RICE UNIVERSITY

Capacity of Low Power Multiuser Systems With Antenna Arrays

by

Tarik Muharemović

A THESIS SUBMITTED
IN PARTIAL FULFILLMENT OF THE
REQUIREMENTS FOR THE DEGREE
Doctor of Philosophy

APPROVED, THESIS COMMITTEE:

Behnaam Aazhang, Chair
J. S. Abercrombie Professor
Electrical and Computer Engineering

Ashutosh Sabharwal, Co-Chair
Faculty Fellow
Electrical and Computer Engineering

Don H. Johnson
J. S. Abercrombie Professor
Electrical and Computer Engineering

Richard A. Tapia
Noah Harding Professor
Computational and Applied Mathematics

Houston, Texas

October, 2004
INFORMATION TO USERS

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleed-through, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

UMI®

UMI Microform 3168108
Copyright 2005 by ProQuest Information and Learning Company.
All rights reserved. This microform edition is protected against unauthorized copying under Title 17, United States Code.

ProQuest Information and Learning Company
300 North Zeeb Road
P.O. Box 1346
Ann Arbor, MI 48106-1346
Capacity of Low Power Multiuser Systems With Antenna Arrays

Tarik Muharemović

Thesis: Doctor of Philosophy
Electrical and Computer Engineering
Rice University, Houston, Texas (October 2004)
Capacity of Low Power Multiuser Systems With Antenna Arrays

Tarik Muharemović

Abstract

In this thesis, we study wireless multiuser communication systems in the regime of low spectral efficiencies, where users and the multiple access point are equipped with antenna arrays. Our first contribution is to develop a generic mathematical framework which captures tradeoffs between fundamental parameters of a low power multiuser system: spectral efficiency and energy per information bit, of each user.

Using the framework that we developed we next consider variable data rate multiple access problem, in low power systems, where we remove the usual assumption of tight user coordination, and we allow users to select their own data rates and transmit powers, without coordinating, and without negotiating with the access point. Here, every user has a set of low power codebooks, that we name the policy, which accommodates a range of small spectral efficiencies, but particular data rates of other users are assumed to be an unknown - compound parameter - at each mobile. In antenna-array transmission and reception, we demonstrate an elegant interpretation of users policies, where each policy is represented by partitioning spatial dimensions into blocks, and each block is dedicated to a different user.

Finally, we address the paradigm of statistically correlated antenna arrays, where we derive the effective number of uncorrelated receive spatial dimensions, which we
partition to represent users policies. As more correlated antennas are packed into a limited area we show that effective receive dimensionality converges to a finite limit which we evaluate for some simple geometries.
Acknowledgments

I am a man of few words ©. I would like to thank my advisor Behnaam for constant interaction, ability to lift the spirits, and most importantly, for great insight into fundamentals of the multiuser communication problem. Also, I would like to thank my Co-advisor Ashu for countless hours of help on this work, and for his extensive effort to teach me, and other students from our group, to think outside of the box. I would also like to thank Professors from Departments of ECE, MATH and CAAM for providing great course materials, which give me a toolbox to tackle this problem. From ECE department I learned fundamentals of Detection Theory and Information Theory from Dr. Don Johnson and Dr. Elza Erkip. From CAAM department I learned Optimization Theory from Dr. Richard Tapia, and from the MATH department I would like to thank Dr. Stephen Semmes and Dr. William Veech for teaching me Real Analysis. Also, I would like to thank my buddies from the trenches of Grad students for constant interaction and mutual support: Jober, Kiran, Sinan, Nasir, Chris, Jacob, Nadeem, Amir, Arnab, Muharem, Sridhar, Vikram, Coskun and Scott Clay. Finally, my special thanks go to my family: wife Amelia, my parents Alija and Nadza and my siblings Adnan and Aida.
Contents

Abstract ii
Acknowledgments iv
List of Illustrations viii
List of Tables x

1 Introduction 1
   1.1 Low Power Antenna-Array Multiuser Communication 1
   1.2 Coarsely Coordinated Multiuser Systems 2
   1.3 Antenna Packing: The Space Limit to Dimensionality 4

2 System Description 6
   2.1 Channel Model 6
   2.2 Notation 6
   2.3 Tightly Coordinated Multiuser Systems 8
   2.4 Coarsely Coordinated Multiuser Systems 11
   2.5 Appendix: Chapter Proofs 15

3 Background on Low Power Systems 18
   3.1 Low Power Systems with Infinite Bandwidth 18
   3.2 Low Power Systems with Finite Bandwidth 21
   3.3 Policy and Payoff in Low Power Systems with Finite Bandwidth 22
3.4 Chapter Summary ............................................. 23
3.5 Appendix: Chapter Proofs ............................. 23

4 Tightly Coordinated Low Power Multiuser Systems 28
4.1 Multiuser Detection Receiver and Tightly Coordinated Users .... 30
4.2 Time Division .............................................. 32
4.3 Single User Detector Receiver and Tightly Coordinated Users .... 34
4.4 Chapter Summary ......................................... 36
4.5 Appendix: Chapter Proofs ............................. 37

5 Antenna Arrays in Low Power Multiuser Systems 42
5.1 Spatially Correlated Codebooks ......................... 46
5.2 Chapter Summary ......................................... 49
5.3 Appendix: Chapter Proofs ............................. 49

6 Coarsely Coordinated Low Power Multiuser Systems 60
6.1 Coarsely Coordinated Users with Multiuser Detection Receiver .... 60
6.2 Coarse User Coordination With Antenna Arrays ................ 65
6.3 Coarsely Coordinated Users with Single User Detection Receiver .... 79
6.4 Appendix: Chapter Proofs ............................. 81

7 Comparison of Multiple Access Strategies 88
7.1 Tightly Coordinated Users: TDMA vs. SUD(R) .................. 88
7.2 Coarsely Coordinated Users: TDMA vs. MUDₜ .................. 90
7.3 Appendix: Chapter Proofs ............................. 92
8 System Design

8.1 Receiver Antenna Packing: Fundamental Limits for the Effective Degrees of Freedom ........................................ 94

8.2 Limiting Degrees of Freedom ........................................ 96
8.2.1 Limiting Degrees of Freedom: Properties ....................... 97
8.2.2 Limiting Degrees of Freedom of Circular Array .............. 98
8.2.3 Small-Sized Circular Arrays ..................................... 101
8.2.4 Approximation for Moderate-Sized Circular Arrays .......... 101

8.3 Antenna Array Design Examples .................................. 103
8.3.1 Numerical Procedures ......................................... 103
8.3.2 Linear Arrays ..................................................... 105
8.3.3 Square Arrays .................................................... 108

8.4 Bandwidth Requirements of A Low Power Multiple Access System .. 109
8.4.1 Frequency Division ............................................ 110
8.4.2 Time Division .................................................... 112
8.4.3 Superposition .................................................... 113

8.5 Optimal Spreading Sequences for Symmetric Policies ............ 116

8.6 Appendix: Chapter Proofs .......................................... 118

9 Conclusions and Future Work ........................................ 130

A Appendix: Additional Facts and Proofs ............................ 132

Bibliography ................................................................ 134
Illustrations

2.1 Linear Multiuser Channel Model .......................... 7
2.2 Rate Region of Two User Channel ...................... 9
2.3 Achievable Rate Regions For Various Multiaccess Strategies 10
2.4 Policy of User 1 is the set of CodeBooks of User 1. .......... 12

4.1 Slope regions for 2 Users. MUD($\mathcal{R}$), TDMA and SUD($\mathcal{R}$). ........ 33

6.1 Robust Slope Region MUD$_*$ for single antenna transmission. .... 62
6.2 Subspace projections at the receiver establish parallel channels. .... 68
6.3 Policies Chosen From MUD$_*$ Represent Dedicated Channels. .... 70
6.4 Characteristics Points on the Robust Slope Region for 4 Receive Antennas ........................................ 73
6.5 Dedicated Channels - Point A on MUD$_*$ from Figure 6.4 ........ 74
6.6 Dedicated Channels - Point B on MUD$_*$ from Figure 6.4 ........ 75
6.7 Dedicated Channels - Point C on MUD$_*$ from Figure 6.4 ........ 76

7.1 Slope regions for 2 Users with 2 Receive Antennas. .......... 89

8.1 Antenna Packing ........................................ 95
8.2 Effective Degrees of Freedom vs. the Number of Antennas Uniformly
Packed on a Circle of Radius $r$. .............................................. 99
8.3 Limiting Degrees of Freedom and the Linear Approximation. ... 102
8.4 Effective Degrees of Freedom vs. Number of Antennas in a Uniform
Linear Array of Length $d$ ...................................................... 105
8.5 Locations of $M = 6$ Antennas in a Linear Array of Size $d = 4\lambda$. . 107
8.6 Locations of $M = 7$ Antennas in a Linear Array of Size $d = 4\lambda$. . 107
8.7 Locations of $M = 8$ Antennas in a Linear Array of Size $d = 4\lambda$. . 107
8.8 Locations of $M = 9$ Antennas in a Linear Array of Size $d = 4\lambda$. . 107
8.9 Effective Antenna-Space vs. Number of Antennas in a Uniform
Square Array of Length $d$ ...................................................... 109
8.10 Locations of $M = 4$ Antennas on a Square Array of Size $\equiv \frac{\lambda}{2}$ . 110
8.11 Locations of $M = 9$ Antennas on a Square Array of Size $\equiv \lambda$ . . . 111
8.12 An illustration which is used for the proof of Theorem 3 . . . . . . . 122
# Tables

8.1 Comparison of Uniform and Non-Uniform Linear Arrays. . . . . . . . . 106

8.2 Non-Uniform Antenna Locations in a Linear Array of Size \( d = 4\lambda \) . . 106

8.3 A Comparison of Uniform and Non-Uniform Square Arrays . . . . . . . . . 108

8.4 Non-Uniform Antenna Locations on a Square Array . . . . . . . . . . . . . 108
Chapter 1

Introduction

A study of achievable communication limits, for wideband low power systems, is motivated by a variety of applications in Ultrawideband, Satellite and Sensor Networks. Here, low power communication is chosen or imposed externally for numerous reasons, such as FCC regulations for Ultrawideband, communication distance for Satellite, and prolonged battery life for Sensor networks. The common characteristic for all such systems, is that users operate with extremely low received powers, and with very low spectral efficiencies.

1.1 Low Power Antenna-Array Multiuser Communication

As a first contribution of this thesis, we provide a generic theory of low power multiple access systems, which deals with the multiuser, multiantenna problem, in the regime of low spectral efficiencies.

In wideband low power systems, the aggregate system bandwidth requirement, along with multiple access constraints on spectral efficiency, are derived by considering spectral efficiency as a function of energy per information bit $E$. Here, spectral efficiency originates from the minimum achievable value for energy per information bit $E_{\text{min}}$, with an initial slope - the slope of spectral efficiency. Two main parameters, which are considered to be information theoretic limits of low power systems, are therefore minimum energy per information bit $E_{\text{min}}$, and the slope $S$, of spectral
efficiency.

The minimum energy per information bit, for each user, is the same in single and multiuser systems, because it is achieved only at zero spectral efficiencies. In multiuser systems, constraints on spectral efficiency of each user are best described by the multiaccess spectral efficiency slope region, which depends on various design choices for the multiple access strategy and the receiver structure, be it time division, or superposition of users signals with single or with multiuser detectors. The first contribution of this thesis is to systematically derive these information theoretic limits of communication, for multiuser systems with transmit/receive antenna arrays.

1.2 Coarsely Coordinated Multiuser Systems

The second contribution of this thesis is to analyze information theoretic limits of low power multiuser systems, where users do not coordinate their choices for data rate - coarsely coordinated multiple access systems.

Namely, the theory of multiuser systems is commonly developed under the assumption that a central agent exercises tight control over users' transmission parameters, which typically includes data rate and power allocation for each user. We will refer to such systems as tightly coordinated multiuser systems. In data networks, which are characterized by bursty multiple access interference, and by delay bounded traffic, tight user coordination necessitates that users re-negotiate their data rates and transmit powers with the base-station, very frequently throughout the call duration. For instance, practically any change in transmission parameters (data rate or power) of only one user, necessitates data rate and power adaptation for all other users. This occurs because multiple access interference changes, from the perspective of all other users. Tight user coordination is therefore often costly and difficult
to achieve even in high power systems, and low power transmission exacerbates the problem as control channels must be strongly encoded, with large contention delays.

Second contribution of this thesis is departure from this classic framework, and a study of communication limits of low power, multiple access systems, where such tight user coordination may not be feasible or desirable. Furthermore, we demonstrate how the advent of multiuser detection, coupled with antenna arrays, entirely removes the requirement of tight user coordination. The quest of ultra-wideband systems has mostly been directed towards cheap and fairly uncoordinated multiple access to large amounts of bandwidth, for a large number of users [6], and in this work we establish a perspective on multiaccess limits, that are simply unavoidable.

We will approach this problem by assuming that user coordination is exercised at a very coarse level: each user is granted a set of codebooks, which accommodate a range of data rates and corresponding transmit powers, also including complete silence. Particular set of codebooks awarded to a given user we will name the policy. Each user locally selects its own data rate, and power to meet the data rate requirements, by choosing the appropriate codebook, out of the entire set.

Any given user is assumed to be totally oblivious to particular codebooks chosen by other users, and therefore, users are oblivious to data rate, and transmit power of others. Each set of codebooks must be carefully designed, such that multiple access capacity constraints are not violated under any circumstances. We label such systems as coarsely coordinated multiuser systems. This problem is as a multiple access compound channel problem, having an unknown parameter at each mobile: data rate and power of other users.

It is apparent that standard orthogonal multiple access schemes, like TDMA or FDMA, trivially enable coarse user coordination: base-station allocates time or
frequency slots, and users can select their rates/power, based on established single-user capacity limits. We consider the following question: is there is a multiple access scheme, which achieves higher spectral efficiencies than orthogonal systems, while allowing local data rate adaptation by each user?

We answer the question in affirmative, and we first provide a representative example of such scheme, for single antenna, multiple access channels. Antenna-array transmission and reception is addressed in the regime of low spectral efficiencies, where we provide the complete family of solutions to the coarsely coordinated multiuser problem. In multi-antenna systems, we give a geometric interpretation of users policies via virtual partition of the receiver space dimension.

1.3 Antenna Packing: The Space Limit to Dimensionality

The third contribution of this thesis is to provide a variety of Low Power System Design Guidelines, where the Limit to Antenna Packing in the bounded space is both theoretical and practical highlight.

As more antennas are packed into a limited space, signals which they measure become correlated, and the communication dimensionality does not correspond to the aggregate number of receivers any more. Based on analysis of information theoretic limits, which we establish in the first two contributions of the thesis, we derive a quantity $\omega$, which depends only on the receiver relative antenna positioning, and where $\omega$ represents the equivalent number of uncorrelated receive dimensions.

Then we consider the problem of having an abundance of cheap receivers, which are to be positioned inside a bounded space. Here, we are not constrained by the number of receivers, but rather by the locality $G$ that they ought be placed. The least upper bound to $\omega$, over all finite antenna configurations inside $G$ represents the
space limitation to receiver dimensionality.

First, we demonstrate that space dimensionality is finite, after which we evaluate it for some finite geometries. For instance, placing 8 antennas on a circumference of a circle, which has radius 1/2 of the carrier wavelength, will result in the equivalent of 6 receive dimensions, after which there is no point in packing more antennas. Similarly, if radius of the circle is 2 carrier wavelengths, then 30 antennas on the circumference provide an equivalent of 20 receive dimensions, which is the point of saturation.
Chapter 2

System Description

2.1 Channel Model

We consider the linear $K$-user multiple access channel, with complex additive white Gaussian noise (AWGN), namely

$$Y = \sum_{k=1}^{K} H_k X_k + \eta,$$  \hspace{1cm} (2.1)

where vector $X_k$, of user $k$ complex symbols, is modulated by the matrix channel $H_k$, and superimposed with signals from other users. Vector $\eta$ represents independent and identically distributed complex AWGN, which is normalized to unit variance per each complex dimension.

Channel model (2.1) encompasses standard multi-antenna transmission and reception techniques, such as space-time codes and MIMO. Also, with appropriate labeling, it may be used to account for users with spreading codes, or even frequency selective channels. Channel Model (2.1) is illustrated in Figure 2.1.

2.2 Notation

In this thesis, we will consider users with variable data rates $R_k$, and variable transmit powers $P_k$. To be able to effectively address variable power transmission, we will establish a notation, which explicitly labels the mutual information between inputs $X_k$ and output $Y$, to be a function of users adjustable transmit powers.
By $\bar{A}$, we denote the complement of set $A \subseteq \{1, \ldots, K\}$, and mutual information between $X_A$ and $Y$ is denoted

$$C_A(P_A) \triangleq I(X_A; Y|X_{\bar{A}} = 0),$$

where $P_A = [P_k]_{k \in A}$ represents the vector of transmit powers for users inside the set $A$, and also $X_A = [X_k]_{k \in A}$. The family of mutual informations $C_A(P_A)$ is implicitly conditioned upon all $\mathcal{H}_k$, because receiver is assumed to be aware of all channel coefficients. At times, we will find it convenient to suppress $A$, when $A = \{1, 2, \ldots, K\}$, and we will simply write $C(P)$, instead of $C_A(P_A)$. Notice that $C(P)$ maps $\mathbb{R}^K \to \mathbb{R}$.

One concrete example of this family $C_A(P_A)$ is given for $K = 2$ users, in multi-
antenna transmission and reception

\[ C_{\{1,2\}}(P_1, P_2) = \mathbb{E} \log \det \left[ I + \mathcal{H}_1 \mathcal{H}_1^\dagger \frac{P_1}{N_1} + \mathcal{H}_2 \mathcal{H}_2^\dagger \frac{P_2}{N_2} \right] \]  
\[ (2.3) \]

\[ C_{\{1\}}(P_1) = C_{\{1,2\}}(P_1, 0) \]  
\[ (2.4) \]

\[ C_{\{2\}}(P_2) = C_{\{1,2\}}(0, P_2), \]  
\[ (2.5) \]

where user 1 has \( N_1 \) transmit antennas, while user 2 has \( N_2 \) transmit antennas.

