)}$. *Furthermore, if*

$$E(p_\perp, m) = \log\left(\frac{m \log n}{m \log n - \log\log n}\right) \tag{3.22}$$

then, $\Pr(|\mathcal{S}_{\epsilon,\rho}| = 0) = O(1/\log n)$

Proof. Similar to the proof of Theorem 3.0.5 we can use a Chernoff bound to bound the probability that $N_\rho > l$. Thus, (3.21) becomes

$$1 - \Pr(|\mathcal{S}_{\epsilon,\rho}| = 0) \geq 1 - \exp\left(-\frac{(np_s - m)^2}{np_s}\right) - \left(1 + p_s\left(e^{(-E(p_\perp,m))} - 1\right)\right)^n$$

So, for sufficiently large n we have $np_s \sim m\log(n)$ and we have

$$\begin{aligned}
\Pr(|\mathcal{S}_{\epsilon,\rho}| = 0) &\leq \Theta(n^{-m}) + \left(1 + \frac{m\log(n)}{n}\left(e^{(-E(p_\perp,m))} - 1\right)\right)^n \\
&\leq \Theta(n^{-m}) + \exp\left(m\log(n)\left(e^{(-E(p_\perp,m))} - 1\right)\right) \qquad (3.23)\\
&= O\left(n^{-km}\right)
\end{aligned}$$

To see the final statement note that by applying (3.22) to (3.23) we have that

$$\begin{aligned}
\Pr(|\mathcal{S}_{\epsilon,\rho}| = 0) &\leq \Theta(n^{-m}) + \exp\left(m\log(n)\left(\frac{m\log n - \log\log n}{m\log n} - 1\right)\right) \qquad (3.24)\\
&\leq \Theta(n^{-m}) + \frac{1}{\log n} = O\left(\frac{1}{\log n}\right) \qquad (3.25)
\end{aligned}$$

□

We now have an explicit expression for the probability of existence as a function of p_s and $p_\perp$. There is, however, still the problem of finding a simple expression for the probability $\delta_c(\theta, 2m)$ and the probability $p_\perp$. In the following sections we present bounds on these two probabilities.

3.3 New Bounds on the Content of Spherical Caps

In this section we derive the probability that a point chosen uniformly at random falls in a cap of half angle θ. Recall, that we previously defined this probability as

$$\delta_c(\theta, 2m) \triangleq 2\frac{\Omega_{2m}(\theta)}{\Omega_{2m}(\pi)}$$

where $\Omega_{2m}(\theta)$ is the content of the spherical cap of half angle θ in $\mathbb{R}^{2m}$. In [25], Shannon provides the following bounds on the probability $\delta_c(\theta, m)$

$$\frac{1}{\psi_{m-2}}\frac{\sin^{m-1}\theta}{\cos\theta}\left(1 - \frac{\tan^2\theta}{m}\right) \leq \delta_c(\theta, m) \leq \frac{1}{\psi_{m-2}}\frac{\sin^{m-1}\theta}{\cos\theta} \tag{3.26}$$

where

$$\psi_m \triangleq \frac{\sqrt{\pi}}{m}\frac{\Gamma\left(\frac{m+1}{2}\right)}{\Gamma\left(\frac{m}{2}\right)}$$

While these bounds are very tight for small θ, they diverge for larger values of θ due to the $\tan\theta$ term in the lower bound and the $\cos\theta$ term in the denominator of the upper bound. Since, in our model, a large half angle θ corresponds to low interference, we refine Shannon's estimates of [25]. First define the function $a(\theta; s, \beta)$ to be

$$a(\theta; \alpha, \beta) \triangleq (1 - \cos\theta)^\alpha (1 + \cos\theta)^\beta$$

The following theorem provides a family of bounds on the fraction of the sphere covered by a cap of half angle θ by finding the region of valid pairs of α and β such that $a(\theta; \alpha, \beta)$ is an upper or lower bound. We will denote the region of valid pairs for the for the lower and upper bounds L and U respectively. These can be seen in Figure 3-6 and the envelope of the these functions can be seen in Figure 3-8. This bound can be optimized over α and β in order to find the best bound for a given θ. Before stating this theorem and addressing the optimizing parameters, we state a general relationship between the bound $a(\theta; \alpha, \beta)$ and the parameters α and β.

Lemma 3.3.1. *If $\alpha > \beta$, then $0 \leq a(\theta; \alpha, \beta) \leq 1$. Otherwise, there exists a θ_0 such that $a(\theta; \alpha, \beta)$ is greater than one for $\theta \in (\theta_0, \pi/2)$.*

Proof. See Appendix B.2 □

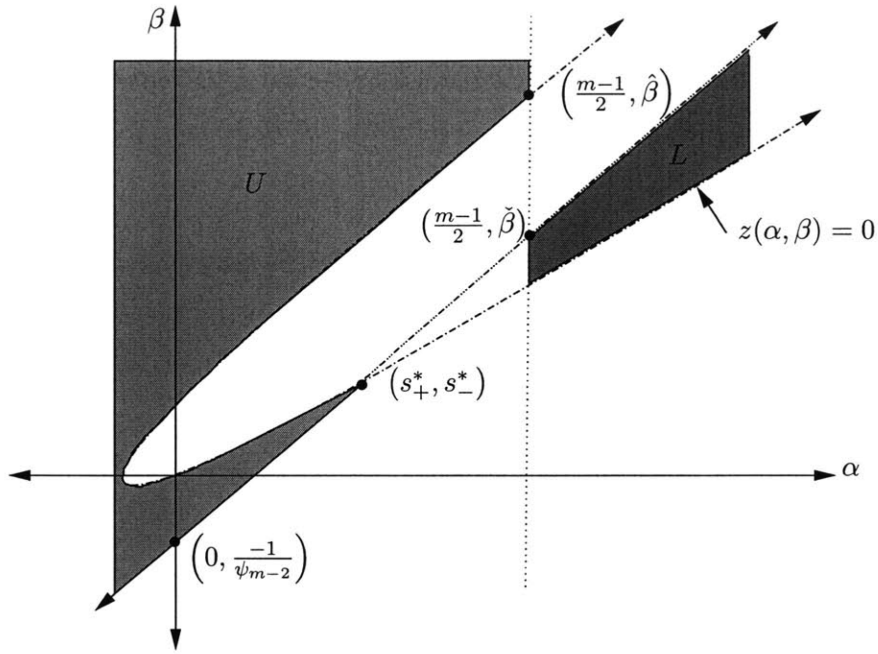

Figure 3-6. The regions for valid α and β pairs. The points corresponding to Corollary 3.3.3, Corollary 3.3.4 and Corollary 3.3.5 are labeled. The point $\left(\frac{m-1}{2}, \breve{\beta}\right)$ corresponds to the lower bound of Corollary 3.3.3, the point (s_+^*, s_-^*) corresponds to the upper bound of Corollary 3.3.4 and the point $\left(\frac{m-1}{2}, \hat{\beta}\right)$ corresponds to the upper bound of Corollary 3.3.5

We have the following general theorem on the probability $\delta_c(\theta, 2m)$.

Theorem 3.3.2. *Let $\delta_c(\theta, m)$ be the fraction of the real unit m-sphere covered by a pair of antipodal caps of half angle θ, where $0 \le \theta \le \frac{\pi}{2}$.*

$$\text{If } (\alpha, \beta) \in L \text{ then,} \quad \delta_c(\theta, m) \ge (1 - \cos\theta)^\alpha (1 + \cos\theta)^\beta$$
$$\text{If } (\alpha, \beta) \in U \text{ then,} \quad \delta_c(\theta, m) \le (1 - \cos\theta)^\alpha (1 + \cos\theta)^\beta$$

where

$$z(\alpha, \beta) \triangleq (\alpha - \beta)^2 - \alpha - \beta$$

$$L = \left\{ (\alpha, \beta) \;\middle|\; \alpha \ge \beta + \frac{1}{\psi_{m-2}} \text{ and } z(\alpha, \beta) \le 0 \text{ and } \alpha \ge \frac{m-1}{2} \right\}$$

$$U = \left\{ (\alpha, \beta) \;\middle|\; \alpha \le \beta + \frac{1}{\psi_{m-2}} \text{ and } z(\alpha, \beta) \ge 0 \text{ and } \alpha \le \frac{m-1}{2} \right\}$$

Proof. See Appendix B.3 $\square$

Corollary 3.3.3. *Let L be the set of all pairs (α, β) such that the function $a(\theta; \alpha, \beta)$ is a lower bound to the probability $\delta_c(\theta, m)$. Then, the set L is convex and for every pair $(\alpha, \beta) \in L$ the bound $a(\theta; \alpha, \beta)$ is a valid probability. Furthermore, for $0 < \theta < \frac{\pi}{2}$, the pair $\left(\frac{m-1}{2}, \frac{m-1}{2} - \psi_{m-2}^{-1}\right)$ can not be improved for the function $a(\theta; \alpha, \beta)$.*

Proof. Now, since $\alpha > \beta + \psi_{m-2}^{-1}$ for every $(\alpha, \beta) \in L$ we have by Lemma 3.3.1 that $a(\theta; \alpha, \beta)$ is a valid probability. Furthermore, L is the intersection of the interior of $z(\alpha, \beta)$ and half spaces, so L is clearly convex. Now, let θ be fixed. Then, the best lower bound is attained for

$$\begin{aligned}
(\alpha^*, \beta^*) &= \arg \max_{(\alpha,\beta)\in L} a(\theta; \alpha, \beta) && (3.27)\\
&= \arg \max_{(\alpha,\beta)\in L} \log a(\theta; \alpha, \beta) && (3.28)\\
&= \arg \max_{(\alpha,\beta)\in L} \alpha \log(1 - \cos\theta) + \beta \log(1 + \cos\theta) && (3.29)
\end{aligned}$$

where the second inequality follows from the fact that log is strictly increasing. Now, consider the polyhedral set

$$\tilde{L} = \left\{ (\alpha, \beta) \;\middle|\; \alpha \geq \beta + \frac{1}{\psi_{m-2}} \text{ and } \beta \geq \frac{1}{2}\left(m - \sqrt{3 + 4m}\right) \text{ and } \alpha \geq \frac{m-1}{2} \right\}$$

We have $L \subset \tilde{L}$ since we have replaced the constraint $z(\alpha, \beta) \leq 0$ with the weaker constraint $z(\frac{m-1}{2}, \beta) \leq 0$. Thus,

$$\max_{(\alpha,\beta)\in L} a(\theta; \alpha, \beta) \leq \max_{(\alpha,\beta)\in \tilde{L}} a(\theta; \alpha, \beta)$$

Now, we may upper bound (3.29) as

$$\max_{(\alpha,\beta)\in L} \alpha \log(1 - \cos\theta) + \beta \log(1 + \cos\theta) \leq \max_{(\alpha,\beta)\in \tilde{L}} (\alpha, \beta) c_L(\theta)^T$$

where the vector $c_L(\theta) = (\,\log(1 - \cos\theta)\,,\ \log(1 + \cos\theta)\,)$. Since this is the maximization of a bounded linear function over a polyhedral set we have by the Fundamental Theorem of Linear Programming [6] that the solution is an extreme point of the set $\tilde{L}$. Further, $c_L(\theta)$ only forms an angles between 0 and $\frac{-\pi}{4}$ with the positive β-axis. So by simple geometry (see Figure 3-7) we have $(\alpha^*, \beta^*) = ((m-1)/2, (m-1)/2 - \psi_{m-1}^{-1})$ for any θ. Since this is also a point in L the proof is complete by noting that every lower bound belongs to L.