To make the mutual information in (2.2) always well defined, we assume that the exact probability distribution on \( X_k \) is uniquely determined by users \( k \) transmit power \( P_k \). For example, in single antenna systems, we always assume \( X_k \) to be zero mean complex Gaussian, with variance \( P_k \). Another example is the two user multi-antenna system (2.3), where we assumed that both vectors \( X_1 \) and \( X_2 \) are independent and identically distributed, and mutual information (2.3) is uniquely specified by powers \( P_1 \) and \( P_2 \).

### 2.3 Tightly Coordinated Multiuser Systems

Multiple access systems are traditionally studied under assumption of tight user coordination, where each user adjusts its transmit power or data rate, in accordance to the multiple access interference. Such systems always require centralized system controller - even when a single user alters its data rate, all other users must re-adjust to different interference level.

The standard mathematical tool to study these systems is the multiuser achievable rate region, which is defined by each user transmit power. Because family of mutual informations \( C_A(P_A) \) is a function of users adjustable transmit powers, so is
achievable rate region [17] of $K$-user channel

$$CMUD = \left\{ (R_1, \ldots, R_K) : \forall A \subseteq \{1, \ldots, K\} \sum_{k \in A} R_k \leq C_A(P_A) \right\}, \quad (2.6)$$

which becomes $2^K - 1$ faced polymatroid, when all transmit powers $P_k$ are fixed.

Rate region (2.6), for a particular pair of transmit powers $P_1$ and $P_2$ is illustrated in Figure 2.2, and its is achieved provided that receiver utilizes the optimum multiuser detection. Figure 2.2 is commonly known as the Cover-Vyner Pentagon [17], and operating points inside the Pentagon in 2.2 require that users coordinate the
choice of codebooks, as illustrated by point B in Figure 2.2.

If the receiver is composed of a bank of single-user detectors, the rate region is

$$\text{CSUD} = \left\{ (R_1, \ldots, R_K) : \forall k \in \{1, \ldots, K\} \quad R_k \leq C(P) - C_k(P_k) \right\}, \quad (2.7)$$

Formula (2.7) comes from the fact that interference cancelation achieves vertices of the rate region (2.6). Relationship between the rate region CSUD and CMUD is illustrated in Figure 2.3.
For time division multiple access strategy (TDMA), the rate region is given as

\[
\text{CTDMA} = \bigcup_{\sum \alpha_k = 1} \left\{(R_1, \ldots, R_k) : \forall k \in \{1, \ldots, K\}, \frac{R_k}{\alpha_k} \leq C_{\{k\}} \left(\frac{P_k}{\alpha_k}\right)\right\}. \tag{2.8}
\]

Here, \(\sum \alpha_k\) represents a partition of a time unit, where time slot \(\alpha_k\) is allocated to users \(k\) transmission. Constraints inside (2.8) have a simple interpretation: because \(C_{\{k\}}\) is a single-user capacity function, during the time slot \(\alpha_k\) only user \(k\) transmits with data rate \(R_k\alpha_k^{-1}\) and power \(P_k\alpha_k^{-1}\), so that average data rate and transmit power of user \(k\) remains \(R_k\) and \(P_k\).

### 2.4 Coarsely Coordinated Multiuser Systems

In coarsely coordinated multiuser systems, each user is allowed to select its own data rate, and transmit power locally, without negotiation with the multiaccess point, and without any knowledge about data rate or transmit power of other users. In order to accommodate variable data rate traffic, each user is provided with a set of codebooks. One should contrast this with classic multiple access problem, where each user has only one codebook, fixed transmit power, and fixed data rate.

**Definition 1 (Policy in Coarsely Coordinated Systems).** The set of codebooks, which are designated to a particular user, and which accommodate various data rates and corresponding average transmit powers, in a one-to-one fashion.

Each users’ set of codebooks, namely the policy, is characterized by data rate versus transmit power relationship for that particular set: from lower data rate-lower power codebooks, to higher data rate-higher power codebooks. This functional characteristic represents benefit versus cost, namely the payoff function, which is designated to a particular user. An illustration of User 1 Policy and the Payoff function is given in Figure 2.4.
Figure 2.4: Policy of User 1 is the set of CodeBooks of User 1.

Each policy has only one payoff function, and the payoff will be used to compare coarse multiple access strategies: superior strategies provide better payoffs. Payoff in coarsely coordinated multiuser system is the analogue to data rate for a tightly coordinated multiuser system - or equivalently, in a system where transmit powers are fixed.

Definition 2 (Payoff Function of a Policy). Date rate versus power $R(P)$ characteristic of a particular policy.
Even though payoff is only one characteristic for the set of codebooks, which are awarded to a particular user, we will use the payoff to specify the entire set, namely the policy. For example, every codebook within the policy of user k is chosen i.i.d. complex Gaussian, with transmit power \( P_k \), and data rate \( R_k(P_k) \), where \( R_k(P_k) \) is the payoff function of user k. As \( P_k \) varies, the choice of payoff function entirely specifies how to construct the policy of user k.

In the same manner that data rate of a particular user restricts data rates of other users, which is described by the rate region, the same holds for payoffs. Payoff awarded to a particular user in a coarsely coordinated multiple access system, will limit payoffs awarded to other users. To demonstrate this, we provide an example of set of policies and payoffs, in regime of moderate and high spectral efficiencies.

**Example 1.** Consider 2-user, single antenna multiaccess channel (2.9), where each users' fading coefficient obeys the same \( p(h) \) distribution, and aggregate capacity function is given as (2.10), namely

\[
Y = h_1X_1 + h_2X_2 + \eta
\]

\[
C_{\{1,2\}}(P_1, P_2) = \mathbb{E} \log [1 + |h_1|^2P_1 + |h_2|^2P_2].
\]

Assume that channel statistics, such as distribution of fading coefficients and noise, remain the same throughout the call duration, but users adjust their data rates locally, each according to its own policy, which is negotiated only once, before any transmission begins.

**Following is the pair of payoff functions**

\[
R_1(P_1) = \frac{1}{2} C_{\{1,2\}}(P_1, P_1) = \frac{1}{2} \mathbb{E} \log [1 + (|h_1|^2 + |h_2|^2)P_1]
\]

\[
R_2(P_2) = \frac{1}{2} C_{\{1,2\}}(P_2, P_2).
\]
Each codebook of transmit power $P_1$, from the policy of user 1, is drawn from Gaussian distribution with data rate $R_1(P_1)$. Notice that $R_1$ is not a direct function of $h_2$ or $h_1$, but rather it only depends on their distribution $p(h)$, through the expectation operator $E$. Here, the payoff function $R_1$ is a direct function only of $P_1$, and vice-versa. Similarly, user 2 selects its policy from the function $R_2(P_2)$.

Once policies are selected, users are free to adjust their data rates and transmit powers in accordance to (2.11) and (2.12), where each user is totally oblivious to data rate and transmit power of other users. The pair $(R_1, R_2)$ is guaranteed to always belong inside the Cover-Wyner pentagon, which becomes defined only once $(P_1, P_2)$ is chosen by users locally. This property of (2.11) and (2.12) always guarantees reliable transmission.

Proof. See Appendix.

Using the convexity property of the aggregate capacity function (2.10), the Appendix first shows that all (three) multiple access capacity constraints are satisfied, for all combination of users transmission rates and powers. Second, the Appendix also shows that the solution (2.11) and (2.12) is superior to time division.

Namely, if we were to divide a time unit into equal portions, each user would have to follow the payoff function which is defined by time division: transmit only half of the time with twice the power, and keep silent otherwise. Payoff function of each user specified by time division, is inferior to (2.11) or (2.12)

$$R_{1}^{TDMA}(P_1) = \frac{1}{2} C_{(1,2)}(2P_1, 0)$$

$$\leq \frac{1}{2} C_{(1,2)}(P_1, P_1) = R_1(P_1),$$

which means that users can achieve superior spectral efficiencies than TDMA, even if they do not coordinate the selection of their codebooks.
To specify payoffs (2.11) and (2.12), we essentially performed guess-and-check. In order to guarantee reliable transmission, it does not suffice to consider a single rate region, which depends on a particular instance of users transmit powers. When considering users which vary their transmit powers and data rates locally, we must verify that multiple access constraints on users data rates remain valid, for all combinations of users transmit powers. In essence, for each possible combination of $P_1$ and $P_2$, we had to verify that multiaccess rate-region is not violated by $R_1$ and $R_2$.

We will follow the same principle in the regime of low spectral efficiencies, where our goal is to provide a systematic construction of all available payoff functions, and policies.

2.5 Appendix: Chapter Proofs

Proof of Example 1. First we demonstrate that policies from (2.11) and (2.12) satisfy all three multiple access constraints. We begin with single user constraints, which are justified by the following chain

\[
1 + (|h_1|^2 + |h_2|^2)P_1 \leq [1 + |h_1|^2P_1] [1 + |h_2|^2P_1]
\]

(2.15)

\[
E \log [1 + (|h_1|^2 + |h_2|^2)P_1] \leq E \log [1 + |h_1|^2P_1] + E \log [1 + |h_2|^2P_1]
\]

(2.16)

Here, (2.16) is reached by first taking the log, and then taking the average. In order to proceed from (2.16), we notice the following two facts. First, two terms on the right hand side of (2.16) are identical, because probability distributions on $h_1$ and $h_2$ are the same. Second, the left hand side of (2.16) is in effect $C_{(1,2)}(P_1, P_1)$. Hence
(2.16) reduces to

\[ C_{(1,2)}(P_1, P_1) \leq 2 \mathbb{E} \log [1 + |h_1|^2 P_1] \]

\[ R_1(P_1) \leq \mathbb{E} \log [1 + |h_1|^2 P_1] \]  

\[ R_1(P_1) \leq C_{(1)}(P_1) \]  

Here, (2.18) is simply by definition (2.11) of users 1 policy and the payoff function \( R_1(P_1) \). Also, (2.19) is by observation that \( C_{(1)}(P_1) = C_{(1,2)}(P_1, 0) \), which is in effect the single-user capacity constraint. By symmetry of the problem, similar argument to (2.15)-(2.19) will hold for user 2 as well, and its single-user capacity constraint.

Because single-user constraints have been verified, next we address the joint multiple access constraint. Straightforward computation provides that (2.20) is equivalent to \( 2P_1P_2 \leq P_1^2 + P_2^2 \), which is true because \( 2P_1P_2 \leq P_1^2 + P_2^2 \Leftrightarrow (P_1 - P_2)^2 \geq 0 \) is true, and we have that

\[
[1 + (|h_1|^2 + |h_2|^2)P_1] \left[ 1 + (|h_1|^2 + |h_2|^2)P_2 \right] \\
\leq [1 + |h_1|^2 P_1 + |h_2|^2 P_2] \left[ 1 + |h_1|^2 P_2 + |h_2|^2 P_1 \right] \\
\Rightarrow \log [1 + (|h_1|^2 + |h_2|^2)P_1] + \log [1 + (|h_1|^2 + |h_2|^2)P_2] \\
\leq \log [1 + |h_1|^2 P_1 + |h_2|^2 P_2] + \log [1 + |h_1|^2 P_2 + |h_2|^2 P_1],
\]

Where (2.21) is simply by taking the logarithm. To proceed from (2.21), we will evaluate the average with respect to fading distribution. The left-hand side of (2.21) is simply expressed using the payoff function of users, as

\[ 2R_1(P_1) + 2R_2(P_2) \leq \mathbb{E} \log [1 + |h_1|^2 P_1 + |h_2|^2 P_2] + \mathbb{E} \log [1 + |h_1|^2 P_2 + |h_2|^2 P_1] \]  

\[ R_1(P_1) + R_2(P_2) \leq \mathbb{E} \log [1 + |h_1|^2 P_1 + |h_2|^2 P_2] \]  

\[ R_1(P_1) + R_2(P_2) \leq C_{(1,2)}(P_1, P_1). \]
Here, (2.23) follows from (2.22) because we assumed that $h_1$ and $h_2$ obey the same probability distribution. Finally, (2.24) is simply the joint capacity constraint for users 1 and 2, which is now verified.

Finally, to complete the claims made in this Example, we verify that policies (2.11) and (2.12) are superior to time division. This will be performed using the concavity ∩ property of $C_{1,2}(P_1, P_2)$, as a function of the vector $(P_1, P_2)$. Notice that with TDMA the payoff function is given as

$$ R_{1}^{TDM} (P_1) = \frac{1}{2} C_{1,2}(2P_1, 0) $$  

$$ = \frac{1}{2} \left[ \frac{1}{2} C_{1,2}(2P_1, 0) + \frac{1}{2} C_{1,2}(0, 2P_1) \right] $$  

$$ \leq \frac{1}{2} C_{1,2}(P_1, P_1) = R_1(P_1) $$  

Again, (2.26) uses the fact that $h_1$ and $h_2$ follow the same probability distribution. Also, (2.27) is true because of concavity of the function $C_{1,2}$, and average value of arguments $(2P_1, 0)$ and $(0, 2P_1)$ equals $(P_1, P_1)$.

From the inequality in (2.27), policies (2.11) and (2.12) are superior to time division, and this completes the two claims made in the Example 1.
Chapter 3

Background on Low Power Systems

In wideband low power systems, both spectral efficiency $R_k$, which is measured in [bits]/[sec×Hz], and transmit power* $P_k$, measured in [Joules]/[sec×Hz], are very small, which is described by the limiting process $R_k \to 0$ and $P_k \to 0$. For this reason, energy per information bit, which is defined by the ratio

$$
\mathcal{E}_k \triangleq \frac{P_k}{R_k} \left[ \frac{\text{Joules}}{\text{bit}} \right],
$$

(3.1)

emerges as a natural metric, which is used to describe the performance of wideband low power systems. The goal of power-efficient communication is to operate at low values of (3.1).

Low power systems are classified in two groups. The first group is infinite bandwidth systems, whose only target is minimizing (3.1), and they have no other considerations. Second group is classified as large-but-finite bandwidth low power systems, where spectral efficiency is also considered, and approximated, as a function of (3.1).

3.1 Low Power Systems with Infinite Bandwidth

Theory of wideband low power systems [1] was originally developed with the assumption of infinite bandwidth, where both spectral efficiency $R_k$, and transmit power $P_k$

* $P_k$ is actually proportional to transmit power in [Watts], where the constant of proportionality is system bandwidth. Perhaps a better name for $P_k$ would be “energy density.” However, we don’t wish to redefine standard terminology.
are exactly zero. Here, the limiting behavior of energy per information bit (3.1), as
$P_k \to 0$, is the only metric, which serves to describe and evaluate the system per-
formance. Assuming multiple users, following theorem establishes achievable lower
bound, for energy per information bit $E_k$.

**Theorem 1 ([1]).** Let $C_{(k)}(P_k)$ be the single-user capacity function for the channel
model (2.1). Also, we denote

$$
\dot{C}_{(k)}(0) \triangleq \frac{d}{dP_k} \left[ C_{(k)}(P_k) \ln 2 \right]_{P_k=0}
$$

(3.2)

to be the first derivative of $C_{(k)}(P_k)$, measured in [nats], and evaluated at $P_k = 0$.

Assuming that (3.2) is non-zero, any $K$-tuple of energies per information bit,
within the following rectangle, is achievable

$$
\left\{ (E_1, \ldots, E_K) : \ E_k \geq \frac{\ln 2}{C_{(k)}(0)} \right\}.
$$

(3.3)

Rectangle (3.3) is achievable by following combinations of multiple access, and
receiver detection strategies: a) superposition\(^\dagger\) of users signals, along with multiuser
detectors at the base-station, b) superposition of users signals, along with with single-
user detectors at the base-station, and c) time division.

Conversely, any point outside of rectangle (3.3) is not achievable by any multi-
access strategy.

**Proof.** See Appendix.

Theorem 1 concludes that for infinite bandwidth systems, there are no joint
multiple access constraints on $E_k$, because $C_{(k)}(P_k)$ from (3.3) represents a single-
user capacity function. Furthermore, theorem 1 concludes that the lower bound for

\(^\dagger\)by superposition we mean that users have perfectly correlated signature waveforms, with no
spectrum spreading
\( \mathcal{E}_k \), namely (3.3), is achievable by all described combination of multiple access and receiver strategies. This lower bound (3.3) is defined as the “minimum energy per information bit”

\[
\mathcal{E}_{\text{min},k} \triangleq \frac{\ln 2}{\hat{C}_{(k)}(0)}. \tag{3.4}
\]

In fact, \( \hat{C}_{(1)}(0) \) has a physical interpretation of the average power gain, which is introduced by the communication channel, and that is the conventional reason for defining derivative in (3.2) and (3.4) in [nats], as opposed to [bits]. This is demonstrated in the following example.

**Example 2.** Consider a single user, single-antenna system, where we have that \( C_{(1)}(P_1) = \log [1 + |h_1|^2 P_1] \), and by (3.2) we have

\[
\hat{C}_{(1)}(0) = \frac{d}{dP_1} \left[ \ln [1 + |h_1|^2 P_1] \right]_{P_1=0} = E[|h_1|^2]. \tag{3.5}
\]

By comparing \( E[X_1^2 X_1] = P_1 \), which is the average transmitted power, from user 1, in a single antenna system, with \( E[(h_1 X_1)^4 (h_1 X_1)] = P_1 E[|h_1|^2] \), which is the average received power, from user 1, we see that \( \hat{C}_{(1)}(0) = E[|h_1|^2] \) from (3.5) represents the channel power gain, relevant for user 1. The same principle holds in antenna array transmission and reception, which we demonstrate later - also see [2].