□

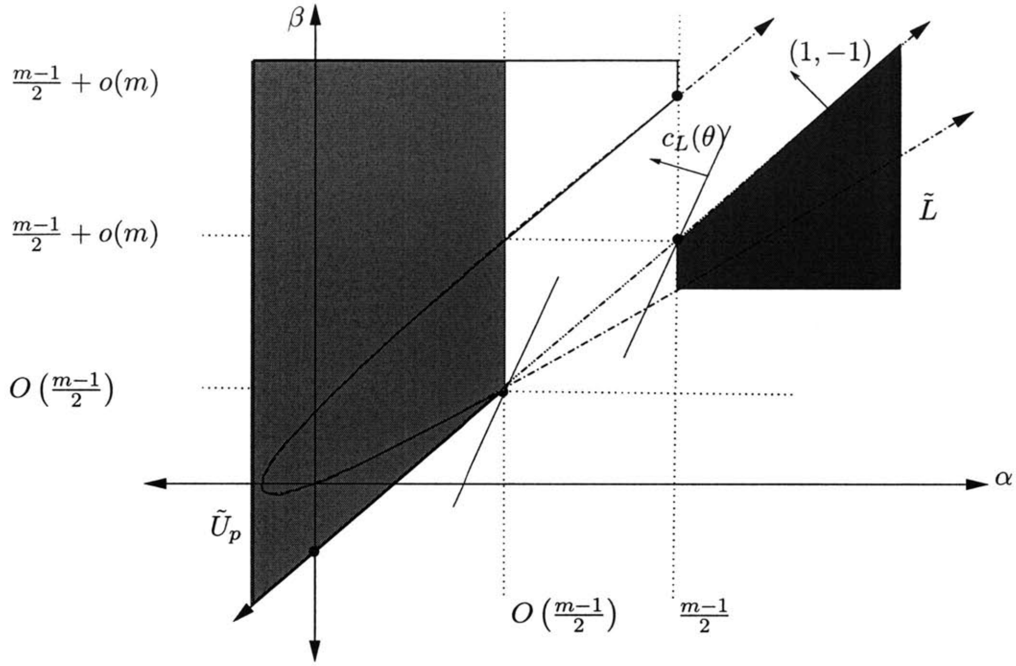

Figure 3-7. A geometric interpretation of the optimal solution to the bounds of Corollary 3.3.3 and Corollary 3.3.4. The polyhedral sets $\tilde{L}$ and $\tilde{U}_p$ are shaded. Note that the optimal pair of α, β from Corollary 3.3.4 are asymptotically off by a constant factor $\sqrt{2/\pi}$ from the lower bound while the pair of α, β from Corollary 3.3.5 are off by a term that is $o(m)$ which yields the correct exponent asymptotically.

We have not found such a simple characterization for the upper bound, since in general the region U is not convex. However, we can consider just the region of U intersected with the half space $\alpha > \beta$ if we wish to find bounds on the probability that are non-trivial over the entire region $0 \leq \theta \leq \frac{\pi}{2}$ as a consequence of Lemma 3.3.1. Considering just this region we have the following Corollary to Theorem 3.3.2.

Corollary 3.3.4. *Let*

$$U_p = \{(\alpha, \beta)\, | \alpha > \beta\} \cap U$$

be the set of all pairs (α, β) *such that the function* $a(\theta; \alpha, \beta)$ *is an upper bound to the probability* $\delta_c(\theta, m)$ *and bounded above by* 1 *where* U *is as in Theorem 3.3.2. Then, the pair* (s_+^*, s_-^*) *can not be improved upon for the function* $a(\theta; \alpha, \beta)$*, where* $s_\pm^* = \frac{1}{2}\left(\psi_{n-1}^{-2} \pm \psi_{n-1}^{-1}\right)$.

Proof. This proof follows exactly as the proof of Corollary 3.3.3, where we instead form a polyhedral set $\tilde{U}_p = \left\{(\alpha, \beta) \;\middle|\; \alpha \leq \beta + \frac{1}{\psi_m} \text{ and } \alpha \leq \frac{1}{2}\left(\psi_{m-2}^{-2} + \psi_{m-2}^{-1}\right)\right\}$

□

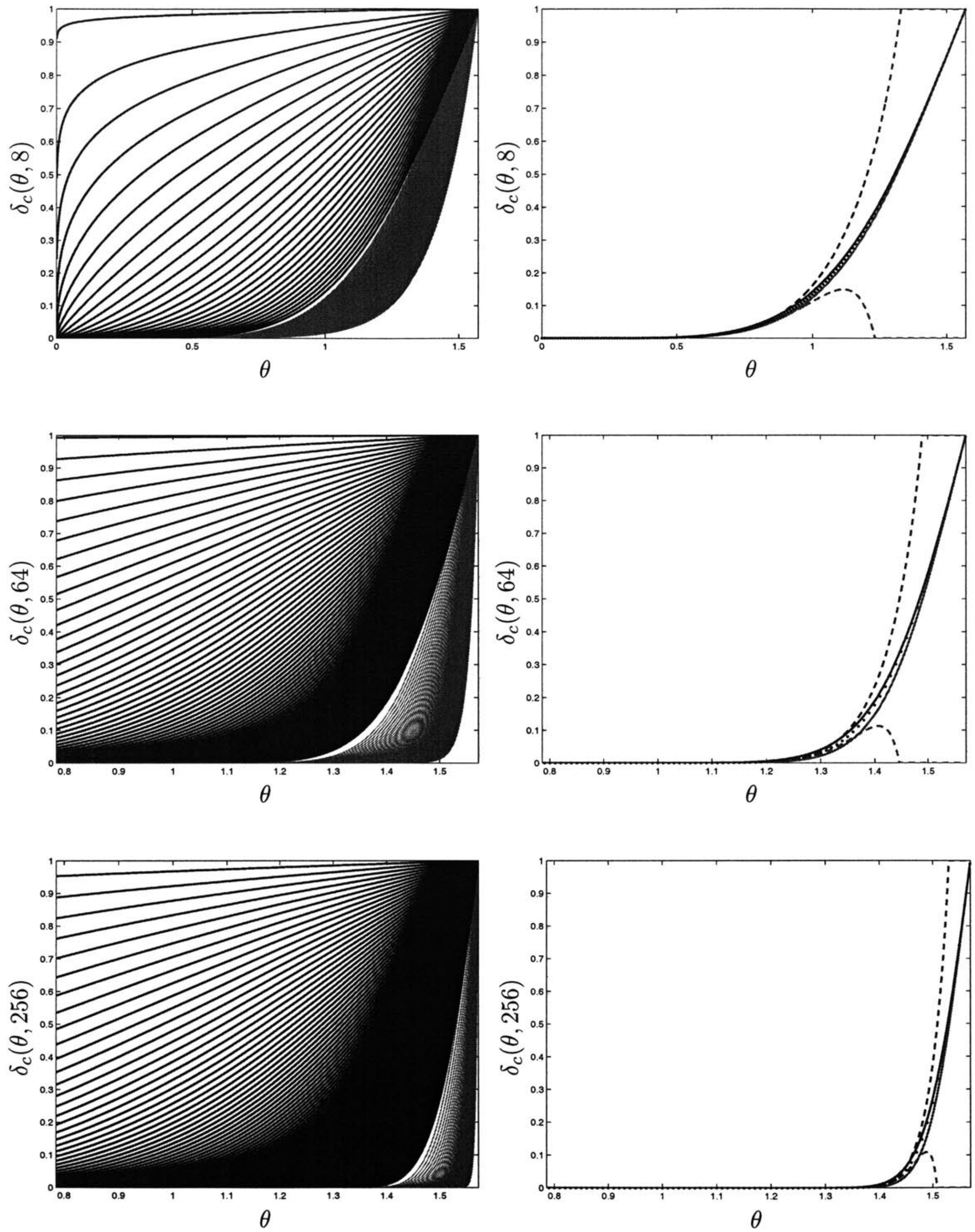

Figure 3-8. Bounds on the fraction of the sphere covered by a cap of half angle θ in 8 dimensions (top) and 64 dimensions (middle) and 256 dimensions (bottom). The envelope of the family of bounds and a comparison of the bounds of Corollary 3.3.4 to known bounds. The new estimates solid and the estimates from [25] dashed and exact expression dotted

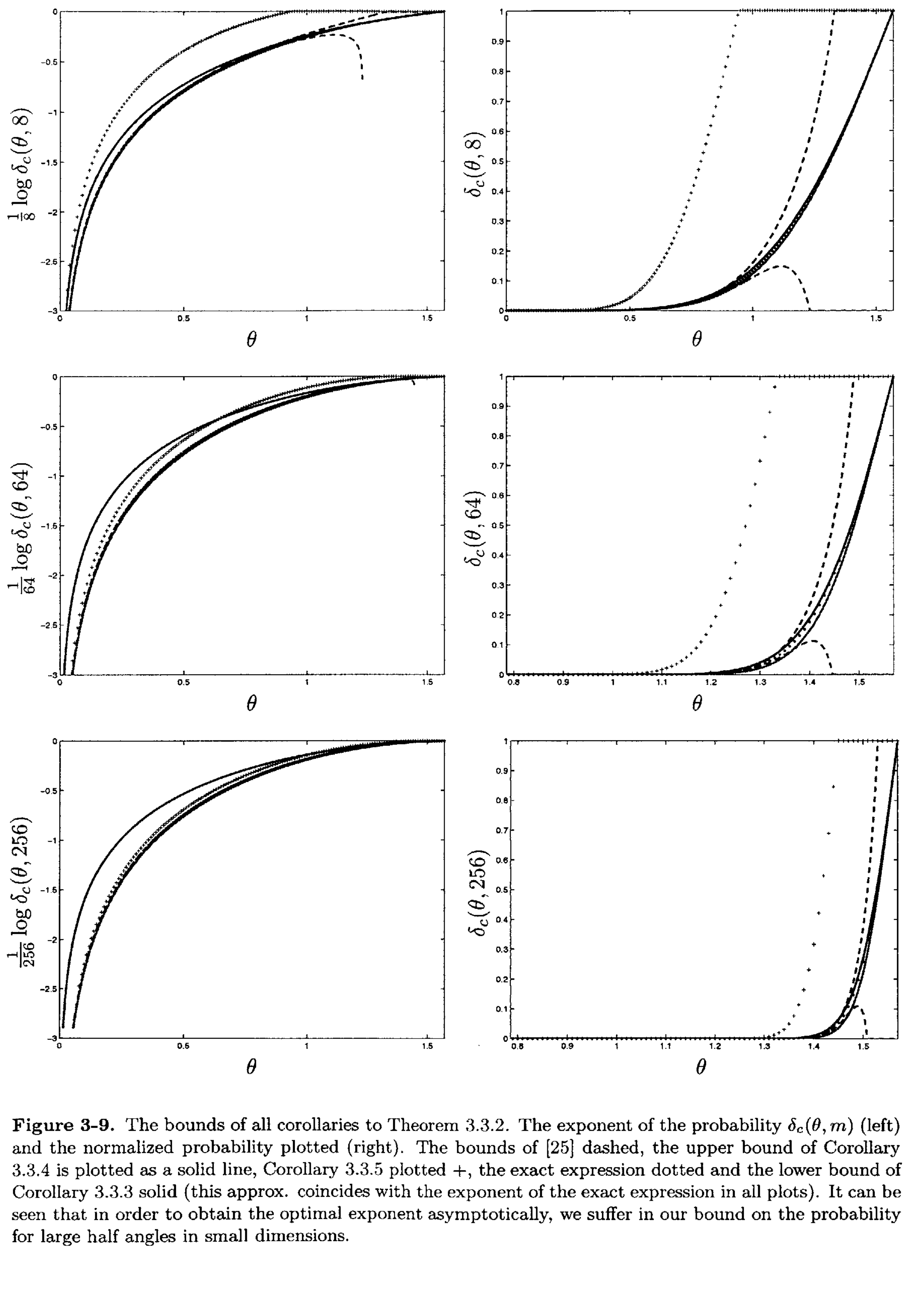

Figure 3-9. The bounds of all corollaries to Theorem 3.3.2. The exponent of the probability $\delta_c(\theta, m)$ (left) and the normalized probability plotted (right). The bounds of [25] dashed, the upper bound of Corollary 3.3.4 is plotted as a solid line, Corollary 3.3.5 plotted +, the exact expression dotted and the lower bound of Corollary 3.3.3 solid (this approx. coincides with the exponent of the exact expression in all plots). It can be seen that in order to obtain the optimal exponent asymptotically, we suffer in our bound on the probability for large half angles in small dimensions.