Finally, it is common to express the minimum energy per information bit from (3.4), in decibel scale, as

\[
10 \log_{10} \mathcal{E}_{\text{min},k} = -1.59 \text{dB} - 10 \log_{10} \hat{C}_{(1)}(0), \tag{3.6}
\]

which holds because \( 10 \log_{10}(\ln 2) = -1.59 \text{dB} \). Here, the quantity \(-1.59 \text{dB}\) is referred to as the “minimum received energy per information bit,” and is commonly used as a benchmark, to evaluate the performance of practical coding strategies, by normalizing the channel gain to unity.
3.2 Low Power Systems with Finite Bandwidth

If we were to regard $\mathcal{E}_k$ as the only metric, which is relevant for low power systems, the design choice for spectral efficiency would be exactly zero. Namely, minimizing $\mathcal{E}_k$ results in $\mathcal{E}_{\text{min},k}$, which is achieved only at $R_k = 0$. Low power systems with finite bandwidth, are defined as systems which operate at small, but non-zero spectral efficiencies. Here, considering only $\mathcal{E}_k$ is insufficient.

In low power systems with finite bandwidth [2], it is natural to also consider the spectral efficiency $R_k$, but now as a function of energy per information bit $\mathcal{E}_k$. This function $R_k(\mathcal{E}_k)$ cannot be evaluated in closed-form, but the low power assumption allows us to approximate the spectral efficiency, by the rate of increase of $R_k$, as $\mathcal{E}_k$ departs away from $\mathcal{E}_{\text{min},k}$. Notice that we toggle our focus between $R_k(P_k)$ and $R_k(\mathcal{E}_k)$, which are two different functions, with different domains; nevertheless, both functions take values inside the same set of spectral efficiencies, and we will often suppress the independent variable.

The rate of increase of spectral efficiency is defined in [2], in decibel scale, as

\[
S_k \triangleq \lim_{\mathcal{E}_k \to \mathcal{E}_{\text{min},k}} \frac{R_k}{10 \log_{10} \mathcal{E}_k - 10 \log_{10} \mathcal{E}_{\text{min},k}} \left[ \frac{10 \log_{10} 2}{\text{bits}} \right] \quad (3.7)
\]

\[
= \lim_{\mathcal{E}_k \to \mathcal{E}_{\text{min},k}} \frac{R_k}{\log_2 \mathcal{E}_k - \log_2 \mathcal{E}_{\text{min},k}} \left[ \frac{\text{sec} \times \text{Hz} \times 3\text{dB}}{\text{bits}} \right]. \quad (3.8)
\]

Reflecting on the unit of $S_k$, 3dB appears because of $\log_2$. For example, if $\mathcal{E}_k$ doubles from $\mathcal{E}_{\text{min},k}$, operating point moves 3 dB away from $\mathcal{E}_{\text{min},k}$, and spectral efficiency is approximated simply with $S_k$. Approximations of this nature are found to be excellent in [2,10], and spectral efficiency in low power systems is approximated as

\[
R_k = S_k [\log_2 \mathcal{E}_k - \log_2 \mathcal{E}_{\text{min},k}] + \epsilon, \quad (3.9)
\]

which represents the tradeoff between spectral efficiency and energy per information
bit.

In low power systems, $\mathcal{E}_{\text{min},k}$ and $\mathcal{S}_k$ may be evaluated by knowing a) properties of the communication channel, for example, single antenna Rayleigh fading channel, and by knowing b) the actual signalling strategy, for example, Gaussian codebooks. Given the multi-user, correlated multi-antenna channel, we will evaluate both $\mathcal{E}_{\text{min},k}$ and $\mathcal{S}_k$, for various signalling strategies, in order to find the tradeoff (3.9), between $R_k$ and $\mathcal{E}_k$, in low power systems.

Tradeoff between spectral efficiency and energy per information bit (3.9) has several important practical implications. First, it clearly serves to approximate the spectral efficiency in low power systems. Second, it may be used to approximate bandwidth requirements for a low power system, given users with constant data rates in [bits]/[sec] and powers in [Joules]/[sec]. Third, we will show how to utilize (3.9), to allow for coarse coordination between users, in dynamic data rate environment. Because of these practical implications, our primary focus will be to evaluate achievable $\mathcal{E}_{\text{min},k}$ and $\mathcal{S}_k$.

### 3.3 Policy and Payoff in Low Power Systems with Finite Bandwidth

Tradeoff (3.9) between spectral efficiency $R_k$ and energy per information bit $\mathcal{E}_k$ is also very convenient to represent the policy of user $k$, in coarsely coordinated multiple access systems. The policy was defined to be the set of codebooks which belong to user $k$, and user $k$ adjusts its data rate, by selecting the appropriate codebook from the entire set.

For a fixed $\mathcal{S}_k$ and $\mathcal{E}_{\text{min},k}$, (3.9) conveniently relates variable data rate of a particu-
lar user \( k \), with its variable energy per information bit. Equivalently, this relationship is between transmit power and data rate, which is the payoff function.

**Proposition 1 (Payoff of a Policy).** If \( S_k \) is non-zero, using (3.1) and (3.8), transmit power satisfies

\[
P_k = \mathcal{E}_{\min,k} R_k + \mathcal{E}_{\min,k} R_k^2 \left( \frac{\ln 2}{S_k} \right) + o(R_k^2).
\]

(3.10)

**Proof.** Manipulation of infinitesimals. See Appendix. □

Given the payoff function (3.10), with \( S_k \) and \( \mathcal{E}_{\min,k} \) fixed, each codebook from users \( k \) policy is chosen from complex Gaussian distribution, independent and identically distributed. Here, each codebook of data rate \( R_k \) has transmit power \( P_k \), which are related as (3.10). Hence, construction of a policy of user \( k \) is fully described, by fixed \( S_k \) and \( \mathcal{E}_{\min,k} \), as in (3.10).

### 3.4 Chapter Summary

Spectral efficiency in low power systems is approximated as (3.9). Theorem 1 summarizes the set of \( \mathcal{E}_{\min,k} \), which are achievable in low power multiple access system.

To fully characterize the set of achievable spectral efficiencies, we must provide a characterization for the set of achievable slopes \( S_k \), which is the target of the next section, where we consider tightly coordinated users - or equivalently - constant data rate transmission.

### 3.5 Appendix: Chapter Proofs

**Proof of Theorem 1.** First we show the converse part, which is the fact that anything outside of the rectangle (3.3) is not achievable. Here, we show that for reliable transmission, \( \mathcal{E}_k \) must always satisfy (3.3).
Because the single-user capacity function $C_{\{k\}}(P_k) \ln 2$ is concave ∩, it must be dominated by its first-derivative approximation at $R_k = 0$, formally

$$C_{\{k\}}(P_k) \ln 2 \leq C_{\{k\}}(0) \ln 2 + \dot{C}_{\{k\}}(0)(P_k - 0), \quad (3.11)$$

where the factor $\ln 2$ appears, because derivative $\dot{C}_{\{k\}}(0)$ is defined after $C_{\{k\}}(P_k)$ is converted into [nats], by multiplying with $\ln 2$. The single-user capacity constraint, which simply states that $R_k \leq C_{\{k\}}(P_k)$, is combined with (3.11), and we have

$$R_k \ln 2 \leq \dot{C}_{\{k\}}(0) P_k. \quad (3.12)$$

$$\frac{\ln 2}{\dot{C}_{\{k\}}(0)} \leq \frac{P_k}{R_k} \quad (3.13)$$

$$\frac{\ln 2}{C_{\{k\}}(0)} \leq \mathcal{E}_k. \quad (3.14)$$

Here, (3.12) - (3.14) are simple algebraic manipulations, and (3.14) completes the converse part of the theorem.

For the direct part of the theorem, which states that all points inside the rectangle (3.3) are achievable, we first consider single user-detector receivers. To demonstrate the principle of (3.3), it suffices to consider $K = 2$ users, because proof is the same if $K > 2$. Proof of (3.3) will be done in two consecutive stages.

The first stage is simply to evaluate the achievable data rate $R_1$, of user 1, assuming single-user detectors are employed at the base-station. To solve for $R_1$, we suppose interference cancelation was used, and users were ordered as $1 \rightarrow 2$. Now we observe the following: because user 1 is the first in the decoding process, his performance is effectively the same, as if single user detectors were utilized, while user 2 experiences the “interference clean” channel and therefore $R_2 = C_{\{2\}}(P_2)$. From the joint multi-access constraint, we have that spectral efficiencies of users 1 and 2 must
satisfy

\[ R_1 + R_2 \leq C_{\{1,2\}}(P_1, P_2). \]  
(3.15)

\[ R_1 + C_{\{2\}}(P_2) \leq C_{\{1,2\}}(P_1, P_2). \]  
(3.16)

The second stage evaluates the energy per information bit, for user 1, which is achieved from (3.16). Here, we first multiply through by \(\ln 2\), and then we utilize the first-order Taylor expansion on functions \(C_{\{2\}}(P_2) \ln 2\), and \(C_{\{1,2\}}(P_1, P_2) \ln 2\), evaluated at \(P_1 = 0\) and \(P_2 = 0\), as follows

\[
R_1 \ln 2 + \frac{dC_{\{2\}}(P_2) \ln 2}{dP_2} \bigg|_0 P_2 + o(P_2) \leq \left. \frac{\partial C_{\{1,2\}}(P_1, P_2) \ln 2}{\partial P_1} \right|_0 P_1 + \left. \frac{\partial C_{\{1,2\}}(P_1, P_2) \ln 2}{\partial P_2} \right|_0 P_2 + o(||[P_1 P_2]||),
\]  
(3.17)

where we have abbreviated \(\big|_0\) to represent evaluating all derivatives at the origin.

Because partials on the right hand side of (3.17) are with respect to only one variable, and also evaluated at the origin, we may simply substitute them with derivatives \(\dot{C}_{\{1\}}(0)\) and \(\dot{C}_{\{2\}}(0)\). Then, we cancel the common term on both sides, and we are left with

\[
R_1 \ln 2 + o(P_2) \leq \dot{C}_{\{1\}}(0) P_1 + o(||[P_1 P_2]||) \]  
(3.18)

\[
\ln 2 + \left[ \frac{o(P_2) + o(||[P_1 P_2]||)}{R_1} \right] \leq \dot{C}_{\{1\}}(0) \left[ \frac{P_1}{R_1} \right].
\]  
(3.19)

We claim that the bracketed term on the left hand side of (3.19) can be made arbitrarily small, by selecting small values for \(P_1\) and \(P_2\).

Because ratio \(P_1/R_1\) is bounded (3.14) below by a real number, then any quantity which is negligible (in the jargon infinitesimal) with respect to \(P_1\), like \(o(P_1)\), must also be infinitesimal with respect to \(R_1\). Because of this, Given \(\mathcal{E}_1\), one can choose
\( P_1 \) and \( P_2 \) small enough, so that the bracketed term on the left of (3.19) becomes negligible. Hence, in (3.19) we replace \( \leq \) by \( < \) we have that

\[
\ln 2 < \frac{\ln 2}{\hat{C}_{(1)}(0)} \mathcal{E}_1
\]

\[
\ln 2 < \frac{\ln 2}{\hat{C}_{(1)}(0)} \mathcal{E}_1,
\]

which achieves \( \mathcal{E}_1 \), that is arbitrary close to the bound (3.21). The second stage of the achievability proof, with single user detectors, is now complete.

Since we proved achievability of rectangle (3.3) with single-user detectors, then the same rectangle is also achievable with multiuser detectors, which are a superset of single-user detectors. We are left only to provide a proof for time division.

With time division, if \( \alpha_1 \) is the time slot which is dedicated to user 1, the generic constraint on data rate is given as \( \alpha_1 R_1 \leq \hat{C}_{(1)}(\alpha_1 P_1) \), from where it follows

\[
\alpha_1 R_1 \ln 2 \leq \left. \frac{d\hat{C}_{(1)}(\alpha_1 P_1) \ln 2}{dP_1} \right|_0 P_1 + o(\alpha_1 P_1).
\]

\[
\alpha_1 R_1 \ln 2 \leq \alpha_1 \left. \frac{d\hat{C}_{(1)}(P_1) \ln 2}{dP_1} \right|_0 P_1 + o(P_1).
\]

The result follows by canceling \( \alpha_k \) from both sides of (3.23), and then following the same steps as before to deal with the infinitesimal term \( o(P_1) \). This completes the proof of Theorem 1. \( \square \)

**Proof of Proposition 1.** To begin with, we suppress the users k subscript, to simplify the notation. Namely, it’s there, but we don’t write it, because its not relevant.

Using (3.1) and (3.8), ‘the slope’ admits an alternate expression

\[
S = \lim_{R \downarrow 0} \frac{R}{\log_2 \left( \frac{P}{R \varepsilon_{\min}} \right)^k},
\]

where we also wrote \( R \downarrow 0 \) instead of \( \varepsilon_k \downarrow \varepsilon_{\min,k} \) because it is the same limiting process, and both numerator and denominator of (3.24) approach zero.
We will first tackle denominator of (3.24) as follows

\[
\log_2 \left( \frac{P}{R \mathcal{E}_{\text{min}}} \right) = \log_2 \left( 1 + \frac{P - R \mathcal{E}_{\text{min}}}{R \mathcal{E}_{\text{min}}} \right) = (\log_2 e) \frac{P - R \mathcal{E}_{\text{min}}}{R \mathcal{E}_{\text{min}}} + o \left( \frac{P - R \mathcal{E}_{\text{min}}}{R \mathcal{E}_{\text{min}}} \right). \tag{3.25}
\]

Because \( S \neq 0 \), then the \( o \)-term inside the denominator (3.26) is also infinitesimal with respect to numerator \( R \), and for this reason we can express (3.24) as

\[
S = \lim_{R \to 0} \frac{R}{(\log_2 e) \frac{P - R \mathcal{E}_{\text{min}}}{R \mathcal{E}_{\text{min}}} + o(R)} = \lim_{R \to 0} \frac{R}{(\log_2 e) \frac{P - R \mathcal{E}_{\text{min}}}{R \mathcal{E}_{\text{min}}}} = \lim_{R \to 0} \frac{R^2 \mathcal{E}_{\text{min}} \ln 2}{P - R \mathcal{E}_{\text{min}}}. \tag{3.28}
\]

Now we solve for \( P \), as

\[
P = \frac{R^2 \mathcal{E}_{\text{min}} \ln 2}{P - R \mathcal{E}_{\text{min}}} \tag{3.30}
\]

\[
P = R \mathcal{E}_{\text{min}} + R^2 \left( \frac{\mathcal{E}_{\text{min}} \ln 2}{S} \right) + (P - R \mathcal{E}_{\text{min}}) o(R), \tag{3.31}
\]

\[
P = \mathcal{E}_{\text{min}} R + \mathcal{E}_{\text{min}} R^2 \left( \frac{\ln 2}{S} \right) + o(R^2), \tag{3.32}
\]

which is the basic claim of Proposition 1.

To complete the proof, we only need to clarify transition (3.31) to (3.32), which is done in two steps. Step 1 claims that \( (P - R \mathcal{E}_{\text{min}}) = o(R) \), which is true because

\[
\lim_{R \to 0} \frac{P - R \mathcal{E}_{\text{min}}}{R} = \mathcal{E}_{\text{min}} - \mathcal{E}_{\text{min}} = 0. \tag{3.33}
\]

Step 2 simply states that \( o(R) o(R) = o(R^2) \), which is straightforward. \( \Box \)
Chapter 4

Tightly Coordinated Low Power Multiuser Systems

The objective of this section is to find the set of jointly achievable \( S_k \), in order to evaluate the tradeoff (3.9) between spectral efficiency \( R_k \) and energy per information bit \( E_k \), in tightly coordinated low power multiuser system. Results provided here generalize single user discussion from [2], in a broader sense than initially provided for two-user systems in [11]. This section will also lay out a series of results, which we will exploit to allow for later coarse user coordination.

For single user systems, the maximum achievable slope of spectral efficiency is established in [2], and it depends on the local behavior of the single-user capacity function \( C_{(1)}(P_1) = C(P) \) at \( P_1 = 0 \), namely

\[
S_1 \leq \frac{2 \left[ \frac{\dot{C}_{(1)}(0)}{-\ddot{C}_{(1)}(0)} \right]^2}{\Delta S_{1,\text{su}}},
\]

where “su” stands for single-user systems. To evaluate the maximum achievable \( S_1 \) in a closed-form, we only require knowledge of the first two derivatives of \( C_{(1)}(P_1) \), and not necessarily the entire \( C_{(1)}(P_1) \).

Extending the formula (4.1), in order to account for multiple access environment, will naturally result in a region of achievable \( S_k \), which will have to be defined in terms of partials of \( C(P) \), because \( C(P) \) maps \( \mathbb{R}^K \rightarrow \mathbb{R} \). We will describe three extensions of (4.1), depending on the specific type of multiple access, and receiver strategies: single user detectors, multi-user detectors, or time division.
We will show below, that regardless of the multiple access strategy, the set of jointly achievable slopes \( \mathcal{S}_k \) is described by constraints which always involve the following matrix, that generalizes the right hand side of (4.1). Because of its' reusability and importance, we immediately take the opportunity to establish its definition.

**Definition 3. Wideband Matrix** \( \mathbf{C} \) for a \( K \)-user multiaccess channel is a \( K \times K \) matrix, which we define as follows

\[
\mathbf{C} \triangleq \frac{1}{2} \text{diag}[\nabla C(0)]^{-1} \left[ -\nabla^2 C(0) \right] \text{diag}[\nabla C(0)]^{-1},
\]

(4.2)

where \( \nabla^2 C(0) \) is the Hessian matrix of second partials, and \( \nabla C(0) \) is the gradient of \( C \), both evaluated at 0.

All partials in (4.2) are evaluated after converting \( C(P) \) to [nats], and operator ‘diag’ produces a diagonal matrix from the gradient vector \( \nabla C(0) \). Finally, by \( \mathbf{C}_A \) we denote the Wideband Matrix, which is restricted to both rows and columns, only from the set \( A \).

Definition (4.2) simply states that entries of \( \mathbf{C} \) are negative second order partials, and cross-partials, of \( C(P) \), which are further divided by the product of corresponding first partials of \( C(P) \), and \( \mathbf{C} \) is therefore symmetric. Also, \( \mathbf{C} \) is non-negative definite, because \( C(P) \) is a log-function, and therefore concave with \( \nabla^2 C(0) \leq 0 \). More importantly, \( \mathbf{C} \) in general depends only physical channel characteristics, such as fading distribution and number of antennas, because all dependency of \( \mathbf{C} \) on transmit power \( P \) has been removed by evaluating its partials at \( P = 0 \). Optional constraints, such as discrete alphabet or a set of spreading codes, may be implicitly incorporated in \( \mathbf{C} \), as part of physical channel characteristics.
Notice that (4.2) reduces to inverse of the right hand side of (4.1), for single user case, when $C(P) : \mathbb{R} \rightarrow \mathbb{R}$. Since $\mathbf{c}_A$ equals the Wideband Matrix for users within $A$, when all other users are silent, setting $A = \{k\}$, one can see that diagonal entries of $\mathbf{c}$ are relevant for each individual single user channel: they are $\mathcal{S}_{k,\text{su}}^{-1}$ spectral efficiency slopes, as if no other users were present.

4.1 Multiuser Detection Receiver and Tightly Coordinated Users

Given the Wideband Matrix $\mathbf{c}$, we next generalize expression (4.1), in order to account for multiple access environment, by finding the set of achievable spectral efficiency slopes $\mathcal{S}_k$. Here, we assume that the receiver employs multiuser detection, where the rate-region is given as (2.6).

**Theorem 2 (Slope Region of Multiuser Detection).** Let $\Sigma$ represent the diagonal matrix of users spectral efficiency slopes

$$\Sigma \triangleq \text{diag}(\mathcal{S}_1 \ldots \mathcal{S}_k),$$

(4.3)

and let $\Sigma_A$ be the submatrix of $\Sigma$, maintaining only rows and columns within $A$. Furthermore, let $\mathcal{R} \triangleq [\mathcal{R}_1 \ldots \mathcal{R}_k]^T$ such that $\|\mathcal{R}\| = 1$, be the direction vector inside the $K$-dimensional space of actual transmission rates, which originates from zero, and similarly, denote by $\mathcal{R}_A$ the sub-vector of $\mathcal{R}$ with indices from subset $A$.

Slope region achieved by multiuser detection $\text{MUD}(\mathcal{R})$ is given as

$$\text{MUD}(\mathcal{R}) = \left\{ \Sigma : \mathcal{R}_A^T [\Sigma_A^{-1} - \mathbf{c}_A] \mathcal{R}_A \geq 0, \forall A \subseteq \{1, \ldots, K\} \right\}.$$  

(4.4)

**Proof.** Follows from Proposition 1 and multi-dimensional Taylor expansion of generic capacity constraint (2.6). See Appendix for more details. $\square$
Notice that (4.4) reduces to (4.1), for single user systems, where \( A = \{1\} \). For multiple users, region \( \text{MUD}(\mathcal{R}) \) captures the important tradeoff (3.9) between spectral efficiency and energy per information bit, by specifying the set of achievable spectral efficiency slopes \( S_k \). Therefore, \( \text{MUD}(\mathcal{R}) \) may be utilized to either approximate spectral efficiency, or to solve for minimum bandwidth requirements of a multiple access system*. The fact that \( \text{MUD}(\mathcal{R}) \) is relevant for tightly coordinated multiuser systems, is given by its explicit dependence on the direction vector of spectral efficiencies \( \mathcal{R} \). For this reason, operating points inside \( \text{MUD}(\mathcal{R}) \) are valid for static data rate environment, or equivalently, they require user coordination. Namely, users are not free to adjust their data rates at will.

Given Theorem 2, it becomes straightforward to evaluate the set of achievable \( S_k \), for a low power multiuser system, because all it takes is evaluation of \( C \). Even when a closed-form expression for \( C(P) \) is mathematically intractable, evaluating its partials at \( P = 0 \), to find \( C \), may be quite simple, as shown in the following example.

**Example 3 (Wideband Matrix and Slope Region).** Consider single-antenna channel model (2.9), which is found in Example 1. The Wideband Matrix for the channel model (2.9) is found by evaluating partials of the aggregate capacity function, which is done in Appendix

\[
C = \frac{1}{2} \begin{bmatrix}
\frac{E|h_1|^4}{|E|h_1|^2} & 1 & 1 \\
1 & \frac{E|h_2|^4}{|E|h_2|^2} & \kappa(h_1) \\
1 & \kappa(h_2) & 1
\end{bmatrix}
= \frac{1}{2} \begin{bmatrix}
\frac{1}{S_{1,\text{su}}} & \frac{1}{S_{2,\text{su}}} \\
\frac{1}{2} & \frac{1}{2}
\end{bmatrix}. \quad (4.5)
\]

---

*Given users with constant data rates in [bits]/[sec], and constant transmit powers in [Joules]/[sec], direction vector \( \mathcal{R} \) is determined. From the other side, if we fix system bandwidth \( W \), spectral efficiency slopes \( \Sigma \) are also determined from (3.8). The "minimum required bandwidth," is the smallest \( W \), for which multiple access constraints (4.4) on \( \Sigma \) remain satisfied.
Here, $\kappa(h_1)$ and $\kappa(h_2)$ are kurtosis of fading distributions\(^1\).

The slope region $\text{MUD}(\mathcal{R})$ is specified by three constraints: first $A = \{1\}$, then $A = \{2\}$, and finally $A = \{1, 2\}$. We expand quadratic forms (4.4), using (4.5)

$$\text{MUD}(\mathcal{R}) = \left\{ \Sigma : \begin{array}{c} \mathcal{S}_1^{-1} - \mathcal{S}_{1,\text{su}}^{-1} \geq 0 \\ \mathcal{S}_2^{-1} - \mathcal{S}_{2,\text{su}}^{-1} \geq 0 \\ \mathcal{R}_1^2 \left[ \mathcal{S}_1^{-1} - \mathcal{S}_{1,\text{su}}^{-1} \right] + \mathcal{R}_2^2 \left[ \mathcal{S}_2^{-1} - \mathcal{S}_{2,\text{su}}^{-1} \right] - \mathcal{R}_1 \mathcal{R}_2 \geq 0 \end{array} \right\}.$$  \hspace{1cm} (4.6)

(4.7)

(4.8)

The above example (4.6 - 4.8) of two-user slope region may also be found in [3], although in a different format. However, using the $C$ matrix, we will be able to address much more intricate channel models, which involve multiantenna transmission and reception, as well as an optional set of spreading codes.

In Figure 4.1, we plotted the $\text{MUD}(\mathcal{R})$ slope region from Example 3, for a specific proportion between users data rates, namely $\mathcal{R}_1 : \mathcal{R}_2 = 0.3 : 0.7$. To plot Figure 4.1, we assumed that fading distributions have kurtosis which satisfy $\kappa(h_1) = \kappa(h_2) = 4/3$. Single-user constraints, such as (4.6), are therefore given as $\mathcal{S}_k \leq \mathcal{S}_{k,\text{su}} = 2\kappa^{-1}(h_k) = 3/2$. The the joint multiple access constraint (4.8) in Figure 4.1 is curved, whereas $\text{MUD}(\mathcal{R})$ is the entire region in between the three constraints.

### 4.2 Time Division

Time division multiple access (TDMA) strategy possess the following two properties, which make it an attractive design alternative, for a low power multiple access system. First, time division enables a simple receiver structure, which does not require multiuser detection, and second, time division trivially enables both coarse and tight user coordination.

---

\(^1\)Here we do not require fading coefficients to obey the same distribution, unlike in Example 1
In order to compare the performance of time division, against other multiple access strategies, we extend the single user formula (4.1), into multiple access setting, by evaluating the set of spectral efficiency slopes, which are achieved by TDMA. We represent the set of achievable $S_k$, in terms of the Wideband matrix $\mathcal{E}$.

**Theorem 3 ([3] Slope Region of TDMA).** Let $S_{k,su}$ be the spectral efficiency slope, which is achieved by user $k$, provided that other users are silent. Notice that $S_{k,su}^{-1}$ may be found on the diagonal of the Wideband Matrix $\mathcal{E}$. 

**Figure 4.1:** Slope regions for 2 Users. MUD($\mathcal{R}$), TDMA and SUD($\mathcal{R}$).
Set of spectral efficiency slopes, which are achieved by TDMA is given as

\[ \text{TDMA} = \left\{ (S_1, \ldots, S_K) : \sum_{k=1}^{K} S_k S_{k,su}^{-1} \leq 1 \right\} \]  \hspace{1cm} (4.9)

\[ = \left\{ \Sigma : \text{tr} \left( \Sigma^{1/2} \mathbf{C} \Sigma^{1/2} \right) \leq 1 \right\}, \]  \hspace{1cm} (4.10)

where \( \text{tr} \) denotes the trace operator, which acts on a matrix, by summing its diagonal entries. Two formats for TDMA, namely (4.9) and (4.10) are equivalent, because of the fact that \( S_{k,su}^{-1} \) comprises the diagonal of \( \mathbf{C} \).

Proof. Partition the time unit into slots of length \( \alpha_k \), where slot \( \alpha_k \) is dedicated to the transmission of user \( k \) data.

To achieve the aggregate spectral efficiency of \( R_k \), user \( k \) transmits with \( R_k \alpha_k^{-1} \), during its transmission interval \( \alpha_k \), and keeps silent otherwise, namely

\[ S_k = \lim_{\varepsilon_k \downarrow \varepsilon_{\min,k}} \frac{R_k}{\log_2 \varepsilon_k - \log_2 \varepsilon_{\min,k}} \]  \hspace{1cm} (4.11)

\[ = \alpha_k \lim_{\varepsilon_k \downarrow \varepsilon_{\min,k}} \frac{R_k \alpha_k^{-1}}{\log_2 \varepsilon_k - \log_2 \varepsilon_{\min,k}} \leq \alpha_k S_{k,su}. \]  \hspace{1cm} (4.12)

Now (4.9) follows, because \( \alpha_k \) is a partition of unity.

Figure 4.1 illustrates the TDMA slope region, for the same multiple access channel from Example 3, which has \( K = 2 \) users. Here, TDMA is defined by a linear constraint (4.9), which connects vertices \( (S_{1,su}, 0) \) and \( (0, S_{2,su}) \). Visually, TDMA is contained within MUD(\( \mathcal{R} \)).

4.3 Single User Detector Receiver and Tightly Coordinated Users

With the assumption of superposition multiple access strategy, low complexity single-user detectors represent an attractive receiver design alternative, for a low power
multiple access system. This combination of multiple access and receiver technique is a natural competitor with TDMA, because both TDMA and single user receivers have relatively comparable complexity, which is very low, when compared to the complexity of full multiuser detectors.

In order to be able to compare the performance of single user detectors, against other multiple access strategies, here we extend the single-user formula (4.1), into multiaccess setting, assuming that single user detection receivers are employed at the access point.

**Theorem 4 (Slope Region of Single User Detection).** Let

\[ \Sigma_{\text{su}} = \text{diag} \left( S_{1,\text{su}} \ldots S_{k,\text{su}} \right) \]  

(4.13)

denote diagonal matrix of spectral efficiency slopes, in the single user communication mode. Notice that \( \Sigma_{\text{su}}^{-1} \) may be found on diagonal of \( \mathcal{C} \). Also, let \( \succeq \) denote the term by term inequality between vectors.

Set of spectral efficiency slopes, which are achievable by superposition multiple access strategy, in combination with simple single-user detector receivers, is given as

\[ \text{SUD}(\mathcal{R}) = \left\{ \Sigma : \Sigma^{-1} \mathcal{R} \succeq \left[ 2\mathcal{C} - \Sigma_{\text{su}}^{-1} \right] \mathcal{R} \right\}. \]  

(4.14)

**Proof.** Idea is that successive interference cancelation achieves vertices of the rate region, which correspond to vertices of MUD(\( \mathcal{R} \)). The user that is decoded first effectively experiences the performance of SUD(\( \mathcal{R} \)). The rest of the proof leverages of the results from Theorem 2. See Appendix for further details.

Clearly, SUD(\( \mathcal{R} \)) is relevant for tightly coordinated multiaccess system, because of its dependence on the direction vector \( \mathcal{R} \), which lies in the \( K \)-dimensional space of spectral efficiencies. As with MUD(\( \mathcal{R} \)) and TDMA, we conclude by giving a concrete example of (4.14).
Example 4. Consider single antenna channel model (2.9), which is found in both Example 1 and Example 3, where the appropriate Wideband Matrix $\mathbf{C}$ is also found in Example 3, equation (4.5).

Vector $\Sigma^{-1}\mathbf{R}$ from (4.14) has two rows, and the slope region SUD($\mathbf{R}$) is specified by two multiple access constraints. Utilising (4.5) we have

$$\text{SUD}(\mathbf{R}) = \left\{ (S_1, S_2) : \begin{array}{c}
\frac{R_1}{S_1} \geq \frac{R_1}{S_{1,\text{su}}} + R_2 \\
\frac{R_2}{S_2} \geq \frac{R_2}{S_{2,\text{su}}} + R_1
\end{array} \right\}.$$  \tag{4.15}

Region SUD($\mathbf{R}$), which is the set of achievable $(S_1, S_2)$ pairs from Example 4, is essentially a rectangle, that is portrayed in Figure 4.1. For this particular Example, TDMA achieves superior spectral efficiency slopes than single-user detectors, because SUD($\mathbf{R}$) is contained within TDMA, from Figure 4.1. However, this preliminary comparison heavily relies on the single-antenna channel model (2.9), which is relevant for the structure of $\mathbf{C}$ in (4.5). The full comparison of multiaccess strategies and different receiver structures, given various antenna array combinations, will be provided in Section 7.

4.4 Chapter Summary

The fundamental tradeoff between spectral efficiency and energy per information bit (3.9), in low power multiaccess systems, is fully described by evaluating the minimum energy per information bit $E_{\text{min},k}$ for each user $k$, and the set of achievable $S_k$, which is referred to as the slope region. With orthogonal multiple access strategies, the slope region is given as TDMA, whereas multiple access by superposition results in either MUD($\mathbf{R}$) or SUD($\mathbf{R}$), depending on whether the receiver employs multiuser detection, or alternatively, a bank of single user detectors.
The central object, which is used to describe all slope regions, is the Wideband Matrix $\mathcal{C}$, which was defined as (4.2). Next section focuses on evaluating $\mathcal{C}$ for multi-antenna systems.

4.5 Appendix: Chapter Proofs

*Proof of Theorem 2.* The basic approach is to generate multiaccess constraints on $S_k$, using the generic rate-region constraints from (2.6), and for that purpose we let $A \subseteq \{1, ..., K\}$. Multi-dimensional Taylor series of $C_A(P_A)$ at $P_A = 0$ gives

$$C_A(P_A) = \sum_{k \in A} \frac{\partial C_A(P_A)}{\partial P_k} \bigg|_0 P_k + \left(\frac{1}{2}\right) \sum_{k, i \in A} \frac{\partial^2 C_A(P_A)}{\partial P_k \partial P_i} \bigg|_0 P_k P_i + o(\|P_A\|^2). \quad (4.17)$$

We will utilize two substitutions to proceed from (4.17).

First, by definition of $C_A(P_A)$ in (2.2), if $k \in A$, then we have

$$\frac{\partial C_A(P_A)}{\partial P_k} \bigg|_0 = \frac{\dot{C}_{[k]}(0)}{\ln 2} = \frac{1}{\mathcal{E}_{\min, k}}, \quad (4.18)$$

which we utilize to substitute the first partials in (4.17). Second, we use (3.10) to eliminate powers $P_k$ from (4.17), for every user $k$. This is justified since (3.10) simply describes the relationship between power, rate and the resultant slope, which is irrespective of multiple access.

The generic multiaccess constraint on data rate is given as $\sum_{k \in A} R_k \leq C_A(P_A)$, and using (4.17), with described substitutions, we have that

$$\sum_{k \in A} R_k \leq \sum_{k \in A} \left(\frac{1}{\mathcal{E}_{\min, k}}\right) \left[ \mathcal{E}_{\min, k} R_k + \mathcal{E}_{\min, k} R_k^2 \left(\frac{\ln 2}{S_k}\right) + o(R_k^2) \right] + \left(\frac{1}{2}\right) \sum_{k, i \in A} \frac{\partial^2 C_A(P_A)}{\partial P_k \partial P_i} \bigg|_0 \left[ \mathcal{E}_{\min, k} R_k + \mathcal{E}_{\min, k} R_k^2 \left(\frac{\ln 2}{S_k}\right) + o(R_k^2) \right] \left[ \mathcal{E}_{\min, i} R_i + \mathcal{E}_{\min, i} R_i^2 \left(\frac{\ln 2}{S_i}\right) + o(R_i^2) \right] + o(\|P_A\|^2). \quad (4.19)$$
We begin by collecting first-order terms, which involve only $R_k$, that are found on the right-hand side of (4.19). Notice that the sum rate $\sum_{k \in A} R_k$, appears on the right in (4.19), which will cancel with the left hand side of (4.19).

Next, we collect the infinitesimal terms, where we use facts that $o(||P_A||^2) = o(||R_A||^2)$, also $R_k R_l^2 = o(||R_A||^2)$, and $o(R_k^2) = o(||R_A||^2)$. All of these infinitesimal relationships are straightforward. After canceling the sum rate on both sides, and collecting the infinitesimal we are left with

$$0 \leq \sum_{k \in A} R_k^2 \left( \frac{\ln 2}{S_k} \right) + \left( \frac{1}{2} \right) \sum_{k, l \in A} \frac{\partial^2 C_A(P_A)}{\partial P_k \partial P_l} \left|_{0}^{\mathcal{E}_{min,k} \mathcal{E}_{min,l}} R_k R_l + o(||R_A||^2) \right). \quad (4.22)$$

To proceed from (4.22), we focus on $\mathcal{E}_{min,k} \mathcal{E}_{min,l}$, which is grouped with the product $R_k R_l$. Here, we use (4.18) to re-write $\mathcal{E}_{min,k} = (\partial C_A / \partial P_k)^{-1}$, which we use along with $R_k R_l$ in (4.21).