It can be seen in Figure 3-8 that the upper bound of Corollary 3.3.4 performs well for large half angles. However, we can improve on the bound for small angles if we remove the constraint that the upper bound is less than one yielding the following corollary.

Corollary 3.3.5. *Let $0 \le \theta \le \frac{\pi}{2}$ be given. Then,* $\delta_c(\theta, m) \le (1+\cos\theta)^{\frac{1}{2}\sqrt{3+4m}}(\sin\theta)^{m-1}$

The differing performance between Corollary 3.3.4 and 3.3.5 can be seen in Figure 3-9. The normalized log of the probability is plotted so that the performance of the bound may be examined for small θ, while the unnormalized probability is plotted to see the difference in performance for large θ. We further provide the following bounds on the exponent of this probability. That is, let

$$R(\theta, m) \triangleq \frac{1}{m}\log\delta_c(\theta, m)$$

then we have the following bounds.

Theorem 3.3.6. *Let m be specified and consider the normalized* log *of the probability $\delta_c(\theta, m)$, $R(\theta, m)$. Then*

$$R(\theta, m) \ge \left(\frac{m-1}{m} - \frac{2}{m\psi_{m-1}}\right)\log(\sin\theta) + \frac{1}{m\psi_{m-1}}\log(1-\cos\theta)$$

and

$$R(\theta, m) \le \frac{m-1}{m}\log(\sin\theta) + \frac{\sqrt{3+4m}}{2m}\log(1-\cos\theta)$$

Additionally,

$$R(\theta, m) = \log(\sin\theta) + \Theta\left(\frac{1}{\sqrt{m}}\right)$$

and as $m \to \infty$, $R(\theta, m) \sim \log(\sin\theta)$

Proof. The upper and lower bounds come through direct application of Corollary 3.3.5 and 3.3.3 respectively. Now it is easily shown that

$$\sqrt{\frac{2m}{\pi}} \le \frac{1}{\psi_m} \le \sqrt{\frac{2(m+1)}{\pi}} \tag{3.30}$$

using the inequalities

$$\sqrt{z+\frac{1}{4}} \le \frac{\Gamma(z+1)}{\Gamma(z+1/2)} \le \sqrt{z+\frac{1}{2}}$$

So, $\frac{1}{\psi_m} = o(m)$ which completes the proof.

□

Note that we can not prove the asymptotic result of the above theorem by using the bound of Corollary 3.3.4. This is because asymptotically, $s^*_\pm = \frac{2}{\pi}\frac{m-1}{2}$, so we are off by a constant as seen in Figure 3-9. Recall that Lemma 3.3.1 stated that if $\alpha > \beta$ there exists a θ_0 such that the bound $a(\theta; \alpha, \beta)$ is greater than 1 for all $\theta \in (\theta_0, \pi/2)$. Thus, it is natural to ask over what region the upper bound of Corollary 3.3.5 yields a valid probability. We will denote the largest such θ as θ^*. Note that since the bound of Corollary 3.3.4 is always a probability we can lower bound this by examining when these two bound cross. This will additionally also provide the region over which the exponent is improved by using Corollary 3.3.4. Now consider θ_0 such that

$$(1-\cos\theta_0)^{s^*_-}(1+\cos\theta_0)^{s^*_+} = (1-\cos\theta_0)^{\alpha^*}(1+\cos\theta_0)^{\beta^*}$$

where $s^*_\pm$ and α^*, β^* are the optimal pairs of Corollary 3.3.4 and 3.3.5 respectively. Taking the log of both sides and rearranging terms we get

$$\frac{\log(1+\cos\theta_0)}{\log(1-\cos\theta_0)} = -\frac{s^*_- - \alpha^*}{s^*_+ - \beta^*}$$

Now, since for $x > -1$, $\frac{x}{1+x} \le \log(1+x) \le x$ [1] we have

$$\frac{s^*_- - \alpha^*}{s^*_+ - \beta^*} \ge 1 - \cos\theta_0$$

So, since $\arccos\theta$ is a strictly decreasing function

$$\theta^* \ge \theta_0 \ge \arccos\left(1 - \frac{s^*_- - \alpha^*}{s^*_+ - \beta^*}\right)$$

Now,

$$1 - \frac{s^*_- - \alpha^*}{s^*_+ - \beta^*} = 1 - \frac{\frac{\alpha^*}{s^*_-} - 1}{\frac{\beta^*}{s^*_-} - \frac{s^*_+}{s^*_-}} \le 1 - \frac{\frac{\alpha^*}{s^*_-} - 1}{\frac{\beta^*}{s^*_-} - 1} \le \frac{s^*_-}{\alpha^*} = O\left(\frac{1}{\sqrt{m}}\right)$$

Similarly, we have that

$$\frac{s^*_- - \alpha^*}{s^*_+ - \beta^*} \ge \frac{1}{1+\cos\theta_0}$$

and by reversing the role of the parameters

$$\frac{s^*_+ - \beta^*}{s^*_- - \alpha^*} - 1 = \Omega\left(\frac{1}{\sqrt{m}}\right)$$

Thus,

$$\theta^* = \frac{\pi}{2} - \Theta\left(\frac{1}{\sqrt{m}}\right)$$

and asymptotically the upper bound yields the correct probability. This agrees with the convergence that can be seen in Figure 3-9.

We now provide a proof of a simpler bound using the same method that will be used to prove theorem 3.3.2 in the following section. The method was used in [4] to provide bounds on the incomplete gamma function which we used in the proof of Theorem 3.0.4.

3.3.1 A simple bound and its proof with a general technique

In this section we consider the family of bounds defined by the function

$$b_m(\theta; s) \triangleq \left(\frac{2\theta}{\pi}\right)^s$$

We will derive bounds on the probability $\delta_c(\theta, m)$ in terms of this function. We begin with a general description of the method that we will use. This method was used in [4] and uses the properties of the error term in order to characterize the best exponents that a specific form of a bound can take.

Begin by noting that in order to derive the fraction of the sphere covered by a cap of half angle $0 < \theta < \frac{\pi}{2}$ we can first project the cap on to the plane orthogonal to its center and then integrate over this region. This can be done in a one to one manner since $\theta < \frac{\pi}{2}$. Now, with out loss of generality, suppose that the cap is centered about $(0, 0, \ldots, 0, 1)$. Then, defining R to be the image of the projection (as seen in Figure 3-10), we have that

$$\delta_c(\theta, m) = \iint_R \frac{dy}{\sqrt{1 - \sum_{i=1}^{m-1} y_i^2}}$$

It is well known that (see for example [26]) letting $r = \sum_{i=1}^{m-1} y_i^2$ and integrating over the spherical shells of radius r yields

$$\begin{aligned}
\delta_c(\theta, m) &= \frac{2}{mC_m} \int_0^{\sin\theta} \frac{(m-1)C_{m-1} r^{m-2}}{\sqrt{1 - r^2}} dr \\
&= \frac{1}{\psi_{m-2}} \int_0^{\theta} \sin^{m-2}\phi \, d\phi
\end{aligned}$$

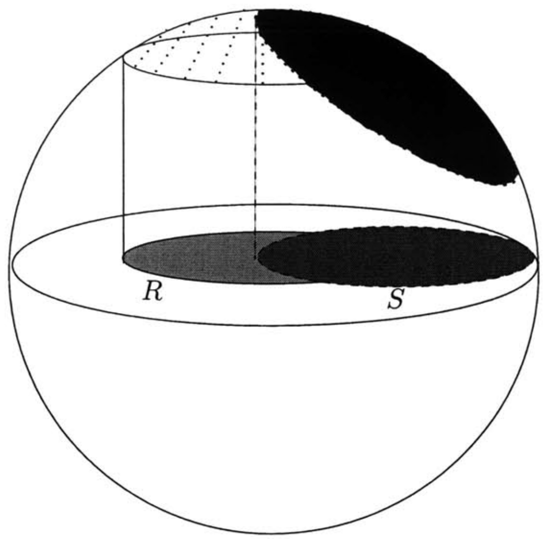

Figure 3-10. Two spherical caps and their projection onto a plane

where

$$\psi_m \triangleq \frac{\sqrt{\pi}\Gamma\left(\frac{m+1}{2}\right)}{m\Gamma\left(\frac{m}{2}\right)} = \int_0^{\frac{\pi}{2}} \sin^m \phi \, d\phi$$

and mC_m is the surface area of the unit m sphere. Thus, we consider bounding the integral

$$\frac{1}{\psi_m} \int_0^{\theta} \sin^m \phi \, d\phi \qquad 0 \leq \theta \leq \frac{\pi}{2}$$

Begin by cosidering the error term

$$D_m(\theta; f(\cdot; s)) \triangleq \int_0^{\theta} \sin^m \phi \, d\phi - \psi_m f(\theta; s) \tag{3.31}$$

Note that if this function is non-negative for every $0 \leq \theta \leq \frac{\pi}{2}$ the function f is a lower bound. Similarly, if D_m is always non-positive then the function f is an upper bound. More generally if there is a θ_0 so that $D_m(\theta; b_m) < 0$ for all $\theta_0 < \theta < \frac{\pi}{2}$ the function $b_m(\theta; s)$ is an upper bound over this interval. A similar statement can be made for $-D_m$ in order for $b_m(\theta; s)$ to be a lower bound.