These manipulations of (4.21) provide the following

$$0 \leq \ln 2 \sum_{k \in A} R_k^2 \left( \frac{\ln 2}{S_k} \right) + \left( \frac{\ln 2}{2} \right) \sum_{k, l \in A} \frac{\partial^2 C_A(P_A) \ln 2}{\partial P_k \partial P_l} \left|_{0}^{\frac{\partial C_A(P_A)}{\partial P_k} \ln 2} \frac{R_k R_l}{\frac{\partial C_A(P_A)}{\partial P_l} \ln 2} \right|_{0} + o(||R_A||^2), \quad (4.23)$$

after which we divide through by $(\ln 2)$, and a quadratic expression in $R_A$ remains. Notice that all partials in (4.23) are found after converting $C_A(P_A)$ to nats.

Finally, (4.23) is expressed using the definition of the Wideband Matrix $C$, where off-diagonal entries are negative second partials, which are divided by the product of the first partials (also multiplied by 2). This exactly appears with $R_k R_l$ in (4.23). Also, $R_k^2$ in the first summation in (4.23) is paired with inverse slopes $S_k$. In short (4.23) is simply expressed as

$$0 \leq R_A^T \left[ \Sigma_A^{-1} - C_A \right] R_A + o(||R_A||^2). \quad (4.24)$$

The result follows from (4.24), divided by $||R_A||^2$, and under limit as $||R_A|| \downarrow 0$. 

because the limiting process $\|R\| \downarrow 0$ takes place along the direction vector $\mathcal{R}$ in the $K$-dimensional space of data rates. This completes the proof of Theorem 2. \hfill \Box

**Proof of Example 3.** Here, we simply evaluate first and second partials of the aggregate capacity function $C(P)$. They are given as follows

$$C_{\{1,2\}}(P_1, P_2) = \text{E} \ln(1 + |h_1|^2P_1 + |h_2|^2P_2) \quad \text{[nats]}$$

$$\left. \frac{\partial C_{\{1,2\}}(P_1, P_2)}{\partial P_1} \right|_0 = \text{E}|h_1|^2$$

$$\left. \frac{\partial^2 C_{\{1,2\}}(P_1, P_2)}{\partial P_1 \partial P_2} \right|_0 = -\text{E}|h_1|^2\text{E}|h_2|^2 \quad \text{Independence of } h_1 \text{ and } h_2$$

$$\left. \frac{\partial^2 C_{\{1,2\}}(P_1, P_2)}{\partial P_1^2} \right|_0 = -\text{E}|h_1|^4. \quad (4.28)$$

From here, the Wideband matrix $\mathcal{C}$, which is in Example 3 follows. \hfill \Box

**Proof of Theorem 4.** To make the proof more readable, we will first provide the basic principle by considering only $K = 2$ users, after which we generalize to $K > 2$. The joint multiuser constraint, if the optimum multiuser detector (4.4) were used, is

$$\begin{bmatrix} \mathcal{R}_1 & \mathcal{R}_2 \end{bmatrix} \begin{bmatrix} \mathcal{S}_1^{-1} - \mathcal{S}_{1,\text{su}}^{-1} & -\mathcal{C}_{1,2} \\ -\mathcal{C}_{2,1} & \mathcal{S}_2^{-1} - \mathcal{S}_{2,\text{su}}^{-1} \end{bmatrix} \begin{bmatrix} \mathcal{R}_1 \\ \mathcal{R}_2 \end{bmatrix} = 0. \quad (4.29)$$

To solve for the performance that user 1 experiences with single user detectors, we make two substitutions inside (4.29). First, we note that if interference cancelation were used, in the order $1 \rightarrow 2$, then user 1 would experience the same performance as with single user detectors, and user 2 would have the interference clean channel with $\mathcal{S}_2 = \mathcal{S}_{2,\text{su}}$. The second substitution that we make inside (4.29) is simple symmetry of the Wideband Matrix $\mathcal{C}$, namely $\mathcal{C}_{1,2} = \mathcal{C}_{2,1}$.

With these two substitutions in (4.29) we have that

$$\mathcal{R}_1^2 \left[ \mathcal{S}_1^{-1} - \mathcal{S}_{1,\text{su}}^{-1} \right] - 2\mathcal{R}_1\mathcal{R}_2\mathcal{C}_{1,2} = 0 \quad (4.30)$$

$$\mathcal{R}_1\mathcal{S}_1^{-1} = \mathcal{R}_1\mathcal{S}_{1,\text{su}}^{-1} + 2\mathcal{R}_2\mathcal{C}_{1,2} \quad (4.31)$$
which solves for the performance, namely the spectral efficiency slope, of user 1.

Similarly, if user 2 was utilising single user detector as well, we simply reverse the indecent inside (4.31), and we have that

\[ R_2 S_2^{-1} = R_2 S_{2,\text{su}}^{-1} + 2R_1 C_{2,1}. \]  

Finally, (4.32) and (4.31) are combined to compactly write \( \Sigma^{-1} R = [2C - \Sigma^{-1}] R \), which completes the proof for \( K = 2 \) users. Next, we utilize the same principle to demonstrate the proof for \( K > 2 \) users.

Assuming that interference cancelation was performed in the order \( 1 \rightarrow K \), we let \( A = \{2, \ldots, K\} \). The joint multiuser constraint, if the optimum multiuser detector were utilized, is given as follows

\[
\begin{bmatrix}
R_1 & R_A^T \\
\end{bmatrix}
\begin{bmatrix}
S_1^{-1} - S_{1,\text{su}}^{-1} & -C_{1,A} \\
-C_{A,1} & \Sigma_A^{-1} - C_A \\
\end{bmatrix}
\begin{bmatrix}
R_1 \\
R_A \\
\end{bmatrix} = 0.
\]  

(4.33)

Again, we perform two substitutions inside (4.33). First, after user 1 has been cancelled, remaining users must be jointly optimal, which means that they must operate on the joint boundary of the slope region. Here \( R_A^T [\Sigma_A^{-1} - C_A] R_A = 0 \), which we utilize inside (4.33). Second, because \( C \) is symmetric we have \( C_{A,1} = C_{1,A}^T \).

With these substitutions, (4.33) becomes

\[
R_1^2 [S_1^{-1} - S_{1,\text{su}}^{-1}] = 2 \sum_{k=2}^{K} C_{1,k} R_1 R_k.
\]  

(4.34)

From here, we evaluate the spectral efficiency slope \( S_1^{-1} \), which is achieved by user 1, who was first in the decoding process. By symmetry from (4.34), we also find the spectral efficiency slopes, which are achieved assuming that all users employ
single-user detectors, namely

\[
\text{SUD}(\mathcal{R}) = \left\{ \Sigma : \forall k = \{1, \ldots, K\} \quad \mathcal{R}_k \Sigma_k^{-1} = \mathcal{R}_k \Sigma_{k,\text{su}}^{-1} + 2 \sum_{l \neq k} \mathcal{C}_{l,k} \mathcal{R}_l \right\}. \tag{4.35}
\]

\[
\text{SUD}(\mathcal{R}) = \left\{ \Sigma : \Sigma^{-1} \mathcal{R} = [2 \mathcal{C} - \Sigma_{\text{su}}^{-1}] \mathcal{R} \right\}. \tag{4.36}
\]

Formats (4.35) and (4.36) of SUD(\mathcal{R}) are equivalent, but (4.36) is simply more compact way of writing the set of equations. Also, (4.36) is the same as stated in Theorem (4), which naturally includes all slopes, which are smaller than (4.36), for the sake of completeness. \qed
Chapter 5

Antenna Arrays in Low Power Multiuser Systems

The only example of $\mathbf{C}$ matrix that we derived thus far is in (4.5), and it is computed under the assumption of single antenna transmission and reception. Because of practicality and tremendous gains in spectral efficiency, which are offered by antenna-array transmission and reception $[2, 4, 5, 32, 33]$, our sole focus in this section is finding the $\mathbf{C}$ matrix in multi-antenna systems. First, we derive $\mathbf{C}$ for the generic linear channel model (2.1), and second, we specialize the result for antenna array systems.

Both Wideband Matrix $\mathbf{C}$, and $\mathcal{E}_{\text{min},k}$, are summarized in the following Theorem.

Theorem 5 (Generic Structure of the Wideband Matrix). Let $\mathcal{K}_k$ represent the matrix channel of user $k$, in the multiple access channel model (2.1), and let the size of $\mathcal{K}_k$ be $M \times N_k$. Furthermore, let each user employ spatially uncorrelated i.i.d. complex Gaussian codebooks, which is described as $\mathbb{E} \left[ X_k X_k^\dagger \right] = (P_k / N_k) I_{N_k \times N_k}$, and where $I_{N_k \times N_k}$ represents the $N_k \times N_k$ identity matrix. We claim that

a) Minimum energy per information bit of user $k$ is given as

$$\mathcal{E}_{\text{min},k} = \frac{N_k \ln 2}{\text{tr} \mathbb{E} \left[ \mathcal{K}_k \mathcal{K}_k^\dagger \right]}.$$  \hspace{1cm} (5.1)

b) The $(k,l)$ entry of the Wideband Matrix $\mathbf{C}$ is

$$\mathbf{C}_{k,l} = \frac{\text{tr} \mathbb{E} \left[ \mathcal{K}_k \mathcal{K}_l^\dagger \mathcal{K}_l \mathcal{K}_k^\dagger \right]}{2 \text{tr} \mathbb{E} \left[ \mathcal{K}_k \mathcal{K}_k^\dagger \right] \text{tr} \mathbb{E} \left[ \mathcal{K}_l \mathcal{K}_l^\dagger \right]}.$$ \hspace{1cm} (5.2)
Proof. The aggregate capacity function $C(P)$ is given as the obvious extension of (2.3) to $K$ users. Given $C(P)$, we evaluate $\mathcal{C}$ as it is defined in (4.2). See Appendix for further details.

Formula (5.1), for minimum energy per information bit, may also be found in the single-user discussion in [2]. Expression for $\mathcal{C}$ in (5.2) encompasses all linear channel models, whether $\mathcal{H}_k$ is stochastic - as in the antenna array communications, or deterministic - as in the case of short spreading codes. Here, we focus on antenna array transmission and reception, where $\mathcal{H}_k$ represents the stochastic matrix of fading coefficients, and we evaluate both (5.1) and (5.2) in a closed form, and in the presence of transmit/receive antenna correlations.

Antenna correlation is commonly modeled as

$$\mathcal{H}_k = \Phi_k^{1/2} \hat{\mathcal{H}}_k \Theta_k^{1/2}, \tag{5.3}$$

where $\Phi_k$ is named the receive antenna correlation, and $\Theta_k$ is the transmit antenna correlation, specific for user $k$. Both matrixes $\Phi_k$ and $\Theta_k$ are deterministic and non-negative definite, whereas the matrix root in (5.3) may be defined based on the eigenvalue decomposition* [22]. Because both $\Theta_k$ and $\Phi_k$ are deterministic, channel randomness is captured by entries of $\hat{\mathcal{H}}_k$, which are assumed to be zero-mean, independent, and identically distributed.

Besides these properties of $\mathcal{H}_k$, we also assume that the probability distribution on $\mathcal{H}_k$ has additional characteristic of left and right rotational invariance, which is defined by the following two conditions. First, given any unitary matrix $Q$, namely

*All final results are direct functions of $\Theta \geq 0$ and $\Phi \geq 0$, and not their roots. For this reason, $\Theta^{1/2}$ may be any old matrix which satisfies $\Theta^{1/2} (\Theta^{1/2})^\dagger = \Theta$. For example, given eigenvalue decomposition $QAQ^\dagger = \Theta$, then $\Theta^{1/2}$ may be defined as either $\Theta^{1/2} = QA^{1/2}Q^\dagger$, which is hermitian and therefore $\Theta^{1/2} \Theta^{1/2} = \Theta$, or alternatively $\Theta^{1/2} = QA^{1/2}$, which is not necessarily hermitian.
\( QQ^\dagger = Q^\dagger Q = I \), probability distributions of \( \hat{\mathcal{H}}_k \) and \( \hat{\mathcal{H}}_k Q \) are identical (right rotational invariance). Second, given any unitary matrix \( U \), probability distributions on \( \hat{\mathcal{H}}_k \) and \( U \hat{\mathcal{H}}_k \) are identical (left rotational invariance). One example, of both left and right rotationally invariant probability distribution, is given by the standard complex Gaussian, which is commonly used to model the antenna correlation in Rayleigh fading channels.

The following corollary to Theorem 5 specializes formulas (5.1) and (5.2) from above, in order to address the case (5.3) of correlated antenna arrays.

**Corollary 1 (Wideband Matrix For Correlated Antenna Arrays).** Let \( \mathcal{H}_k \) represent the random matrix of antenna coefficients, for the \( K \)-user multiaccess channel model (2.1), and let antenna correlation be modeled by (5.3). Furthermore, let users employ codebooks, which are spatially uncorrelated i.i.d. complex Gaussian, namely \( \mathbb{E} \left[ X_k X_k^\dagger \right] = (P_k/N_k) I_{N_k \times N_k} \). Let \( \kappa_k \) be the kurtosis of a generic entry of \( \hat{\mathcal{H}}_k \).

First, we claim that the minimum energy per information bit, for user \( k \), satisfies

\[
\mathcal{E}_{\text{min},k} = \frac{N_k \ln 2}{\text{tr} (\Theta_k) \text{tr} (\Phi_k) \mathbb{E} \left[ |\hat{\mathcal{H}}_k|^2 \right]}.
\]

(5.4)

Second, to specify the Wideband Matrix \( \mathcal{C} \), a distinction is made between diagonal elements of \( \mathcal{C} \), where \( k = l \), and off-diagonal elements where \( k \neq l \), because channels \( \mathcal{H}_k \) and \( \mathcal{H}_l \) are assumed independent for different users. The Wideband Matrix \( \mathcal{C} \) is given by the following expressions

\[
\mathcal{C}_{k,l} = \frac{\text{tr} (\Phi_k \Phi_l)}{2 \text{tr} (\Phi_k) \text{tr} (\Phi_l)}, \quad \text{for} \quad k \neq l.
\]

(5.5)

\[
\mathcal{C}_{k,k} = \frac{\text{tr} (\Phi_k^2) \text{tr} (\Theta_k^2)}{2 (\text{tr} \Phi_k)^2 (\text{tr} \Theta_k)^2} (\kappa_k - 2) + \frac{\text{tr} (\Phi_k^2)}{2 (\text{tr} \Phi_k)^2} + \frac{\text{tr} (\Theta_k^2)}{2 (\text{tr} \Theta_k)^2}.
\]

(5.6)

**Proof.** Meticulous evaluation of averages in (5.1) and (5.2). See Appendix. \( \square \)
Note that $\mathcal{C}_{k,k}$, which is given in (5.6), can be also represented as $\mathcal{S}_{k,su}^{-1}$, which is straightforward from the definition of $\mathcal{C}$. For single-user uncorrelated multi-antenna systems, $\mathcal{S}_{k,su}$ is found in [2], which was further extended in [10], for the case study of correlated Rayleigh faded antennas, which have $\kappa_k = 2$. The Wideband matrix $\mathcal{C}$, which we provided above, is a natural extension of these results for multiuser systems.

We proceed by providing a special case of Corollary 1, in order to explicitly address the benchmark channel model of multi-antenna systems, where both transmit and receive antenna arrays are uncorrelated, and Rayleigh faded.

**Corollary 2 (Wideband Matrix For Uncorrelated Antenna Arrays).** Suppose all propagation coefficients in the channel model (2.1) are uncorrelated and Rayleigh faded, with unit variance.

The minimum energy per information bit of user $k$ is given as

$$
\mathcal{E}_{\min, k} = \frac{\ln 2}{M},
$$

(5.7)

where we used (5.4), $\Theta_k = I_{N_k \times N_k}$, along with $\Phi_k = I_{M \times M}$, and $\mathbb{E}[|\hat{h}|^2] = 1$.

To specify the wideband matrix, we use (5.5), and (5.6), together with $\kappa_k = 2$. It follows that

$$
\mathcal{C}_{k,l} = \frac{1}{2M}, \quad \text{for} \quad k \neq l.
$$

(5.8)

$$
\mathcal{C}_{k,k} = \frac{1}{\mathcal{S}_{k,su}} = \frac{1}{2M} + \frac{1}{2N_k}.
$$

(5.9)

Corollary 2, as well as Corollary 1 and Theorem 5 are all derived for the case study of spatially independent codebooks, which was described by the condition $\mathbb{E}[X_kX_k^\dagger] = (P_k/N_k)I_{N_k \times N_k}$. However, it is fairly straightforward to remove this assumption, which is what we do next. The motivation for this step comes from
the fact that for correlated antenna arrays, spatially independent codebooks are no longer optimal, and users may now opt to exploit the antenna correlation, by resorting to spatially correlated codebooks.

5.1 Spatially Correlated Codebooks

To evaluate the $\mathcal{C}$ matrix, for the case of spatially correlated codebooks, we represent the spatial correlation of the signal as $E \left[ X_k X_k^\dagger \right] = (P_k/N_k) F_k$, where $F_k \geq 0$ is a positive semidefinite matrix. Here, $F_k$ is implicitly normalized to satisfy $\text{tr}(F_k) = \text{tr}(I_{N_k \times N_k}) = N_k$, so that the average transmitted power of user $k$ remains $\text{tr} E[X_k X_k^\dagger] = P_k$. Spatial signal correlation is incorporated inside the generic linear channel model (2.1), by introducing a simple substitution $X_k = F_k^{\frac{1}{2}} \hat{X}_k$, where $\hat{X}_k$ is i.i.d complex Gaussian.