We will examine functions that are tight at $\theta = 0$ and $\theta = \frac{\pi}{2}$. Thus for any function that is a lower bound D_m must be increasing at 0. Otherwise f would fail to be a lower bound in a neighborhood about $\theta = 0$. Similarly, D_m must be decreasing at $\theta = \frac{\pi}{2}$ for f to be a lower bound. Thus in order to characterize the regions where f is a lower bound we

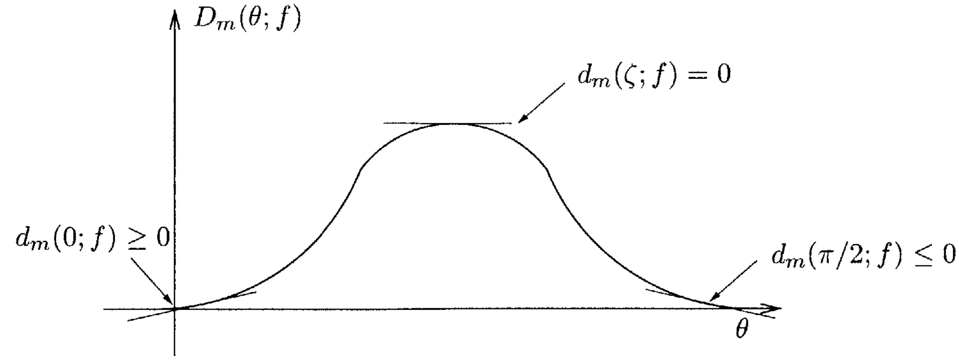

Figure 3-11. A graphical depiction of sufficient conditions for d_m to be a lower bound

examine a scaled version of the derivative of the difference

$$d_m(\theta; f(\cdot; s)) \triangleq \frac{1}{\sin^m \theta} \frac{\partial}{\partial \theta} D_m(\theta; f(\cdot; s))$$

Note that if the function $d_m(\theta; f(\cdot; s))$ only has a single root on the interval $(0, \pi/2)$ then the function D_m is either positive or negative since D_m can never cross zero. The constraints on d_m are depicted in Figure 3-11.

For the functions b_m

$$d_m(\theta; b_m) = 1 - \frac{2\psi_m s}{\pi} \left(\frac{b_m(\theta; (s-1)/m)}{\sin \theta} \right)^m$$

Note that this implies that the derivative is zero precisely when

$$g(\theta; r) = \left(\frac{2\psi_m s}{\pi} \right)^{\frac{-1}{m}} \quad \text{where} \quad g(\theta; r) \triangleq \frac{b_m(\theta; r)}{\sin \theta}$$

where $r = (s-1)/m$. We will show that for any r there exists a θ^* such that $g(\theta; r)$ is decreasing for $\theta \in (0, \theta^*)$ and increasing for $\theta \in (\theta^*, \pi/2)$. That is, we will show that

$$h(\theta; r) = \left(\frac{2}{\pi} \right)^r \frac{\theta^{r-1}}{\sin \theta} \frac{\partial}{\partial \theta} g(\theta; r) \quad = r - \theta \cot \theta$$

is convex. This is clear, since

$$\frac{\partial^2}{\partial \theta^2} h(\theta; r) = (2 \csc^2 \theta)(1 - \theta \cot \theta) > 0 \; \forall \; \theta \in (0, \pi/2) \tag{3.32}$$

which is independent of r. Now, if $r < 1$, let ζ_r be such that $r - \zeta_r \cot \zeta_r = 0$. Then, $g(\theta; r)$ is strictly decreasing on $(0, \zeta)$ and strictly increasing on $(\zeta, \pi/2)$. Otherwise if $r > 1$, $g(\theta; r)$ is strictly increasing since $h(\theta; r)$ is convex and

$$\lim_{\theta \to \frac{\pi}{2}} h(\theta; r) = r \text{ and } \lim_{\theta \to 0} h(\theta; r) = r - 1$$

Now, we trivially have that $D_m(0; s) = D_m(\pi/2; s) = 0$. Further,

$$\begin{array}{rrcl} & \frac{\partial}{\partial \theta} D_m(\theta; s)\big|_{\theta = \frac{\pi}{2}} & \geq & 0 \\ \iff & g(\pi/2; r) & \geq & \left(\frac{2\psi_m s}{\pi}\right)^{\frac{-1}{m}} \\ \iff & 1 & \geq & \left(\frac{2\psi_m s}{\pi}\right)^{\frac{-1}{m}} \\ \iff & s & \leq & \frac{\pi}{2\psi_m} \end{array}$$

Now, let $s_u(m)$ be the largest exponent that achieves the upper bound. That is

$$s_u(m) \triangleq \frac{\pi}{2\psi_m}$$

Note, that the constraint on $s_u(m)$ comes from the behavior at the endpoint $\frac{\pi}{2}$. So, in order for s to be an upper bound over $(\theta_0, \pi/2)$ we must have $s \leq \frac{\pi}{2\psi_m}$ and θ_0 can be taken to be zero.

We now turn to finding the smallest sequence $s_l(m)$ such that the function $b_m(\theta; s_l)$ is a lower bound. We note, that the main constraint on this is that the difference $D_m(\theta; s)$ must be increasing at $\theta = 0$. A second constraint comes from our derivation of the upper bound. Indeed if $s \leq s_u(m)$ then there would exist a θ^* such that $D_m(\theta^*; s)$ is negative for all $\theta > \theta^*$. Now, for $r < 1$ we have

$$\lim_{\theta \to 0} g(\theta, r) = \infty$$

which implies that $D_m(\theta; s)$ is decreasing at 0 if $s < m + 1$. So, $s_l(m) = m + 1$.

We now provide a useful example on a method to apply the upper bound. Note that since $b_m(\theta; s)$ is strictly increasing we may take *any* function $f(\theta)$ such that $f(\theta) > \theta$ for all

$\theta \in (0, \pi/2)$. For example, it is often useful to work with $\cos\theta$ instead of θ itself. So, since

$$\theta = \arccos\cos\theta \leq \frac{\pi}{2} - \cos\theta$$

the function

$$\hat{b}_m(\theta, s) \triangleq \left(1 - \frac{\pi}{2}\cos\theta\right)^{s(m)}$$

is also an upper bound, but is now not tight at $\theta = 0$. Additionally, we may arrive at better lower bounds over a smaller interval. In fact, we have already shown that if $s \in (s_u, m+1)$ then there is a unique θ^* such that $b(\theta; s)$ is a lower bound for $\theta > \theta^*$. Conversely, if we are given a θ_0 we can find the best s by solving the equation $D_m(\theta_0; s) = 0$.

We now present the following theorem that generalizes the result on the function b_m. First define the function $c(\theta; s, \beta)$ to be

$$c(\theta; s, \beta) \triangleq \left(\frac{2}{\pi}\theta\right)^s \sin^\beta\theta$$

Note that $b_m(\theta; s) = c(\theta; s, 0)$. We have the following theorem.

Theorem 3.3.7. *Let $\delta_c(\theta, m)$ be the fraction of real unit m sphere covered by antipodal caps of half angle θ. Then, for any $\beta \geq m-1$ and $s^* = \frac{\pi}{2\psi_{m-2}}$ then*

$$\delta_c(\theta, m) \geq c(\theta; s, \beta)$$

Furthermore, for any $s \leq s*$ and for all $0 < \beta \leq \frac{(s-1)s}{\pi}$

$$\delta_c(\theta, m) \leq c(\theta; s, \beta)$$

Proof. See Appendix B.4. □

3.4 Lower Bounds to Multiplexing Rates

The probability of existence of a near-orthogonal subset among n independently formed m-dimensional complex circularly symmetric channel vectors have been bounded in Section 3.1 and a bound on the probability, $p_\perp$, that any set is ϵ-orthogonal was found in Section 3.3. It is evident from the bounds of Theorem 3.2.3 that for any given $\theta_{\epsilon,\rho}$ the probability of finding an ϵ-orthogonal set of channels quickly jumps to 1. This phase transition behavior has direct application to algorithm design for scheduling in the MIMO broadcast channel.

Recall that it was our ultimate goal to use the bound of Theorem 3.2.3 for user selection and that we have shown in Theorem 3.0.4 that the lower bound on the channel norms can grow at a rate of at most $\log n$ for $\Pr(|\mathcal{S}_\epsilon| > 0) > 0$. Let us now briefly consider the implications of these theorems. These two theorems imply that the SNR can increases at a rate on the order of $\log n$, which will correspond to sum-rate increasing on the order of at most $\log \log n$. As such, if we have that $\Pr(|\mathcal{S}_\epsilon| > 0) > 1 - \delta$ for some small δ we will have to examine exponentially more users to realize a relatively small gain in rate. We will let n_δ be the fewest number of users that need to be examined to have $\Pr(|\mathcal{S}_\epsilon| > 0) > 1 - \delta$.

From inspecting Theorem 3.2.3, it is easy to see that n_δ is finite for $p_\perp > 0$ and $p_s > 0$ and that the threshold can be sharp. The sharpness of this threshold has two possible explanations. First, the number of points that fall in a spherical shell follows a binomial distribution which is known to exhibit a thresholding behavior [18]. Secondly, it is reasonable to expect that there will additionally be a rapid emergence of an ϵ-orthogonal set since it can be roughly modeled by the existence of a clique of size m in a binomial random graph [27,23].

The existence of such a threshold implies that the search among the set of n users for a set to maximize the objective function can be restricted to n_δ users. This is very significant, of course, in reducing algorithm complexity. Indeed, we have shown that any algorithm that tries to approximate the optimal set using say k users and a criterion based inner products and norms of the channel vectors will have a small probability of success if $k \ll n_\delta$. However, if the same algorithm is employed in the scenario where $k \gg n_\delta$, then we can expect a high probability of success. Moreover, using this search does not result in an appreciable loss in throughput provided that ρ^- can be increased sufficiently fast.

The threshold n_δ also has a valuable application in providing lower bounds for the expected rate of any multiplexer in a channel with choice. If we have that $n > n_\delta$, then we can randomly choose any set of users $\mathcal{A} \in \mathcal{S}_\epsilon$ and expect good performance. To be more precise, let f_{rate} be the sum rate expression for a given multiplexer. Then, we can use our SIR and SNR guarantees to provide simple bounds for $f_{\text{rate}}(\mathcal{A})$, say $f_{\text{bnd}}(\text{SNR}, \text{SIR})$. Then,

$$\begin{aligned}
\mathbb{E} \max_{\mathcal{A}} f_{\text{rate}}(\mathcal{A}) &\geq \mathbb{E} \max_{\mathcal{A} \in \mathcal{S}_\epsilon} f_{\text{rate}}(\mathcal{A}) && (3.33)\\
&\geq \Pr(|\mathcal{S}_\epsilon| > 0) f_{\text{bnd}}(\text{SNR}, \text{SIR}) && (3.34)\\
&\geq (1 - \delta(n)) f_{\text{bnd}}(\text{SNR}, \text{SIR}) && (3.35)
\end{aligned}$$

We can further optimize the above bound over the parameters ϵ, ρ^- and ρ^+. Recall from Theorem 2.2.2, in the case of zero forcing multiplexing, with power constraint P, we have for any set in $\mathcal{S}_{\text{zf}}$,

$$f_{\text{zf}}(\mathbf{H}_\mathcal{A}, \mathbf{q}(\mathcal{A})) = \log\left(1 + \frac{P}{|\mathcal{A}|}\mathcal{H}\left(\{b_i(\mathcal{A})\}\right)\right) + D(q(\mathcal{A}) \| b_\mathcal{A})$$

Now, since $b_i(\mathcal{A})$ is the projection of the channel vector $\mathbf{h}_i$ away from the space spanned by every other user in $\mathcal{A}$, b_i is decreasing in the magnitude of the inner product between every user in the set. Thus, in order to arrive at a lower bound we may assume that the inner product between every user is ϵ. In this direction, let $W_{\epsilon,\rho}$ be the matrix with diagonal elements equal to ρ and off diagonal elements equal to ϵ. That is, $W_{\epsilon,\rho}$ is the Wishart matrix of the set of users with the worst possible mutual interference.

Lemma 3.4.1. *Let $0 \leq \epsilon \leq \rho$. Then,*

$$\operatorname{Tr}\left(W_{\epsilon,\rho}^{-1}\right) = \frac{1}{\rho} \frac{m + (m^2 - 2m)\frac{\epsilon}{\rho}}{\left(1 - \frac{\epsilon}{\rho}\right)\left(1 + (m-1)\frac{\epsilon}{\rho}\right)}$$

Proof. We will prove this fact by finding the eigenvalues of $W_{\epsilon,\rho}$. First note that the all one vector, $\mathbf{j}$, is an eigenvector since $\mathbf{v} = \mathbf{j}\ W_{\epsilon,\rho}$ simply has the sum of each column in each entry. That is $v_i = \rho + (m-1)\epsilon$ for each i. Thus, $\mathbf{j}$ is an eigenvector with eigenvalue $\rho + (m-1)\epsilon$. We now claim that any vector that has -1 in the ith position and $1/(m-1)$ elsewhere is an eigenvector. Indeed, if $\mathbf{v}_i$ is this vector then $\mathbf{v}_i W_{\epsilon,\rho} = (\rho - \epsilon)\mathbf{v}_i$. Since the set of vectors $\mathbf{v}_i$ span a space of dimension $(m-1)$ (note that the sum of all $\mathbf{v}_i$ is the zero vector) we have an eigenvalue $\lambda_1 = \rho + (m-1)\epsilon$ with multiplicity 1 and an eigenvalue of

$\lambda_2 = \rho - \epsilon$ with multiplicity $m-1$. Thus,

$$\text{Tr}(W_{\epsilon,\rho}) = \frac{1}{\lambda_1} + \frac{m-1}{\lambda_2} = \frac{m\rho + (m-1)^2\epsilon - \epsilon}{(\rho-\epsilon)(\rho+(m-1)\epsilon)}$$

□

Note that we may write,

$$f_{\text{zf}}(\mathbf{H}_{\mathcal{A}}, \mathbf{q}(\mathcal{A})) \geq \log\left(1 + \frac{P}{|\mathcal{A}|}\mathcal{H}\left(\{b_i(\mathcal{A})\}\right)\right) = \log\left(1 + \frac{P}{\text{Tr}(W_{\mathcal{A}}^{-1})}\right)$$

since $D(\cdot\|\cdot)$ is always non-negative [11]. This yields the following lower bound.

Theorem 3.4.2. *Let $\epsilon < \rho^- < \rho^+$ be given. Then, the* SIR *between any pair of users is lower bounded by* $\text{SIR}_{\epsilon,\rho} = \frac{\rho^-}{\epsilon}$ *and*

$$\mathbb{E}_{\mathbf{H}}\, R_{\text{sum}}^{\text{zf}} \geq \max_{0<l\leq m} \Pr\left(N_\epsilon^{(l)} > 0\right)\; l\, \log\left(1 + P\rho^- \frac{\left(1 - \frac{1}{\text{SIR}_{\epsilon,\rho}}\right)\left(1 + (l-1)\frac{1}{\text{SIR}_{\epsilon,\rho}}\right)}{l + (l^2 - 2l)\frac{1}{\text{SIR}_{\epsilon,\rho}}}\right) \tag{3.36}$$

Furthermore, letting $\rho^- = \log(n)$ and $\rho^+ = \log(n) + d$ for $d > 0$ and sufficiently large n

$$\mathbb{E}_{\mathbf{H}}\, R_{\text{sum}}^{\text{zf}} \geq m\left(1 - \Theta(n^{-\gamma m})\right) \log\left(1 + P\log(n)\, \frac{\left(1 - \frac{1}{\text{SIR}_{\epsilon,\rho}}\right)\left(1 + (m-1)\frac{1}{\text{SIR}_{\epsilon,\rho}}\right)}{m + (m^2 - 2m)\frac{1}{\text{SIR}_{\epsilon,\rho}}}\right) \tag{3.37}$$

Proof. Note that the lower bound (3.36) follows by applying Lemma 3.4.1, (3.35), and the discussion preceding (3.35). The bound (3.37) follows by noting that for significantly large n and any ϵ and ρ we have $\Pr(N_\epsilon^{(m)} > 0) = 1 - \Theta(n^{-m})$.

□

There is a tradeoff between the probability of existence and the SIR constraint in (3.36). That is, the probability of existence is decreasing in SIR while the rate function is increasing in SIR. An example of how one may achieve a good point on this trade-off (in the case of zero-forcing multiplexing and four transmit antennas) can be seen in Figure 3-12. First consider the contours of equal existence probability 0.9 as seen in Figure 3-12. The curves show that to find an ϵ-orthogonal set with this probability we must either select only pairs of users or have $\epsilon/\rho^+ > 0.8$. Clearly a high rate (*i.e.* small ϵ) requires taking $|\mathcal{A}| = 2$ in this example.

Now, one would like to take ϵ at the knee of the level contour of a desired existence probability. Such an ϵ not only guarantees high rate at a low number of users, which lowers complexity, but also provides robust performance when the number of users can vary. The effect of choosing a large ϵ is exhibited in Figure 3-12(b), where $\epsilon = 0.98$. Here, the probability quickly jumps to one. However, the rate remains the small constant f_{bnd} for all n thereafter. Alternatively, if ϵ is taken to be too small then the existence probability will be too low, as seen Figure 3-12(b) in the case $\epsilon = 0.2$. Note that the climb in rate indicates that the probability of existence is still increasing to 1.

In a fashion similar to that of Theorem 3.4.2 we may additionally upper and lower bound the beamforming and dirty paper coding rates. We omit the statement of the scaling properties of dirty paper coding since it is easily seen to have an expected rate that is $\Theta(\log\log n)$ since it always dominates the zero forcing rate and is dominated by the rates of a parallel channel. However, we have the following characterization of the beamforming rate.

Theorem 3.4.3. *Let $\epsilon < \rho^- < \rho^+$ be given. Then the* SIR *of between ever pair of users is lower bounded by* $\text{SIR}_{\epsilon,\rho} = \frac{\rho^-}{\epsilon}$ *and*

$$\mathbb{E}_{\mathbf{H}}\, R^{\text{sp}}_{\text{sum}} \geq \max_{0<l\leq m} \Pr\left(N_\epsilon^{(l)} > 0\right)\; l\, \log\left(1 + \frac{P}{l}\frac{\text{SIR}_{\epsilon,\rho}}{l-1+1/\epsilon}\right) \tag{3.38}$$

Proof. The proof of (3.38) follows directly as the proof of (3.36). This is just a simple application of Theorem 3.2.3 and Theorem 2.2.1. □

Now, note that in the right hand side of (3.36) the SNR is multiplied by a constant that reflects the effects of our choice of target SIR on the expected rate under zero forcing multiplexing. We expect this term to vary depending on how we choose to multiplex our signal. For example, in the case of supposition coding we saw that this constant is just

$$\frac{\text{SIR}_{\epsilon,\rho}}{l-1+1/\epsilon}$$

Thus, if we can not grow the SIR faster than $\log n$ we expect there to be an asymptotic SNR gap in the expected rate under different multiplexers. Figure 3-13 bears on this question.

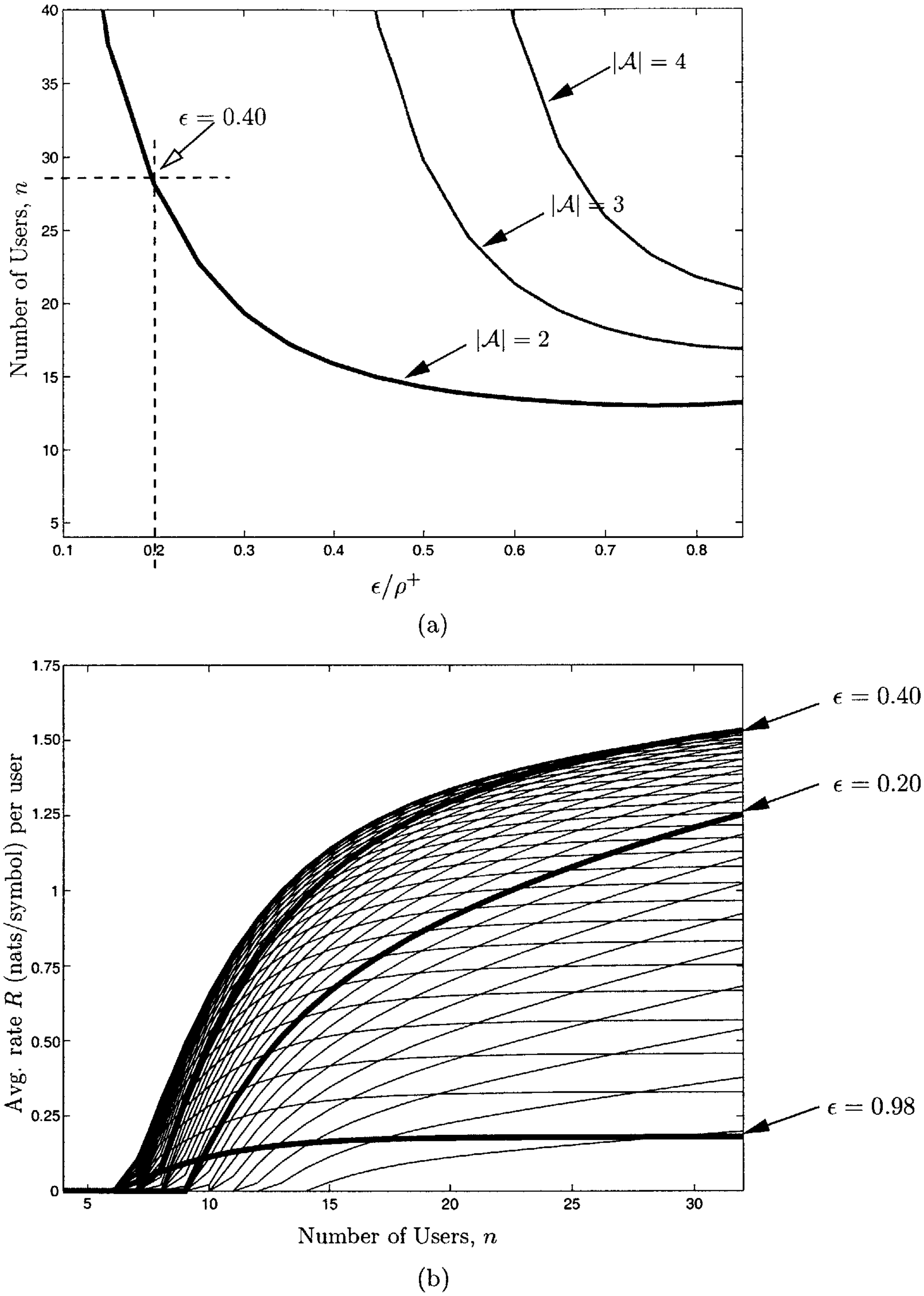

Figure 3-12. An example of using ϵ-orthogonal selection for a 4 transmit antenna system with $\rho^- = 1$ and $\rho^+ = 2$ and zero forcing multiplexing. (a) The level curves for the lower bound on the probability of existence for $\delta = 0.1$. (b) The lower bound on the expected throughput of selecting 2 users for various ϵ. The simulated throughput using optimal selection can be seen above the convex hull of these bounds