The following Corollary to Theorem 5 is a more general version of Corollary 1, in order to account for spatially correlated codebooks.

Corollary 3 (Wideband Matrix for Spatially Correlated Codebooks). Let $\mathbf{H}_k$ represent the random matrix of antenna coefficients, for the $K$-user multiaccess channel model (2.1), and let the antenna correlation be modeled by (5.3). In contrast to Corollary 1, we now assume that users transmit with codebooks which satisfy $E \left[ X_k X_k^\dagger \right] = (P_k/N_k) F_k$, and $\text{tr}(F_k) = N_k$.

We claim that

a) Minimum energy per information bit, for user $k$, is given as

$$E_{\text{min},k} = \frac{N_k \ln 2}{\text{tr}(F_k \Theta_k) \text{tr}(\Phi_k) E \left[ |\hat{h}|^2 \right]}.$$  \hspace{1cm} (5.10)

b) Wideband Matrix $\mathcal{C}$ is again specified by making a distinction, between off-diagonal entries $k \neq l$, and diagonal entries $k = l$. For the case of spatially correlated
codebooks, we have that

\[ C_{k,l} = \frac{\text{tr} \left( \Phi_k \Phi_l \right)}{2 \text{tr} \left( \Phi_k \right) \text{tr} \left( \Phi_l \right)}, \quad \text{for} \quad k \neq l. \quad (5.11) \]

\[ C_{k,k} = \frac{\text{tr} \left( \Phi_k^2 \right) \text{tr} \left[ \left( F_k \Theta_k \right)^2 \right]}{2 \left( \text{tr} \left( \Phi_k \right) \right)^2 \left[ \text{tr} \left( F_k \Theta_k \right) \right]^2} + \frac{\text{tr} \left( \Phi_k^2 \right)}{2 \left( \text{tr} \left( \Phi_k \right) \right)^2} + \frac{\text{tr} \left[ \left( F_k \Theta_k \right)^2 \right]}{2 \left[ \text{tr} \left( F_k \Theta_k \right) \right]^2}. \quad (5.12) \]

**Proof.** Spatially correlated codebooks are described by introducing the substitution \( X_k = F_k^{\frac{1}{2}} \hat{X}_k \), where \( \hat{X}_k \) is i.i.d. complex Gaussian, and \( E \left[ \hat{X}_k \hat{X}_k^\dagger \right] = \left( P_k/N_k \right) I_{N_k \times N_k} \). Notice that this substitution results in spatial covariance, which is given as \( E \left[ X_k X_k^\dagger \right] = \left( P_k/N_k \right) F_k \), and it follows that

\[ \mathcal{H}_k X_k = \left[ \Phi_k^{\frac{1}{2}} \hat{\mathcal{H}}_k \Theta_k^{\frac{1}{2}} F_k^{\frac{1}{2}} \hat{X}_k \right] \quad (5.13) \]

\[ = \left[ \Phi_k^{\frac{1}{2}} \hat{\mathcal{H}}_k \left( \Theta_k^{\frac{1}{2}} F_k^{\frac{1}{2}} \right) \right] \hat{X}_k. \quad (5.14) \]

We now compare the structure of (5.14) and (5.3). The rest of the proof is the same as the proof of Corollary 1.

In effect, matrix \( F_k^{\frac{1}{2}} \neq I \) is a pre-coding filter, which may be designed to exploit the transmit antenna correlation \( \Theta_k \). Namely, in low power systems, the primary reason for introducing \( F_k \) is to minimize the \( \mathcal{E}_{\min,k} \), from (5.10), which is accomplished by maximizing \( \text{tr} \left( F_k \Theta_k \right) \), under constraints \( F_k \geq 0 \), and \( \text{tr}(F_k) = N_k \).

This optimization problem is addressed in the following Proposition.

**Proposition 2.** Suppose the maximum eigenvalue of \( \Theta_k \) has multiplicity \( n_k \), and let columns of the matrix \( U \), which is of size \( N_k \times n_k \), represent corresponding principal \( n_k \) eigenvectors. To maximize \( \text{tr} \left( F_k \Theta_k \right) \), under constraints \( F_k \geq 0 \), and \( \text{tr}(F_k) = N_k \), we set \( F_k = (N_k/n_k) U U^\dagger \), and the maximum becomes \( \text{tr} \left( F_k \Theta_k \right) = N_k \lambda_{\max}(\Theta_k) \).

**Proof.** See Appendix.

Proposition 2 may be interpreted as statistical form of beamforming, because it suggests that user \( k \) should concentrate all transmit energy in the subspace which
is spanned by principal eigenvectors of $\Theta_k$, in order to exploit the maximum gain from the communication channel. This energy is concentrated using $F_k^\frac{3}{2}$, and because $F_k = (N_k/n_k)UU^\dagger$, then we may simply set $F_k^\frac{3}{2} = \sqrt{N_k/n_k} U$.

For this particular choice of $F_k$, Corollary 3 is specialized as follows.

**Corollary 4 (Wideband Matrix for Codebooks with Optimized Spatial Correlation).** Let $H_k$ represent the random matrix of antenna coefficients, for the $K$-user multiaccess channel model (2.1), and let the antenna correlation be modeled by (5.3). Let users transmit with codebooks which satisfy $E[X_kX_k^\dagger] = (P_k/N_k)F_k$, where $F_k$ is provided in Proposition 2.

We have that

$$E_{\min,k} = \frac{\ln 2}{\lambda_{\max}(\Theta_k) \text{tr}(\Phi_k) E[|\hat{h}|^2]}.$$  \hspace{1cm} (5.15)

Notice from (5.11) that off-diagonal entries of $\mathcal{C}$ are unaffected by the choice of $F_k$. To specify diagonal entries of $\mathcal{C}$, we use (5.12) and $F_k$ from Proposition 2, for which we compute that $\text{tr}[(F_k\Theta_k)^2] = N_k^2\lambda_{\max}(\Theta_k)/n_k$. It now follows

$$\mathcal{C}_{k,k} = \frac{\text{tr}(\Phi_k^2)}{2(\text{tr} \Phi_k)^2} \left( \frac{\kappa_k - 2}{n_k} + 1 \right) + \frac{1}{n_k}.$$ \hspace{1cm} (5.16)

To conclude, the optimum $E_{\min,k}$ is given in (5.15), which is achieved when users are aware of transmit antenna correlations $\Theta_k$, and when they concentrate transmit energy inside the subspace which is spanned by principal eigenvectors of $\Theta_k$. On the other hand, knowing receive antenna correlations $\Phi_k$ does not help to optimize $F_k$. Furthermore, (5.16) along with (5.11) is the format of the Wideband matrix $\mathcal{C}$, provided that each user utilizes the optimum $F_k$, which minimizes $E_{\min,k}$.
5.2 Chapter Summary

The most generic form of the Wideband Matrix $\mathcal{C}$, for linear channel models, is given in Theorem 5. For correlated multi-antenna channels, Theorem 5 is specialized into Corollary 1, provided that users transmit with spatially white codebooks. Even simpler version of Corollary 1 is given in Corollary 2, which provides $\mathcal{C}$, for the benchmark case study of uncorrelated Rayleigh faded antenna arrays.

On the other hand, when users are allowed to transmit with spatially correlated codebooks, Theorem 5 is used to provide Corollary 3, which incorporates the precoding filter. If this filter is given as in Proposition 2, then the $\mathcal{C}$ matrix is provided in Corollary 4, which is a special case of Corollary 3.

5.3 Appendix: Chapter Proofs

*Proof of Theorem 5.* Because of equal power allocation across all transmit antennas, vector $X_k$ of users $k$ transmit symbols, becomes independent and identically distributed complex Gaussian, which is described as

$$
E \left[ X_k X_k^\dagger \right] = \left( \frac{P_k}{N_k} \right) I_{N_k \times N_k}.
$$

(5.17)

Here, $I_{N_k \times N_k}$ is the identity matrix, and normalization by $N_k$ is there in (5.17), because the aggregate transmit power of user $k$ must satisfy $\text{tr} E \left[ X_k X_k^\dagger \right] = P_k$.

Given a particular realization of the random channel $\mathcal{H}_k$, the covariance of the signal part of user $k$ is given as follows

$$
E \left[ \mathcal{H}_k X_k X_k^\dagger \mathcal{H}_k^\dagger \left| \mathcal{H}_k \right. \right] = \mathcal{H}_k \mathcal{H}_k^\dagger \left( \frac{P_k}{N_k} \right).
$$

(5.18)

On the other hand, covariance of the additive white Gaussian noise is identity $E \left[ \eta \eta^\dagger \right] = I_{M \times M}$. From (5.18), we will directly express the mutual information $C(P)$
in nats, namely \( C(P) \) is given by

\[
C(P) \ln 2 = \mathbb{E} \ln \det \left[ I + \sum_{k=1}^{K} \mathcal{H}_k \mathcal{H}_k^\dagger \left( \frac{P_k}{N_k} \right) \right].
\] (5.19)

In order to evaluate the entries of the Wideband Matrix \( \mathcal{C} \) for (5.19), we will require the following lemma, to deal with the partials of \( C(P) \).

**Lemma 1.** Let \( B \) be a generic matrix, and let \( P \) be a scalar.

The \( \ln \det \) function satisfies following identities

\[
\frac{d}{dP} \ln \det [I + BP] \bigg|_{P=0} = \text{tr}(B) 
\] (5.20)

\[
\frac{d^2}{dP^2} \ln \det [I + BP] \bigg|_{P=0} = -\text{tr}(B^2) 
\] (5.21)

\[
\frac{\partial^2}{\partial P_1 \partial P_2} \ln \det [I + B_1 P_1 + B_2 P_2] \bigg|_{P_1=0, P_2=0} = -\text{tr}(B_1 B_2).
\] (5.22)

**Proof.** To see (5.20), let \( B = QAQ^\dagger \) represent the eigenvalue decomposition of \( B \), where \( \Lambda \) is diagonal matrix, which is built from eigenvalues of \( B \). Since eigenvalues of \( BP \) are \( \Lambda(m, m)P \), it follows that eigenvalues of \( (I + BP) \) are \( 1 + \Lambda(m, m)P \).

Determinant in (5.20) is expressed as product of eigenvalues of \( (I + BP) \), and we have that

\[
\frac{d}{dP} \ln \det [I + BP] \bigg|_{P=0} = \frac{d}{dP} \sum_m \ln [1 + \Lambda(m, m)P] \bigg|_{P=0} 
\] (5.23)

\[
= \sum_m \frac{\Lambda(m, m)}{1 + \Lambda(m, m)P} \Bigg|_{P=0} 
\] (5.24)

\[
= \text{tr}(\Lambda) = \text{tr}(B).
\] (5.25)

This completes the proof of identity (5.20).

To proceed, and to prove (5.21), we use the similar approach as (5.20), and take another derivative of (5.24), namely

\[
\frac{d^2}{dP^2} \ln [1 + \Lambda(m, m)P] = \frac{-\Lambda^2(m, m)}{[1 + \Lambda(m, m)P]^2}.
\] (5.26)
From here it follows that
\[
\frac{d^2}{dP^2} \ln \det [I + BP] \big|_{P=0} = -\frac{d^2}{dP^2} \sum_m \ln [I + \Lambda(m, m)P] \big|_{P=0} \quad (5.27)
\]
\[
= -\sum_m \Lambda^2(m, m) = -\text{tr}(\Lambda^2). \quad (5.28)
\]
Finally, we note that \( \text{tr}(B^2) = \text{tr}(Q\Lambda Q^\dagger Q\Lambda Q^\dagger) = \text{tr}(Q\Lambda^2 Q^\dagger) = \text{tr}(\Lambda^2 Q^\dagger Q) = \text{tr}(\Lambda^2), \)
where we used the standard fact about the trace operator \( \text{tr}(B_1B_2) = \text{tr}(B_2B_1) \).
Hence, \( \text{tr}(\Lambda^2) = \text{tr}(B^2), \) which completes the proof of (5.21).

The proof of (5.22) is more involved, and we begin by the following expansion
\[
I + B_1P_1 + B_2P_2 = (I + B_1P_1) \left[ I + (I + B_1P_1)^{-1} B_2P_2 \right] \quad (5.29)
\]
\[
\ln \det [I + B_1P_1 + B_2P_2] = \ln \det[I + B_1P_1] + \ln \det \left[ I + (I + B_1P_1)^{-1} B_2P_2 \right] \quad (5.30)
\]
\[
= \text{tr}(B_1) P_1 + o(P_1) + \text{tr} \left[ (I + B_1P_1)^{-1} B_2 \right] P_2 + o(P_2). \quad (5.31)
\]
In transition from (5.30) to (5.31), we used the expansion \( \ln \det[I + B_1P_1] = \text{tr}(B_1) P_1 + o(P_1), \) that holds true by (5.20), because the first derivative of \( \ln \det[I + B_1P_1] \) at \( P_1 = 0 \) is \( \text{tr}(B_1) \). The same expansion is used for the other term in (5.30), in order to arrive at (5.31).

Next, to deal with nuisance term \( (I + B_1P_1)^{-1} \) in (5.31), we observe the following
\[
(I + B_1P_1) (I - B_1P_1) = I - B_1^2 P_1^2 = I + o(P_1) \quad (5.32)
\]
\[
(I - B_1P_1) = (I + B_1P_1)^{-1} [I + o(P_1)]. \quad (5.33)
\]
Now we use the fact that \( (I + B_1P_1)^{-1} o(P_1) = o(P_1), \) which is true because \( (I + B_1P_1) \) simply converges to a constant as \( P_1 \to 0 \). Using this with (5.33) we have
\[
(I - B_1P_1) = (I + B_1P_1)^{-1} + o(P_1) \quad (5.34)
\]
\[
(I + B_1P_1)^{-1} = (I - B_1P_1) + o(P_1). \quad (5.35)
\]
With (5.35), and (5.31) we have that
\[
\ln \det[I + B_1 P_1 + B_2 P_2] = \text{tr}(B_1) P_1 + o(P_1) + \text{tr}[(I - B_1 P_1) B_2] P_2 + o(P_2)
\]
\[
= -\text{tr}(B_1 B_2) P_1 P_2 + \text{tr}(B_1) P_1 + \text{tr}(B_2) P_2 + o(P_1) + o(P_2).
\]
(5.37)

Here, (5.37) simply re-groups (5.36).

To complete the proof of (5.22), we evaluate the cross-partial of (5.37), with respect to \( P_1 \) then with respect to \( P_2 \). First, fix \( P_2 \). After evaluating partial with respect \( P_1 \), at \( P_1 = 0 \), (5.37) becomes \(-\text{tr}(B_1 B_2) P_2 + \text{tr}(B_1)\). After evaluating the partial of this with respect to \( P_2 \), only \(-\text{tr}(B_1 B_2)\) is left standing tall, and completing the proof of (5.22), which completes the proof of Lemma 1.

To finalize the proof of Theorem 5, we utilize the Lemma 1, with the labeling \( B_1 = (\mathcal{H}_k \mathcal{H}_k^\dagger)/N_k \) as well as \( B_2 = (\mathcal{H}_l \mathcal{H}_l^\dagger)/N_k \) in the expression for mutual information (5.19). Partial of (5.19) are therefore given as follows
\[
\left. \frac{\partial C(P) \ln 2}{\partial P_k} \right|_{P=0} = \frac{\text{E tr} \left[ \mathcal{H}_k \mathcal{H}_k^\dagger \right]}{N_k}
\]
(5.38)
\[
\left. \frac{\partial^2 C(P) \ln 2}{\partial P_k^2} \right|_{P=0} = \frac{\text{E tr} \left[ \mathcal{H}_k \mathcal{H}_k^\dagger \mathcal{H}_k \mathcal{H}_k^\dagger \right]}{N_k}
\]
(5.39)
\[
\left. \frac{\partial^2 C(P) \ln 2}{\partial P_k \partial P_l} \right|_{P=0} = \frac{\text{E tr} \left[ \mathcal{H}_k \mathcal{H}_k^\dagger \mathcal{H}_l \mathcal{H}_l^\dagger \right]}{N_k}
\]
(5.40)

Here, (5.38), (5.39) and (5.40) follow from the Lemma 1, and \( C(P) \) in (5.19).

Generic entry of the Wideband Matrix \( \mathcal{C}_{k,l} \) is (5.40), divided by twice the product of (5.38) for \( k \) and \( l \). Notice that we required (5.39) to cover for the case \( k = l \). This completes the proof of Theorem 5.
Proof of Corollary 1. Here, we are set to evaluate the averages (5.38), (5.39) and (5.40) in the presence of antenna correlation which is described in (5.2). In order to do so, we break the proof into two parts.

The first part consists of two Lemmas, which deal with averages of $\mathcal{H}$.

**Lemma 2.** Let $\mathcal{H} = \Phi^{1/2} \tilde{\mathcal{H}} \Theta^{1/2}$, where $\tilde{\mathcal{H}}$ is zero-mean $M \times N$ matrix, with independent, identically distributed, and rotationally invariant entries, both from left and from the right. Generic entry of $\tilde{\mathcal{H}}$ is denoted as $\tilde{h}$, and its second moment $E[|\tilde{h}|^2]$.

We claim that $\mathcal{H}$ satisfies

$$E[\mathcal{H} \mathcal{H}^\dagger] = \Phi \text{tr}(\Theta) E[|\tilde{h}|^2].$$

(5.41)

**Proof.** We begin by direct computation

$$E[\mathcal{H} \mathcal{H}^\dagger] = E \left[ (\Phi^{1/2} \tilde{\mathcal{H}} \Theta^{1/2}) \left( \Theta^{1/2} \tilde{\mathcal{H}}^\dagger \Phi^{1/2} \right) \right]$$

(5.42)

$$= \Phi^{1/2} E \left[ \tilde{\mathcal{H}} \Theta \tilde{\mathcal{H}}^\dagger \right] \Phi^{1/2}.$$  

(5.43)

To proceed from (5.43), we will use the eigenvalue decomposition of $\Theta = Q \Lambda Q^\dagger$, where $\Lambda$ is diagonal matrix, which contains all eigenvalues of $\Theta$.

(5.43) now becomes

$$E[\mathcal{H} \mathcal{H}^\dagger] = \Phi^{1/2} E \left[ \tilde{\mathcal{H}} (Q \Lambda Q^\dagger) \tilde{\mathcal{H}}^\dagger \right] \Phi^{1/2}$$

(5.44)

$$= \Phi^{1/2} E \left[ \tilde{\mathcal{H}} \Lambda \tilde{\mathcal{H}}^\dagger \right] \Phi^{1/2}.$$  

(5.45)

Here, (5.45) holds because distribution of $\tilde{\mathcal{H}}$ is assumed to be rotationally invariant, and hence probability distribution of $\tilde{\mathcal{H}}$ is the same as probability distribution of $\tilde{\mathcal{H}} Q$, which allows us to effectively eliminate $Q$.

From (5.45) we proceed on entry-by-entry basis. Because $\Lambda$ is diagonal, it simply scales $n$-th column of $\tilde{\mathcal{H}}$ by $\Lambda(n, n)$. Because of this, the entry $(m, n)$ of $\tilde{\mathcal{H}} \Lambda$ therefore
becomes \( \hat{\mathcal{H}}(m, n)\Lambda(n, n) \). Using the standard formula for \((m, m')\) entry of matrix product \( (\hat{\mathcal{H}}\Lambda)\hat{\mathcal{H}}^\dagger \), where brackets denote association, we note that \((m, m')\) entry of \( \hat{\mathcal{H}}\Lambda\hat{\mathcal{H}}^\dagger \) equals the following
\[
\sum_{n=1}^{N} \hat{\mathcal{H}}(m, n)\Lambda(n, n) \left[ \hat{\mathcal{H}}(m', n) \right]^*,
\]
where * denotes conjugate.

The average of (5.46) equals 0 if \( m \neq m' \), because \( \hat{\mathcal{H}}(n, m) \) is independent from \( \hat{\mathcal{H}}(m', n) \) if and only if \( m \neq m' \). On the other hand, when \( m = m' \), then the average of (5.46) equals \( \mathbb{E} \left[ |\hat{h}|^2 \right] \sum_n \Lambda(n, n) \). It follows that the average of the entire matrix \( \hat{\mathcal{H}}\Lambda\hat{\mathcal{H}}^\dagger \) is in fact \( \mathbb{E} \left[ |\hat{h}|^2 \right] \text{tr}(\Lambda) \) multiple of the identity matrix.

From here (5.45) becomes
\[
\mathbb{E} \left[ \mathcal{H}\mathcal{H}^\dagger \right] = \Phi \text{tr}(\Lambda) \mathbb{E} \left[ |\hat{h}|^2 \right] = \Phi \text{tr}(\Theta) \mathbb{E} \left[ |\hat{h}|^2 \right],
\]
where we used standard eigenvalue decomposition result \( \text{tr}(\Theta) = \text{tr}(\Lambda) \). The last expression completes the proof of Lemma 2.

In order to find the fourth order averages, we require the following Lemma

**Lemma 3.** Under the same conditions of Lemma 2, we denote the fourth moment of \( \hat{h} \) to be \( \mathbb{E} \left[ |\hat{h}|^4 \right] \). The following holds
\[
\mathbb{E} \text{tr} \left[ \mathcal{H}\mathcal{H}^\dagger \mathcal{H}\mathcal{H}^\dagger \right] = \text{tr} (\Phi^2) \text{tr} (\Theta^2) \left( \mathbb{E} \left[ |\hat{h}|^4 \right] - 2 \mathbb{E} \left[ |\hat{h}|^2 \right]^2 \right)
\]
\[
+ \left[ \text{tr} (\Theta^2) (\text{tr} \Phi)^2 + \text{tr} (\Phi^2) (\text{tr} \Theta)^2 \right] \mathbb{E} \left[ |\hat{h}|^4 \right].
\]

**Proof.** We will frequently below use standard trace identity \( \text{tr}(B_1B_2) = \text{tr}(B_2B_1) \).

We begin by direct computation
\[
\mathbb{E} \text{tr} \left[ \mathcal{H}\mathcal{H}^\dagger \mathcal{H}\mathcal{H}^\dagger \right] = \mathbb{E} \text{tr} \left[ \left( \Phi^{1/2}\hat{\mathcal{H}}\Theta^{1/2} \right) \left( \Theta^{1/2}\hat{\mathcal{H}}^\dagger\Phi^{1/2} \right) \left( \Phi^{1/2}\hat{\mathcal{H}}\Theta^{1/2} \right) \left( \Theta^{1/2}\hat{\mathcal{H}}^\dagger\Phi^{1/2} \right) \right]
\]
\[
= \mathbb{E} \text{tr} \left[ \hat{\mathcal{H}}\Theta\hat{\mathcal{H}}^\dagger\Phi \hat{\mathcal{H}}\Theta\hat{\mathcal{H}}^\dagger\Phi \right].
\]
To clarify (5.50), we note that it abbreviates three steps. First, it collects roots like \( \Theta^{1/2} \Theta^{1/2} = \Theta \). Second, it moves \( \Phi^{1/2} \) from left to right, using the fact that 
\[
\text{tr}(B_1 B_2) = \text{tr}(B_2 B_1).
\]
Third, it collects roots \( \Phi^{1/2} \Phi^{1/2} = \Phi \).

To proceed from (5.50) to (5.51), we will use eigenvalue decompositions of matrixes \( \Theta = Q \Lambda Q^\dagger \) and of \( \Phi = U D U^\dagger \). We have that
\[
E \text{tr} \left[ \mathcal{K} \mathcal{H}^\dagger \mathcal{K} \mathcal{H}^\dagger \right] = E \text{tr} \left[ \hat{\mathcal{K}} (Q \Lambda Q^\dagger) \hat{\mathcal{H}}^\dagger (U D U^\dagger) \hat{\mathcal{K}} (Q \Lambda Q^\dagger) \hat{\mathcal{H}}^\dagger (U D U^\dagger) \right] \tag{5.51}
\]
\[
= E \text{tr} \left[ (U^\dagger \hat{\mathcal{K}} Q) \Lambda (Q^\dagger \hat{\mathcal{H}}^\dagger U) D (U^\dagger \hat{\mathcal{K}} Q) \Lambda (Q^\dagger \hat{\mathcal{H}}^\dagger U) D \right] \tag{5.52}
\]

Notice that (5.52) is simply (5.51), where we moved \( U^\dagger \) from right to left, again using the trace identity \( \text{tr}(B_1 B_2) = \text{tr}(B_2 B_1) \). Terms are re-grouped in brackets in (5.52) because this grouping will make use of rotational invariance of probability distribution on \( \hat{\mathcal{K}} \), which is the next step.

Namely, because probability distribution of \( \hat{\mathcal{K}} \) is the same as the probability distribution of \( U^\dagger \hat{\mathcal{K}} Q \), we have from (5.52) that
\[
E \text{tr} \left[ \mathcal{K} \mathcal{H}^\dagger \mathcal{K} \mathcal{H}^\dagger \right] = E \text{tr} \left[ \hat{\mathcal{K}} \Lambda \hat{\mathcal{H}}^\dagger D \hat{\mathcal{K}} \Lambda \hat{\mathcal{H}}^\dagger D \right] \tag{5.53}
\]
\[
= E \text{tr} \left[ (\hat{\mathcal{K}} \Lambda \hat{\mathcal{H}}^\dagger D)^2 \right], \tag{5.54}
\]

where (5.54) is nothing but an algebraic observation.

From (5.54) we must continue on a term-by-term basis. We already have expression (5.46) for \( (m, m') \) entry of \( \hat{\mathcal{K}} \Lambda \hat{\mathcal{H}}^\dagger \). Because \( D \) is diagonal we have that \( (m, m') \) entry of \( \hat{\mathcal{K}} \Lambda \hat{\mathcal{H}}^\dagger D \) equals
\[
\sum_{n=1}^{N} \hat{\mathcal{K}}(m, n) \Lambda(n, n) \left[ \hat{\mathcal{K}}(m', n) \right]^* D(m', m'). \tag{5.55}
\]

To address the square in (5.54) we note the following: if \( B \) is a generic \( M \times M \) matrix, then each \( (m, m) \) entry of \( B^2 \) equals \( \sum_{m'} B(m, m') B(m', m) \). From here
we evaluate the trace of $B^2$ as $\text{tr}(B^2) = \sum_{m,m'} B(m,m')B(m',m)$. With this and (5.55) we have that (5.54), without the averaging operator $E$, admits a compact representation as follows

$$\text{tr} \left[ \left( \hat{H} \Lambda \hat{H}^\dagger D \right)^2 \right] = \sum_{m,m'} \sum_n \hat{H}(m,n) \Lambda(n,n) \left[ \hat{H}(m',n) \right]^* D(m',m') \times \sum_{n'} \hat{H}(m',n') \Lambda(n',n') \left[ \hat{H}(m,n') \right]^* D(m,m) \quad (5.56)$$

To evaluate the average of (5.56) we will distinguish between four possible cases.

First, we eliminate the easy case when $m \neq m'$ and $n \neq n'$, and the average of (5.56) equals 0. This holds because of both zero-mean, and the independence assumption, on $\hat{H}(m,n)$.

Second case is given when $m = m'$ and $n = n'$, and average of (5.56) equals

$$E \sum_m \sum_n \left| \hat{H}(m,n) \right|^4 |\Lambda(n,n)|^2 |D(m,m)|^2 = E \left[ |\hat{h}|^4 \right] \text{tr} \left( \Lambda^2 \right) \text{tr} \left( D^2 \right) \quad (5.57)$$

$$= E \left[ |\hat{h}|^4 \right] \text{tr} \left( \Phi^2 \right) \text{tr} \left( \Theta^2 \right). \quad (5.58)$$

The last equality (5.58) is true, because $\text{tr} \left( \Phi^2 \right) = \text{tr} \left( Q \Lambda Q^\dagger Q \Lambda Q^\dagger \right) = \text{tr} \left( \Lambda^2 \right)$, which holds because $Q$ is unitary matrix.

Third case is given when $m = m'$ and $n \neq n'$ in the summation (5.56). We have

$$\sum_m \sum_{n,n'} \sum_{n \neq n'} E \left[ \left| \hat{H}(m,n) \right|^2 \right] E \left[ \left| \hat{H}(m,n') \right|^2 \right] \Lambda(n,n) |D(m,m)|^2 \Lambda(n',n') \quad (5.59)$$

$$= E \left[ |\hat{h}|^2 \right]^2 \sum_m |D(m,m)|^2 \sum_{n,n'} \sum_{n \neq n'} \Lambda(n,n) \Lambda(n',n') \quad (5.60)$$

Notice that the last double summation over $n$ and $n'$ is in fact a square of $\sum_n \Lambda(n,n)$,
except for the terms when \( n = n' \). To complete the third case, (5.60) is expressed as

\[
E \left[ |\hat{h}|^2 \right]^2 \text{tr} (D^2) \left[ \sum_{n, n'} \Lambda(n, n)\Lambda(n', n') \right] - \sum_{n} \Lambda^2(n, n) \tag{5.61}
\]

\[
= E \left[ |\hat{h}|^2 \right]^2 \text{tr} (D^2) \left[ (\text{tr} \Lambda)^2 - \text{tr} (\Lambda^2) \right] \tag{5.62}
\]

\[
= E \left[ |\hat{h}|^2 \right]^2 \text{tr} (\Theta^2) \left[ (\text{tr} \Phi)^2 - \text{tr} (\Phi^2) \right]. \tag{5.63}
\]

The fourth case of (5.56) is when \( n = n' \) and \( m \neq m' \), which is similar to the third case. Hence by symmetry, we may interchange \( \Phi \) and \( \Theta \) in (5.63) to arrive at

\[
E \left[ |\hat{h}|^2 \right]^2 \text{tr} (\Phi^2) \left[ (\text{tr} \Theta)^2 - \text{tr} (\Theta^2) \right]. \tag{5.64}
\]

To compute the desired average of (5.56), we combine all three non-trivial cases, starting from (5.58), then (5.63), and finally (5.64). The addition of these three terms provides desired sum from Lemma 3, which completes its proof. \( \square \)

Now we proceed for the second part of the proof, using the two previous Lemmas.

First we give the minimum energy per information bit as follows

\[
\mathcal{E}_{k,\text{min}} = \frac{N_k \ln 2}{E \text{tr} \left[ \mathcal{H}_k \mathcal{H}_k^\dagger \right]} \tag{5.55}
\]

\[
= \frac{N_k \ln 2}{\text{tr} [\Theta_k] \text{tr} [\Phi_k] E \left[ |\hat{h}|^2 \right]} \tag{5.56}
\]

Next, we evaluate the Wideband Matrix.

First, we tackle diagonal terms, when \( k \neq l \). To evaluate the off-diagonal entries \( k \neq l \), of the wideband matrix \( \mathcal{C} \) we use (8.67), independence between \( \mathcal{H}_k \) and \( \mathcal{H}_l \), and Lemma 2, namely

\[
\mathcal{C}_{k,l} = \left( \frac{1}{2} \right) \frac{\text{tr} \left( E \left[ \mathcal{H}_k \mathcal{H}_k^\dagger \right] E \left[ \mathcal{H}_l \mathcal{H}_l^\dagger \right] \right)}{\text{tr} E \left[ \mathcal{H}_k \mathcal{H}_k^\dagger \right] \text{tr} E \left[ \mathcal{H}_l \mathcal{H}_l^\dagger \right]} \tag{5.67}
\]

\[
= \left( \frac{1}{2} \right) \frac{\text{tr} \left( \Phi_k \text{tr}(\Theta_k) E \left[ |\hat{h}_k|^2 \right] \Phi_l \text{tr}(\Theta_l) E \left[ |\hat{h}_l|^2 \right] \right)}{\text{tr} \left( \Phi_k \text{tr}(\Theta_k) E \left[ |\hat{h}_k|^2 \right] \right) \text{tr} \left( \Phi_l \text{tr}(\Theta_l) E \left[ |\hat{h}_l|^2 \right] \right)} \tag{5.68}
\]
Now, observe the all scalars in (5.68), like $\text{tr}(\Theta_k)$ and $E\left[|\hat{h}_k|^2\right]$ may be “pulled out” from the outer trace operation and canceled from both numerator and denominator.

This produces, for $k \neq l$

$$\mathcal{C}_{k,l} = \left(\frac{1}{2}\right) \frac{\text{tr}(\Phi_k \Phi_l)}{\left(\text{tr} \Phi_k\right) \left(\text{tr} \Phi_l\right)}, \quad (5.69)$$

which end the computation for off-diagonal entries of the wideband matrix $\mathcal{C}$.

Finally, to specify diagonal entries, we use the Lemma 3

$$\mathcal{C}_{k,k} = \frac{\text{tr}(\Phi_k^2) \text{tr}(\Theta_k^2) \left( E\left[|\hat{h}_k|^4\right] - 2E\left[|\hat{h}_k|^2\right]^2 \right)}{2 \text{tr}(\Phi_k \text{tr}(\Theta_k) \left| E\left[|\hat{h}_k|^2\right] \right| \text{tr}(\Phi_k \text{tr}(\Theta_k) \left| E\left[|\hat{h}_k|^2\right] \right|)} + \frac{\text{tr}(\Theta_k^2) \left( \text{tr} \Phi_k \right)^2 + \text{tr}(\Phi_k^2) \left( \text{tr} \Theta_k \right)^2 E\left[|\hat{h}_k|^2\right]^2}{2 \text{tr}(\Phi_k \text{tr}(\Theta_k) \left| E\left[|\hat{h}_k|^2\right] \right| \text{tr}(\Phi_k \text{tr}(\Theta_k) \left| E\left[|\hat{h}_k|^2\right] \right|)}. \quad (5.70)$$

The last expression is written as

$$\mathcal{C}_{k,k} = \left(\frac{1}{2}\right) \left[ \frac{\text{tr}(\Phi_k^2) \text{tr}(\Theta_k^2)}{(\text{tr} \Phi_k)^2 (\text{tr} \Theta_k)^2} \right] (\kappa_k - 2) + \left(\frac{1}{2}\right) \frac{\text{tr}(\Phi_k^2)}{(\text{tr} \Phi_k)^2} + \left(\frac{1}{2}\right) \frac{\text{tr}(\Theta_k^2)}{(\text{tr} \Theta_k)^2}, \quad (5.71)$$

which completes the proof of the Theorem 1. \qed

**Proof of Proposition 2.** Let $\Theta_k = QDQ^\dagger$ represent the eigenvalue decomposition of the transmit correlation matrix, for user $k$. We have $\text{tr}(F_k \Theta_k) = \text{tr}(F_k QDQ^\dagger) = \text{tr}\left([Q^\dagger F_k Q] \ D\right)$.

Because $D$ is diagonal, it follows that

$$\text{tr}\left([Q^\dagger F_k Q] \ D\right) = \sum_n [Q^\dagger F_k Q] \ (n, n) \ D(n, n) \quad (5.72)$$

$$\leq \text{tr}\left(Q^\dagger F_k Q\right) \max_n D(n, n). \quad (5.73)$$

$$= \text{tr}(F_k) \max_n D(n, n) = N_k \lambda_{max}(\Theta_k). \quad (5.74)$$
To achieve the upper bound we set \( F_k = Q \Lambda Q^\dagger \), where \( \Lambda = \text{diag}(\frac{N_k}{n_k}, \ldots, \frac{N_k}{n_k}, \ldots, 0) \) and non-zero entries correspond to those vectors within \( Q \), which produce maximum eigenvalues of \( \Theta_k \).
Chapter 6

Coarsely Coordinated Low Power Multiuser Systems

In coarsely coordinated multiuser systems, each user selects its own data rate and transmit power locally, without being directed by the access point, and without coordination with other users. In order to allow for adjustable data rate transmission, each user is allocated a set of low power codebooks, which we name the policy. Because users policies accommodate a range of small spectral efficiencies, each user is totally oblivious to particular codebook, which is chosen by other users.