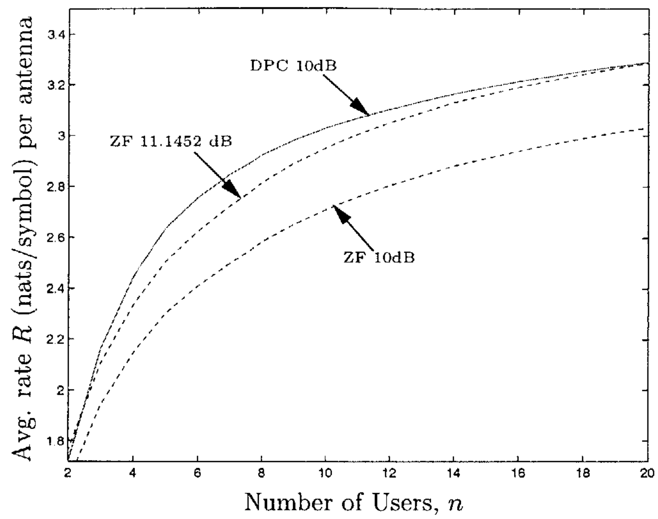

Figure 3-13. The empirical average rate of optimal zero forcing multiplexing and dirty paper coding with nominal SNR $= P$ of 10dB

Note by Theorem 3.4.2, we would have

$$\mathbb{E}_{\mathbf{H}}\, R^{\text{zf}}_{\text{sum}} \sim m \log\left(1 + \frac{P}{m}\log(n)\right)$$

if we can take SIR $= \log n$ while still having the probability of existence going to one. Additionally, note that we can not address the rate at which simple superposition coding scales due to the fact that we do not know the rate at which we can scale $\text{SIR}_{\epsilon,\rho}$. However, by Corollary 3.2.4, if we can take $\epsilon(n) = o(\log n)$ and $\rho^{-}(n) = \log n$ and still have

$$p_{\perp} = \log\left(\frac{m \log n}{m \log n - \log\log n}\right)$$

as $n \to \infty$ then $\text{SIR}_{\epsilon,\rho} \sim \log n$. Then, by Theorem 3.4.3

$$\begin{aligned}\mathbb{E}_{\mathbf{H}}\, R^{\text{sp}}_{\text{sum}} &\geq \left(1 - O\left(\frac{1}{\log n}\right)\right)\left(m \log\left(1 + \frac{P}{m}\log(n)\right)\right) \\ &\sim m \log\left(1 + \frac{P}{m(m-1)}\log(n)\right)\end{aligned}$$

This naturally leds us to ask, at what rate can we scale the SIR. However, it is not clear the fastest rate one can scale the SIR while still maintaining a non-zero probability that there exists a set that meets this SIR constraint.

Chapter 4

Conclusions

In this thesis we have examined the problem of joint multiplexing and scheduling of independent data streams for the MIMO broadcast channel. In order to provide a solution to the scheduling problem that yields a low complexity selection algorithms with close to optimal performance, we defined sets with guaranteed SIR and SNR values. These guarantees were defined by placing restrictions on the norm of users channels and the magnitudes of the inner products. We showed that interpreting the pairwise guarantee as a random packing yielded a simple graph structure that enabled us to provide good bounds on the probability of an ϵ-orthogonal set.

In order to derive good bounds on the probability of existence we employed the theory of random geometric graphs. In this context, we showed that the existence of a complete subgraph in the larger random graph was equivalent to the existence of an ϵ-orthogonal set. After completely characterizing this existence probability as a function of the parameters p_s and $p_\perp$ (the probability of falling in a spherical shell and the probability that any given set forms a packing respectively), we examined characterizing the probability $p_\perp$.

In the direction of finding a good bound on $p_\perp$, we refined the bounds of Shannon [25] in Section 3.3 on the fraction of area covered by a spherical caps in arbitrary dimensions by providing a family of bounds $a(\theta; \alpha, \beta)$. We classified the optimal parameters for the lower bound and the upper bound under the constraint that the function is bounded below by zero and above by 1. We provided a new proof of the exponent of the probability that a point chosen uniformly on the unit sphere falls inside a cone of half angle θ. Further, it was shown that this exponent converges like $\Theta(\frac{1}{\sqrt{m}})$. This exponent has direct applications to the reliability function for the scalar AWGN channel (see [25]).

Using the lower bound on the probability of existence of ϵ-orthogonal sets we provided bounds on the expected rate in a channel with choice in Section 3.4. We used the the fact that there is a quick increase in the probability of existence of an ϵ-orthogonal set to develop a novel method to select sets of users that will obtain close to optimal performance. We further discussed how one can use this to additionally provided close to optimal performance with a complexity constraint.

We left as an open question the rate at which we can scale the SIR as function of the number of users n while still maintaining a non-zero probability of existence. This is the focus of future work. Further, we will additionally examine the question of weather there is an asymptotic gap in SNR between different multiplexers.

Appendix A

Proofs for Chapter 2

A.1 Proof of Theorems 2.2.1 and 2.2.2

We provide the proof for zero forcing multiplexing. The proofs for beamforming follow the same derivation. Under zero-forcing our problem becomes,

$$f(q, \mathcal{A}) = \max_{\sum_i P_i/b_i \le P} \sum_i q_i \log(1 + P_i)$$

Using convexity, writing the Lagrangian and taking the derivative yields the necessary and sufficient condition

$$P_i = \lambda q_i b_i - 1 \qquad \text{where} \qquad \sum_i \left(q_i \lambda - \frac{1}{b_i}\right)_+ = P \tag{A.1}$$

Now, suppose every user gets a strictly positive power. Then, noting that

$$\lambda = \frac{P + \sum_i \frac{1}{b_i}}{\sum_{i \in \mathcal{A}} q_i}$$

yields

$$f(q, \mathcal{A}) = \sum_i q_i \log(1 + P_i) = \left(\sum_i q_i\right) \log(\lambda) + \sum_i q_i \log(b_i\, q_i) \tag{A.2}$$

Now using the cofactor expansion for the inverse we have,

$$b_i = \frac{\det(\mathbf{H}_{\mathcal{A}} \mathbf{H}_{\mathcal{A}}^\dagger)}{\det(M_{ii}(\mathbf{H}_{\mathcal{A}} \mathbf{H}_{\mathcal{A}}^\dagger))} = \frac{A(\mathcal{A})}{M_{i,i}(\mathcal{A})}$$

Returning to (A.2) we have

$$
\begin{aligned}
f(q,\mathcal{A}) &= \left(\sum_{i\in\mathcal{A}} q_i\right)\log\left(\frac{P+\frac{1}{A(\mathcal{A})}\sum_i M_{i,i}(\mathcal{A})}{\sum_{i\in\mathcal{A}} q_i}\right) - \sum_i q_i \log\left(\frac{M_{i,i}(\mathcal{A})}{q_i A(\mathcal{A})}\right) \\
&= \left(\sum_{i\in\mathcal{A}} q_i\right)\left(\log\left(PA(\mathcal{A})+\sum_i M_{i,i}(\mathcal{A})\right) - \sum_{i\in\mathcal{A}} \frac{q_i}{\sum_{i\in\mathcal{A}} q_i}\log\left(\frac{\left(\sum_{i\in\mathcal{A}} q_i\right) M_{i,i}(\mathcal{A})}{q_i}\right)\right) \\
&= \left(\sum_{i\in\mathcal{A}} q_i\right)\left(\log\left(\frac{PA(\mathcal{A})}{\sum_i M_{i,i}(\mathcal{A})}+1\right) + \sum_{i\in\mathcal{A}} \tilde{q}_i \log\left(\frac{\tilde{q}_i}{\tilde{m}_i}\right)\right) \\
&= \left(\sum_{i\in\mathcal{A}} q_i\right)\left(\log\left(\frac{P}{\sum_{i=1}^{m}\frac{1}{b_i}}+1\right) + D(q_{\mathcal{A}}\|m_{\mathcal{A}})\right)
\end{aligned}
$$

where

$$m_{\mathcal{A}i} = \frac{M_{i,i}(\mathcal{A})}{\sum_{i\in\mathcal{A}} M_{i,i}(\mathcal{A})}$$

is the distribution of the minors of $\mathbf{H}_{\mathcal{A}}$

Appendix B

Proofs for Chapter 3

B.1 Proof of Theorem 3.0.4

We now prove the rate at which one can hope to scale channel norms and asymptotically have a non-zero probability. In this direction note that from Alzer's bound [4] we have for $m > 1$

$$\left(1 - e^{-s_l x}\right)^m \le \tilde{\gamma}_{\mathrm{sf}}(m, x) \le \left(1 - e^{-x}\right)^m$$

where $s_l \triangleq \Gamma(1+m)^{-1/m}$ and

$$\tilde{\gamma}_{\mathrm{sf}}(m, x) = \frac{1}{\Gamma(1+m)} \int_0^x t^{m-1} e^{-t} dt$$

So,

$$\begin{aligned} p_s &\ge \left(1 - e^{-s_l \rho^+}\right)^{2m} - \left(1 - e^{-\rho^-}\right)^{2m} \\ &= \sum_{j=0}^{2m} \binom{2m}{j} (-1)^{j+1} \left(e^{j\rho^-} - e^{j s_l \rho^+}\right) \end{aligned}$$

Now, we note that in order for the bound to be non-zero we must have $\rho^- < s_l \rho^+$ so that the probability is non-zero. However, implicit in the proof of the bound given in [4] if we replace the constant s_l in the lower bound by any number $s \in (s_l, 1)$ then there exists a x^* such that

$$\left(1 - e^{-sx}\right)^m \le \tilde{\gamma}_{\mathrm{sf}}(m, x)$$

for all $x \in [x^*, \infty)$. So, asymptotically we can replace the constant s_l by $1 - \epsilon$ for any ϵ such that $1 > \epsilon > 0$.

Now, taking $s < 1$ and $m\rho^+(n) = c\log n$ and $m\rho^-(n) = \log n - d$ yields

$$\begin{aligned} p_s &\geq \sum_{j=0}^{2m} \binom{2m}{j} (-1)^{j+1} \left(e^{-j\log n + jd} - e^{-jcs\log n}\right) \\ &\geq \sum_{j=0}^{2m} \binom{2m}{j} (-1)^{j+1} e^{-j\log n} \left(e^{jd} - e^{-j(cs-1)\log n}\right) \\ &= \sum_{j=0}^{2m} \binom{2m}{j} (-1)^{j+1} n^{-j} \left(e^{jd} - n^{-j(cs-1)}\right) \end{aligned} \tag{B.1}$$

Thus for $cs < 1$ as $n \to \infty$ then

$$np_s \to -\infty$$

Further, for $cs \geq 1$ as $n \to \infty$ then

$$2m(e^d - 1) \leq np_s \leq 2me^d$$

where the lower bound corresponds to $cs = 1$ and the upper bound corresponds to $cs = \infty$. From the above derivation (interchanging the role of s in the upper and lower bound) it should be clear that if $\log(n) = o(\rho^-(n))$, then

$$\lim_{n\to\infty} np_s \to 0$$

B.2 Proof of Lemma 3.3.1

To see this note we can rewrite $a(\theta; \alpha, \beta)$ as

$$a(\theta; \alpha, \beta) = (1 - \cos\theta)^{\alpha-\beta} \sin^{2\beta}\theta$$

and note that the derivative of the function at $\theta = \frac{\pi}{2}$ is $\alpha - \beta$. Thus if $\alpha < \beta$ there is some θ such that the function is greater than 1, since $a(\pi/2; \alpha, \beta) = 1$ and the function is continuous. Otherwise, if $\alpha > \beta$, $a(\theta; \alpha, \beta)$ is the product of functions bound by above by 1 and below by 0 on the interval $(0, \pi/2)$ and thus the function $a(\theta; \alpha, \beta)$ is bounded between 0 and 1.

■ B.3 Proof of Theorem 3.3.2

Again, consider the difference function

$$D_m(\theta; a(\cdot; s, \beta)) \triangleq \int_0^\theta \sin^m \phi \, d\phi - \psi_m a(\theta; s, \beta)$$

and examine a scaled version of the derivative of the difference

$$d_m(\theta; a(\cdot; s, \beta)) \triangleq \frac{1}{\sin^n \theta} \frac{\partial}{\partial \theta} D_m(\theta; a(\cdot; s, \beta)) \tag{B.2}$$

Now, since $D_m(0; a(\cdot; s, \beta)) = D_m(0; a(\cdot; s, \beta)) = 0$, we must have $d_m(0; a(\cdot; s, \beta)) \geq 0$ for $a(\cdot; s, \beta)$ to be a lower bound. Similarly, $d_m(\pi/2; a(\cdot; s, \beta)) \leq 0$ for $a(\cdot; s, \beta)$ to be a lower bound. Now, we have that

$$\lim_{\theta \to 0} d_m(\theta) = 0 \text{ and } d_m\left(\frac{\pi}{2}\right) = 1 - \psi_m(\alpha - \beta) \tag{B.3}$$

So, we must have $\alpha \geq \frac{1}{\psi_m} + \beta$ for $a(\theta; \alpha, \beta)$ to be a lower bound and $\alpha \leq \frac{1}{\psi_m} + \beta$ for $a(\theta; \alpha, \beta)$ to be an upper bound.

We will characterize the α, β pairs such that the derivative of $d_m(\theta; a(\cdot; s, \beta))$ only changes sign once. Now, let

$$g(\theta) \triangleq 2(1 - \cos\theta)^{2-\alpha}(1 + \cos\theta)^{2-\beta} \sin^{m-2}\theta \frac{\partial}{\partial \theta} d_m(\theta; a(\cdot; s, \beta)) \tag{B.4}$$

This is just a scaled version of the derivative of $d_m(\theta; a(\cdot; s, \beta))$ and since the term $2(1 - \cos\theta)^{2-\alpha}(1 + \cos\theta)^{2-\beta} \sin^{m-2}\theta \geq 0$ we have that

$$\operatorname{sign}\left(\frac{\partial}{\partial \theta} d_m(\theta; a(\cdot; s, \beta))\right) = \operatorname{sign}(g(\theta))$$

With some arithmetic is can be shown

$$\begin{aligned} g(\theta) = {} & 3\alpha^2 - \alpha(2 + 2\beta + m) + \beta(3\beta - m - 2) \\ & + 2(\alpha - \beta)(2(\alpha + \beta) - m - 1)\cos\theta + (\alpha + \beta)(\alpha + \beta - m)\cos 2\theta \end{aligned} \tag{B.5}$$

Note, that the function $d + e\cos\theta + f\cos 2\theta$ can have at most one root on the interval $(0, \frac{\pi}{2})$

if d, e and f are fixed. Furthermore,

$$g(0) = \alpha(2\alpha - m - 1) \text{ and } g\left(\frac{\pi}{2}\right) = 2((\alpha - \beta)^2 - \alpha - \beta) \tag{B.6}$$

Since we must have $g(\theta)$ positive at $\theta = 0$ for the function $a(\theta; \alpha, \beta)$ to be a lower bound we must have $\alpha \geq \frac{m+1}{2}$ for $a(\theta; s, \beta)$ to be a lower bound. Similarly, we need $\alpha \leq \frac{m+1}{2}$ for $a(\theta; s, \beta)$ to be an upper bound. We make a similar argument at $\theta = \frac{\pi}{2}$. Let,

$$z(\alpha, \beta) \triangleq (\alpha - \beta)^2 - \alpha - \beta = \frac{1}{2} g\left(\frac{\pi}{2}\right)$$

Then we need $g(\pi/2) \leq 0$ for $a(\theta; \alpha, \beta)$ to a lower bound. Considering the curve where this equation is zero gives

$$\beta_{\pm}(\alpha) \triangleq \frac{1}{2}\left(1 + 2\alpha \pm \sqrt{1 + 8\alpha}\right)$$

Note, that $\beta_{-}(0) = 0$ and $\beta_{+}(0) = 3$. Further, $z(0, 1) = -4$ so we must have $\beta_{-}(\alpha) \leq \beta \leq \beta_{+}(\alpha)$ for $a(\theta; \alpha, \beta)$ to be a lower bound. We summarize this as follows. Let L be the set of all α and β pairs such that $a(\theta; \alpha, \beta)$ is a lower bound. Then,

$$L = \left\{ (\alpha, \beta) \;\middle|\; \alpha \geq \beta + \frac{1}{\psi_m} \text{ and } z(\alpha, \beta) \leq 0 \text{ and } \alpha \geq \frac{m+1}{2} \right\}$$

In a similar fashion we define the set U be the set of all α and β pairs such that $a(\theta; \alpha, \beta)$ is an upper bound. That is,

$$U = \left\{ (\alpha, \beta) \;\middle|\; \alpha \leq \beta + \frac{1}{\psi_m} \text{ and } z(\alpha, \beta) \geq 0 \text{ and } \alpha \leq \frac{m+1}{2} \right\}$$

These regions can be seen plotted in Figure 3-6. Now, note that the function $\beta_{+}(\alpha)$ is concave in α since it is the sum of a square root and linear terms. Thus, since

$$\beta_{+}(0) = 1 \quad \text{and} \quad \beta_{+}((m+1)/2) = \frac{1}{2}\left(2 + m + \sqrt{5 + 4m}\right)$$

the function $\beta_{+}(\alpha)$ intersects the line $\beta = \alpha - \frac{1}{\psi_m}$ for some $\alpha_0 > \frac{m+1}{2}$. A simple calculation additionally yields the function $\beta_{-}(\alpha)$ intersects the line $\beta = \alpha - \frac{1}{\psi_m}$ at $\alpha = \frac{1}{2}(\frac{1}{\psi_m} + \frac{1}{\psi_m^2})$. Taking these points of intersection yields the corollaries to Theorem 3.3.2. The regions of valid α and β pairs can be seen in Figure 3-6 of Section 3.3.

■ B.4 Proof of Theorem 3.3.7

We consider a function that is similar in form to that of the RHS of (3.26). Note that from the example of Section 3.3 we replaced the function θ by a different function that was possibly more convenient for analysis. Now, we always have the elementary inequality $\cos\theta \leq \frac{\sin\theta}{\theta}$. We wish to use this in place of the cosine in (3.26). This motivates the new estimate

$$c(\theta; s, \beta) = \left(\frac{2}{\pi}\theta\right)^s \sin^\beta \theta$$

It is natural to again consider the difference $D_m(\theta; c(\cdot; s, \beta))$ and more importantly the scaled derivative

$$d_m(\theta; c(\cdot; s, \beta)) = 1 - \left(\frac{2}{\pi}\right)^s \psi_m s g(\theta; s, \beta) \tag{B.7}$$

where g is

$$g(\theta; s, \beta) = \sin^{\beta - m}\theta \left(\frac{\beta}{s}\theta^s \cot\theta + \theta^{s-1}\right) \tag{B.8}$$

Now, note that for any β and s

$$\lim_{\theta \to 0} g(\theta; s, \beta) = 0$$

Similarly, for any β

$$\lim_{\theta \to \frac{\pi}{2}} g(\theta; s, \beta) = \left(\frac{2}{\pi}\right)^{1-s}$$

So that the derivative is decreasing if $s \geq \frac{\pi}{2\psi_m}$ at $\theta = \frac{\pi}{2}$ and increasing otherwise. Note that this is precisely the same constant that appears for the upper bound $b_m(\theta, s)$. That is the maximal sequence for the upper bound $b_m(\theta; s)$ coincides with the minimal sequence for the lower bound $c(\theta; s, \beta)$. We will hereafter denote this sequence as

$$s^* \triangleq \frac{\pi}{2\psi_m}$$

We will now show that for $\beta \geq m + 1$ and $s \geq s^*$ there exists a θ^* such that $g(\theta; s, \beta)$ is increasing for $\theta \in (0, \theta^*)$ and decreasing for $\theta \in (\theta^*, \pi/2)$. We do this by showing that

$$h_m(\theta; s, \beta) = \tfrac{s}{b} x^{2-r} (\sin\theta)^{m-\beta} \tan\theta \tfrac{\partial}{\partial\theta} g(\theta; s, \beta) \tag{B.9}$$

$$= h_m^{(1)}(\theta; s, \beta) + h^{(2)}(\theta; s, \beta) \tag{B.10}$$

is concave. Where

$$h_m^{(1)}(\theta; s, \beta) = c(m, s, \beta)\theta + (\beta - m - 1)\theta^2 \cot\theta \tag{B.11}$$

and $c(m, s, \beta)$ is a constant depending on m, s and β and

$$h^{(2)}(\theta; s, \beta) = \frac{s(s-1)}{\beta} \tan\theta - 2\theta^2 \csc(2\theta) + \theta^2 \cot\theta \tag{B.12}$$

Note that since the sum of two concave functions is again concave it is sufficient to show that for all $s \geq s^*$ and $\beta > n+1$ that $h_m^{(1)}(\theta; s, \beta)$ and $h_m^{(2)}(\theta; s, \beta)$ are concave. This is proved through the following lemmas.

Lemma B.4.1. *The function*

$$h^{(1)}(\theta; \alpha) = \alpha\theta^2 \cot\theta$$

is concave for $\alpha > 0$ and convex for $\alpha < 0$.

Lemma B.4.2. *The function*

$$h^{(2)}(\theta; q) \triangleq q \tan\theta - 2\theta^2 \csc 2\theta + \theta^2 \cot\theta$$

is concave for $\theta \in (0, \pi/2)$ if and only if $q \leq \frac{\pi^2}{4}$. Further the function $h^{(2)}$ is convex for $q > \frac{2\pi^2 + 3\pi + 27}{18}$

Additionally, using the requirement that the derivative d_m be negative at $\frac{\pi}{2}$ gives that it is sufficient for $s^* < s < \frac{\pi}{2}\sqrt{m+1} + 1$ for $c(\theta; s, \beta)$ to be a lower bound. Now, since $s^* < \sqrt{\frac{\pi}{2}}\sqrt{m+1}$ we have that such an s exists for all m.

Now, we can arrive at an upper bound in a similar fashion. Indeed, the function $h_m(\theta; s, \beta)$ is convex if[1] we have that the functions $h_m^{(i)}(\theta; s, \beta)$ are convex and $s \leq s*$.

[1] We upper bound the constant $\frac{2\pi^2+3\pi+27}{18}$ by π. One can obtain a slightly better constant by better approximating the smallest q for which the the function $l(\theta; q)$ is positive for all θ.

Combining these one can see that the minimal set of conditions for the upper bound is

$$s \leq \frac{\pi}{2\psi_m} \text{ and } \beta \leq \frac{(s-1)s}{\pi} \tag{B.13}$$

B.4.1 Proof of Lemma B.4.1

We have that

$$\begin{aligned} l(\theta, \alpha) &= \frac{\sin^2\theta \cot\theta}{2\alpha} \frac{\partial^2}{\partial\theta^2} h^{(1)}(\theta; \alpha) && \text{(B.14)} \\ &= \cos^2\theta - 2\theta\cot\theta + \theta^2\cot^2\theta && \text{(B.15)} \end{aligned}$$

Using that $\theta \cot\theta > \cos\theta$ yields

$$\begin{aligned} l(\theta, \alpha) &\leq 2\theta^2\cot^2\theta - 2\theta\cot\theta && \text{(B.16)} \\ &= 2\theta\cot\theta\,(\theta\cot\theta - 1) && \text{(B.17)} \\ &\leq 0 && \text{(B.18)} \end{aligned}$$

B.4.2 Proof of Lemma B.4.2

Now, it is sufficient to show that

$$\begin{aligned} l(\theta; q) &\triangleq -\cos\theta \frac{\partial^2}{\partial\theta^2} h^{(2)}(\theta; q) && \text{(B.19)} \\ &= 4\theta + \tan\theta\left(1 - 2q + 2\theta^2 + \cos(2\theta)\right) && \text{(B.20)} \end{aligned}$$

is greater than zero iff $q \leq \frac{\pi^2}{4}$. First note that

$$\frac{\partial}{\partial q} l(\theta; q) = -2\tan\theta < 0 \tag{B.21}$$

so $l(\theta; r)$ is decreasing in q. It is sufficient to show that $l(\theta; (\pi/2)^2) > 0$. Now,

$$\begin{aligned} 1 - 2q + 2\theta^2 + \cos(2\theta) &\geq -2q + 2\theta^2 && \text{(B.22)} \\ &\geq -2q + 2\left(\frac{\pi}{2}\right)^2 && \text{(B.23)} \\ &\geq 0 \quad \left(\text{if } q \leq \left(\frac{\pi}{2}\right)^2\right) && \text{(B.24)} \end{aligned}$$

so that $l(\theta; q) > 0$ for all $q \leq \left(\frac{\pi}{2}\right)^2$. Conversely suppose that $h^{(2)}(\theta; q)$ was concave for some $q \geq \frac{\pi^2}{4}$. Then,

$$\lim_{\theta \to \frac{\pi}{2}} l(\theta; q) < 0$$

which contradicts the concavity of $h^{(2)}(\theta; q)$. Now, to see that the function is convex for $q > \frac{2\pi^2+3\pi+27}{18}$ note that it is sufficient for this to be true for $\theta = \frac{\pi}{3} = \arg\min 4\theta + \tan\theta$. Some simple arithmetic show that this implies that $l(\theta; \pi/3) < 0$

List of Symbols

$c_m(\mathcal{A};\theta)$ 31
union of spherical cap of half angle θ centered at points of $\mathcal{A}$.

m 15
number of transmit antennas.

n 15
number of single antenna receivers.

$p_{\epsilon,l}$ 31
probability any set is ϵ-orthogonal .

$p_{\perp}$ 31
conditional probability any set is ϵ-orthogonal given it meets norm constraint.

p_s 29
probability any channel vector meets norm constraints ρ^+ and ρ^-.

$\mathbf{H}_{\mathcal{A}}$ 16
the channel matrix of the set of users $\mathcal{A}$.

$\mathbf{W}_{\mathcal{A}}$ 21
the Wishart matrix of the channel matrix defied by the users in $\mathcal{A}$.

$\mathcal{C}_{\text{sp}}(\mathbf{H}, P)$ 20
the set of all rate vectors achievable by beamforming under a power constraint P.

$\mathcal{C}_{\text{zf}}(\mathbf{H}, P)$ 22
the set of all rate vectors achievable by zero-forcing multiplexing under a power constraint P.

Bibliography

[1] M. Abramowitz and I. Stegun, *Handbook of Mathematical Functions*, Dover, New York, 1972.

[2] M. Airy, S. Shakkottai, and R. W. H. Jr, *Spatially greedy scheduling in multi-user MIMO wireless systems*, in Proc. of IEEE Asilomar Conf. on Signals, Systems, and Computers, Pacific Grove, CA, November 2003.

[3] D. Aktas and H. El Gamal, *Multiuser scheduling for multiple antenna systems*, in Proc. IEEE Vehicular Tech. Conf., vol. 3, Orlando, USA, October 2003, pp. 1743–1747.

[4] H. Alzer, *On some inequalities for the incomplete Gamma function*, Mathematics of Computation, 66 (1997), pp. 771–778.

[5] G. Andrews, R. Askey, and R. Roy, *Speacial Functions*, no. 71 in Encyclopedia of Mathematics and its Applications, Cambridge University Press, Cambridge, 1999.

[6] D. P. Bertsekas, *Convex Analysis and Optimization*, Athena Scientific, Belmont,Massachusetts, 2003.

[7] B. Bollobás, *Modern Graph Theory*, no. 184 in Graduate Texts in Mathematics, Springer-Verlag, New York, 1998.

[8] ——, *Random Graphs*, no. 73 in Cambridge studies in advanced mathematics, Cambridge University Press, Cambridge, second ed., 2001.

[9] G. Caire and S. Shamai (Shitz), *On the achievable throughput of a multi-antenna Gaussian broadcast channel*, IEEE Transactions on Information Theory, 49 (2003), pp. 1691–1706.

[10] M. Costa, *Writing on dirty paper*, IEEE Transactions on Information Theory, 29 (1983), pp. 439–441.

[11] T. M. Cover and J. Thomas, *Elements of Information Theory*, John Wiley and Sons Inc., New York, 1997.

[12] P. Fernandes, L. T. Berger, J. Mrtires, and P. Kyritsi, *Effects of multi user MIMO scheduling freedom on cellular downlink system throughput*, in Proc. IEEE 60th Vehicular Technology Conference, Los Angeles, USA, September 2004.

[13] G. J. Foschini, *Layered space-time architectures for wireless communication in a fading environment when using multi-element antennas*, Bell Labs. Tech. J., 1 (1996), pp. 41–59.

[14] C. Godsil and G. Royle, *Algebraic Graph Theory*, no. 207 in Graduate Texts in Mathematics, Springer-Verlag, New York, 2001.

[15] B. Hochwald, T. L. Marzetta, and V. Tarokh, *Multi-antenna channel-hardening and its implications for rate feedback and scheduling*, IEEE Transactions on Information Theory, (2004). to appear.

[16] S. Janson, *Large deviations for sums of partly dependent random variables*, Random Structures Algorithms, 24 (2004), pp. 234–248.

[17] S. Janson, T. Luczak, and A. Rucinski, *An exponential bound for the probability of nonexistence of a specified subgraph in a random graph*, in Random Graphs '87, John Wiley and Sons, Inc., 1990, pp. 73–87.

[18] ——, *Random Graphs*, John Wiley, New York, 2000.

[19] G. A. Kabatyanskii and V. I. Lenenshtein, *Bounds for packings on a sphere and in space*, Problems Inform. Transm., 14 (1978), pp. 1–17.

[20] M. Lopez, *Multiplexing, Scheduling and Multicasting Strategies for Antenna Arrays in Wireless Networks*, PhD thesis, MIT, Cambridge, Massachusetts, 2002.

[21] K. Marton, *A coding theorem for the discrete memoryless broadcast channel*, IEEE Transactions on Information Theory, 25 (1979), pp. 306–311.

[22] M. Neely, E. Modiano, and C. Rohrs, *Power allocation and routing in multibeam satellites with time-varying channels*, IEEE/ACM Transactions on Networking, 11 (2003), pp. 138–152.

[23] M. Penrose, *Random Geometric Graphs*, no. 5 in Oxford Studies in Probability, Oxford University Press, Oxford, 2003.

[24] S. Serbetli and A. Yener, *Time-slotted multiuser MIMO systems: Beamforming and scheduling strategies*, EURASIP Journal on Wireless Communications and Networking, (2004), pp. 286–296.

[25] C. E. Shannon, *Probability of error for optimal codes in a Gaussian channel*, Bell System Technical Journal, 38 (1959).

[26] D. M. Y. Sommerville, *The Geometry of n Dimensions*, Dover Publicatons, Inc., New York, 1958.

[27] C. Swannack, E. Uysal-Biyikoglu, and G. W. Wornell, *Low complexity multiuser scheduling for maximizing throughput in the MIMO broadcast channel*, in Proc. 42nd Annual Allerton Conf. on Communication, Control, and Computing, Monticello, Illinois, 2004.

[28] L. Tassiulas and A. Ephremides, *Stability properties of constrained queueing systems and scheduling policies for maximum throughput in multihop radio networks*, IEEE Transactions on Automatic Control, 37 (1992), pp. 1936–1948.

[29] I. E. Teletar, *Capacity of multi-antenna Gaussian channels*, Eur. Trans. Telecom., 10 (1999), pp. 585–595.

[30] D. Tse and P. Viswanath, *On the capacity of the multiple antenna broadcast channel*, in Multiantenna Channels: Capacity, Coding and Signal Processing, vol. 62 of DIMACS eries in Discrete Mathematics and Theoretical Computer Science, Providence, 2003, American Mathematical Association, pp. 87–105.

[31] ——, *Fundamentals of Wireless Communication*, Cambridge University Press, Cambridge, 2005.

[32] D. N. C. Tse and S. V. Hanly, *Multiacces fading channels–Part I: Polymatroid structure and throughput capacities*, IEEE Transactions on Information Theory, 44 (1998).

[33] P. Vishwanath and D. N. C. Tse, *Sum capacity of the vector Gaussian broadcast channel and uplink-downlink duality*, IEEE Transactions on Information Theory, 49 (2003), pp. 1912–1921.

[34] S. Vishwanath, N. Jindal, and A. Goldsmith, *Duality, achievable rates, and sum-rate capacity of Gaussian MIMO broadcast channels*, IEEE Transactions on Information Theory, 49 (2003).

[35] ——, *On the duality of Gaussian multiple-access and broadcast channels*, IEEE Transactions on Information Theory, 50 (2004), pp. 768–783.

[36] E. Visotsky and U. Madhow, *Optimum beamforming using transmit antenna arrays*, in Proc. 1999 IEEE 49th Vehicular Technology Conference, vol. 1, May 1999, pp. 851–856.

[37] H. Weingarten, Y. Steinberg, and S. Shamai (Shitz), *The capacity region of the Gaussian MIMO broadcast channel*, in Proc. CISS, Princeton University, March 2004.

[38] A. D. Wyner, *Random packings and coverings of the unit n-sphere*, BSTJ, 46 (1967), pp. 2111–2118.

[39] E. Yeh and A. Cohen, *Information theory, queueing, and resource allocation in multi-user fading communications*, in Proc. CISS, Princeton University, March 2004.

[40] W. Yu, *Uplink-downlink duality via minimax duality*, IEEE Transactions on Information Theory, (2004). submitted.