In this chapter, we provide a systematic characterization and construction of all available policies, for each user. Our primary consideration, when designing users policies, is that reliable transmission remains preserved, regardless of which particular codebook any user selects, from its policy.

6.1 Coarsely Coordinated Users with Multiuser Detection Receiver

Focusing on a particular user $k$, each codebook inside its policy is described by transmit power, data rate, and a given typical distribution - which we assume i.i.d. complex Gaussian. In order to specify the entire policy of user $k$, we next describe how to select each codebook, inside the policy of user $k$.

To select each codebook, we utilize the payoff function, which is allocated to the
user k, and which relates variable data rate, to the adjustable transmit power. In low power multiple access systems, the payoff function is specified by utilising the Proposition 1, where $E_{\text{min},k}$ and $S_k$ are fixed. Given $E_{\text{min},k}$ and $S_k$, each codebook within users k policy is chosen i.i.d. complex Gaussian, with transmit power $P_k$ and data rate $R_k$, which are related as in Proposition 1.

By the described policy construction procedure, we only require $E_{\text{min},k}$ and $S_k$ to specify the entire policy, which is the set of low power codebooks, that are allocated to user k. Because data rates of other users are unknown - compound parameter - at each mobile, we will specify users policies, utilising the set of spectral efficiency slopes $S_k$, which are achieved independently from other users data rates.

This set is defined as follows, provided that the receiver utilizes the optimum multiuser detection.

**Definition 4 ([15]). Robust Slope Region of Multiuser Detection** is defined as intersection of slope regions $\text{MUD}(\mathcal{R})$, which is taken over all possible data rate proportions $\mathcal{R} = [R_1 \ldots R_K]$, namely

$$\text{MUD}_* = \bigcap_{\mathcal{R}} \text{MUD}(\mathcal{R}).$$

(6.1)

A visual interpretation of this approach is given in Figure 6.1, for $K = 2$ users, and single antenna transmission and reception. In Figure 6.1, the interior envelope of all regions $\text{MUD}(\mathcal{R})$ defines $\text{MUD}_*$, and slopes within $\text{MUD}_*$ are by construction valid for all proportions between users variable spectral efficiencies. The name “robust multiaccess slope region” is derived from this property, because $\text{MUD}_*$ is insensitive to variations in users data rates.

In coarsely coordinated systems, any pair of slopes from $\text{MUD}_*$ can be utilized to define both the policy, and the payoff function for each user. An example construc-
Figure 6.1: Robust Slope Region MUD$_*$ for single antenna transmission.

tion of users policies, for $K = 2$ users is given as follows, in a two-step procedure.

First, a single point, which is represented by a pair of slopes $(S_1, S_2)$, is chosen from MUD$_*$. For example, in reference to the Figure 2, we selected the point $A$. Second, now that $S_1$, and $E_{\text{min,1}}$ are fixed, a set of low power codebooks is chosen for user 1, with the help of Proposition 1. Namely, each codebook of user 1 with spectral efficiency $R_1$, is selected i.i.d. Gaussian, with power $P_1$, where $P_1$ and $R_1$ are related as in Proposition 1. Codebooks of user 2 are selected in a similar fashion, utilising $S_2$, and $E_{\text{min,2}}$.

This policy construction procedure ensures transmission reliability, within variable data rate environment, because slopes $(S_1, S_2)$, from MUD$_*$ are also within MUD($\mathcal{R}$), regardless of what direction vector $\mathcal{R}$ happens to be. Therefore, all multiple access constraints on spectral efficiency slopes $S_k$ will remain satisfied, no matter
which particular codebook users happen to select. Because all slope regions are derived from the rate regions, namely the Cover-Wyner pentagon, it follows that all multiple access constraints on data rates will always remain preserved, even if users do not coordinate the codebook selection.

In the next theorem, we provide a full closed-form characterization of the robust multiaccess slope region $\text{MUD}_\star$.

**Theorem 6 (Robust Region Characterization).** Given a Wideband Matrix $\mathbf{C}$, the robust slope region $\text{MUD}_\star$, which is achieved provided that receiver utilizes multiuser detection, is given as

$$\text{MUD}_\star = \{ \Sigma : \lambda_{\text{max}} \left( \Sigma^{1/2} \mathbf{C} \Sigma^{1/2} \right) \leq 1 \}. \quad (6.2)$$

*Proof.* Based on the Rayleigh quotient. See Appendix.

Next, we demonstrate the relationship between “generic” coarsely coordinated multiuser systems, as they were defined in Section II, and “low power” coarsely coordinated multiuser systems. For this purpose, we examine the Example 1, from Section II, but now from the perspective of low power communication.

**Example 5.** Assume $K = 2$ users, with channel model (2.9), where the matrix $\mathbf{C}$ for channel model (2.9) is found in (4.5). Symmetric point $A$, which is defined as $S_1 = S_2$, on the boundary of $\text{MUD}_\star$, as shown in Figure 6.1, represents policies for User 1 and User 2, back from the introductory Example 1.

To demonstrate this, we first evaluate the spectral efficiency slopes, which are achieved by users policies from Example 1, as follows

$$S_1 = \frac{2 \left( \hat{R}_1(0) \right)^2}{\hat{R}_1(0)} = \frac{2}{1 + \kappa(h)}, \quad (6.3)$$
where $R_1(P_1)$ is the payoff function for user 1, from Example 1. The same (6.3) holds for User 2, namely $\mathcal{S}_1 = \mathcal{S}_2$, because we assumed in Example 1 that all fading coefficients are identically distributed. Calculation to arrive at (6.3) is performed in the Appendix.

Second, to verify that slopes (6.3) are on the boundary of the robust slope region $\text{MUD}_*$, we note the following two facts. Fact one: from $\mathcal{S}_1 = \mathcal{S}_2$ and (6.3) we have that $\Sigma = 2[1 + \kappa(h)]^{-1} I$. Fact two: because we assumed that $\kappa(h_1) = \kappa(h_2)$, the Wideband matrix $\mathcal{C}$ from Example 3 and (4.5), satisfies $\lambda_{\text{max}}(\mathcal{C}) = (1/2) [1 + \kappa(h)]$, which is just single row-sum of $\mathcal{C}$.

We have

$$\lambda_{\text{max}} \left( \Sigma^{1/2} \mathcal{C} \Sigma^{1/2} \right) = \left[ \frac{2}{1 + \kappa(h)} \right] \lambda_{\text{max}}(\mathcal{C}) = 1,$$

(6.4)

and therefore slopes $\mathcal{S}_1 = \mathcal{S}_2$, which are achieved by both users policies from the introductory Example 1, belong to the boundary of the robust slope region $\text{MUD}_*$ as shown in Figure 6.1, point A.

The robust multiaccess slope region $\text{MUD}_*$ depends only on $\mathcal{C}$, which we see from Theorem 6. If $\mathcal{C}$ were simply diagonal, namely if $\mathcal{C} = \Sigma^{-1}_{\text{su}}$, then the robust slope region $\text{MUD}_*$ would occupy the entire box $\mathcal{S}_1 \leq \mathcal{S}_{1,\text{su}}$, and $\mathcal{S}_2 \leq \mathcal{S}_{2,\text{su}}$ from the Figure 6.1. In this case, we may utilize $\mathcal{S}_1 = \mathcal{S}_{1,\text{su}}$, with $\mathcal{S}_2 = \mathcal{S}_{2,\text{su}}$, in order to design policies for users 1 and 2. Here, the policy of each user is designed as if the other user were non-existent, which is analogous to the scenario where users transmissions are in orthogonal subspaces. In fact, we will show below that if users transmissions were orthogonal, then $\mathcal{C}$ becomes exactly diagonal.

Heuristically, if $\mathcal{C}$ is close to diagonal, which occurs when $\Sigma_{k,\text{su}}^{-1}$ dominate off-diagonal entries by a substantial amount, we expect each users policy to be able to
achieve \( S_1 \approx S_{1,\text{su}} \), with \( S_2 \approx S_{2,\text{su}} \), and there is really no need for user coordination, even if it were available. To demonstrate this point, we revisit the Example 5 above, for the channel model (2.9).

If user 2 were absent from (2.9), then policy (set of codebooks) of user 1 would be designed with \( \mathcal{E}_{\min,1} \) and \( S_1 = 2 [\kappa(h)]^{-1} = S_{1,\text{su}} \). Once user 2 arrives, then policy of user 1 is re-designed with \( S_1 = 2 [1 + \kappa(h)]^{-1} \). If \( \kappa(h) \) is large as compared to 1, then addition of user 2 makes negligible impact on the policy of user 1, namely \( S_1 \approx S_{1,\text{su}} \). This resembles a scenario where user 2 is simply added inside a subspace, which is orthogonal to user 1, because users are free to adjust their data rates locally with no coordination. With a relatively large kurtosis, the independently faded channel of two users effectively performs user selection, like in TDMA, by amplifying user 1, while attenuating user 2, and vice-versa.

6.2 Coarse User Coordination With Antenna Arrays

Given a receiver antenna array, one straightforward solution for coarsely coordinated users can be drawn from analogy with classic orthogonal division systems, and the notion of parallel channels. Namely, a single design principle governs time, frequency, and orthogonal code division, where users are assigned orthogonal subspaces, from the vector space of time-frequency limited waveforms [20]. Such assignment makes signals from different users easily separated, by subspace projections at the receiver. Here, users can select their own data rates, as long as their transmissions do not infringe into subspaces, which are dedicated to other users.

Subspace projections can be exploited with the receiver antenna array as well, in order to establish parallel (orthogonal) channels, and thereby enable coarse user coordination. However, a more careful scrutiny will reveal that this straightforward
approach is severely suboptimal method for enabling coarse user coordination. To
demonstrate this point, we review the classic antenna array subspace projection
strategy, which is the zero-forcing* receiver. When detecting symbols of user 1, the
zero-forcing receiver first projects signal from the antenna array, onto the largest sub-
space, which is orthogonal to the subspace of multiple access interference. Because
multiple access interference is nulled away, the zero-forcing receiver enables coarse
user coordination, and users do not have to coordinate the choice of codebooks.

**Example 6 (Subspace Projections).** Consider the multiple access system (2.1)
with \( K = 2 \) users, and where each user has only one transmit antenna \( N_k = 1 \). The
receiver is assumed to be composed of \( M = 2 \) receive antennas, and all propagation
coefficients are Rayleigh distributed, which have \( \kappa(h) = 2 \). Note that both \( \mathcal{K}_1 \) and
\( \mathcal{K}_2 \) are column vectors of size \( 2 \times 1 \), as in Figure 6.2. Finally, suppose that the
zero-forcing receiver is utilized to enable coarse user coordination.

We have

\[
Y = \mathcal{K}_1 X_1 + \mathcal{K}_2 X_2 + \eta \tag{6.5}
\]

\[
\mathcal{K}_2^\dagger Y = \mathcal{K}_2^\dagger \mathcal{K}_1 X_1 + \mathcal{K}_2^\dagger \eta. \tag{6.6}
\]

Here, \( \mathcal{K}_2^\dagger \) denotes the zero-forcing projection, which is simply the unique unit-norm
row vector, that satisfies \( \mathcal{K}_2^\dagger \mathcal{K}_2 = 0 \). Zero-forcing receiver produces the effective channel (6.6) for user 1, which is now free of multiple access interference. To construct
the policy of user 1, we examine characteristics of the channel (6.6).

For any particular realization of \( \mathcal{K}_2^\dagger \), the random variable \( \mathcal{K}_2^\dagger \mathcal{K}_1 \) is zero-mean
complex Gaussian, because it is a linear combination of random variables in \( \mathcal{K}_1 \).

*The name “zero-forcing” comes from the property of forcing the multiple access interference to
zero.
Variance of $\mathcal{H}_2^\perp \mathcal{H}_1$ is $E[\mathcal{H}_2^\perp \mathcal{H}_1 (\mathcal{H}_2^\perp \mathcal{H}_1)^\dagger | \mathcal{H}_2^\perp]$ which equals $\mathcal{H}_2^\perp E[\mathcal{H}_1 \mathcal{H}_1^T] (\mathcal{H}_2^\perp)^\dagger = \mathcal{H}_2^\perp (\mathcal{H}_2^\perp)^\dagger = 1$. Probability distribution $p(\mathcal{H}_2^\perp | \mathcal{H}_1^\perp)$ is therefore $\mathcal{N}_c(0, 1)$, for any $\mathcal{H}_1^\perp$. Averaging $p(\mathcal{H}_2^\perp | \mathcal{H}_1^\perp)$ over $p(\mathcal{H}_2^\perp)$ results in $p(\mathcal{H}_2^\perp | \mathcal{H}_1)$, which remains $\mathcal{N}_c(0, 1)$.

The same argument extends for $\mathcal{H}_2^\perp \eta$, which is also unit variance complex Gaussian. Furthermore, the noise random variable $\mathcal{H}_2^\perp \eta$ is independent from the channel $\mathcal{H}_2^\perp \mathcal{H}_1$ because they are both complex Gaussian and uncorrelated, namely it holds that $E[\mathcal{H}_2^\perp \mathcal{H}_1 (\mathcal{H}_2^\perp \eta)^\dagger] = 0$, which follows from $E[\eta] = 0$.

In essence, channel (6.6) satisfies $I(\mathcal{H}_2^\perp Y; X_1) = E \log(1 + |\mathcal{H}_2^\perp \mathcal{H}_1|^2 P_1)$, which is equivalent to the single-antenna Rayleigh faded channel, with additive noise of unit variance. User 1 achieves

$$E_{\min,1} = \frac{\ln 2}{E[|\mathcal{H}_2^\perp \mathcal{H}_1|^2]} = \ln 2,$$

(6.7)

and hence, channel (6.6) does not produce the optimum value of minimum energy per information bit. Namely, from (5.7) we conclude that multiuser detection reduces (6.7) by additional factor of 2.

By sacrificing $E_{\min,1}$ the zero forcing receiver (6.6) enables coarse user coordination, and we compute that it warrants the spectral efficiency slope of

$$S_1 = \frac{2}{\kappa (\mathcal{H}_2^\perp \mathcal{H}_1)} = 1.$$

(6.8)

Namely, if receiver is designed as (6.6), then policy of user 1 is constructed using $E_{\min,1}$ from (6.7) and $S_1$ from (6.8).

By symmetry of the problem, if $\mathcal{H}_1^\perp$ is applied to (6.5) before detecting $X_2$, then user 2 would experience the same performance of (6.7) and (6.8). To summarize, the zero-forcing projection receiver does enable coarse user coordination, only at a severe cost in terms of $E_{\min,k}$. 

Example 6 is best portrayed by the Figure 6.2, where the zero forcing projection receiver converts the multiaccess channel into two orthogonal channels, in order to enable coarse user coordination. Here, each channel achieves the minimum energy per information bit (6.7), and the spectral efficiency slope (6.8) of a single antenna Rayleigh faded channel, which is described in Figure 6.2. Explanation for suboptimality of the conversion, which is portrayed in Example 6 and Figure 6.2, is simple: because of rapid random channel variations - fading, spatial dimensions cannot be explicitly dedicated to some users, and protected from others by orthogonalising. Here, the subspace projection at the receiver (6.6) eliminates a significant portion (half) of signal power, besides eliminating multiple access interference, which comes from user 2.

In order to contrast Example 6 against the optimal information theoretic solution for coarsely coordinated users, with antenna arrays, we distinguish between two disparate forms of advantages, which are offered by the receive antenna array, when compared to the single antenna reception.

- Antenna Arrays offer a boost in aggregate received power, for each user. This is portrayed by the expression (5.7) for the minimum energy per information
bit, which is inversely proportional to the number of receivers. This form of gain does not require independent fading.

- Antenna Arrays offer an increase in receiver dimensionality. Namely, if (and only if) receivers are sufficiently separated and suffer independent fading, then different signals are induced in different antennas, and the receiver is equipped with more dimensions to combat multiple access interference.

The zero-forcing receiver, which is based on subspace projections, partitions both of the above forms of advantages, between the two users, in order to enable coarse coordination. Namely, by operation (6.6), user 1 is stripped from half of the received power, and from half of the aggregate receive dimension, all of which is portrayed in Figure 6.2.

In contrast to this orthogonal solution, we demonstrate how policies which are represented by the optimum $\mathcal{E}_{\text{min},k}$, and $S_k$ from MUD, maintain the aggregate received power, while dividing only the receiver dimensionality among users. This approach is portrayed in Figure 6.3, which ought to be contrasted against Figure 6.2. To clarify this point, we next elaborate on how to represent each policy, by an equivalent single-user channel, which is dedicated to a particular user.

**Dedicated Channel Representation**

It is irrelevant for user 1 whether communication takes place in a multiaccess system which is designed in Example 6, or whether communication takes place in a single-user system, having one transmitter and one receiver - as in Figure 6.2. The very same set of codebooks, with the very same payoff function, is applicable for both systems, and from the perspective of user 1, the two systems are equivalent. Here,
Figure 6.3: Policies Chosen From MUD, Represent Dedicated Channels.

the pair of policies from the multiaccess channel in Example 6 is represented by a pair of dedicated single-user channels from Figure 6.2. By equating the payoff function of a policy with the capacity function of a single-user channel, we provide a policy with a dedicated single-user channel representation.

Dedicated channel representation of a policy gives the most intuitive perspective on the coarsely coordinated multiuser problem, because a pair of policies represents partition of the multiaccess channel into a pair of dedicated\textsuperscript{\textdagger} single-user channels. Here, each user exploits its dedicated channel, and doesn’t coordinate with other users.

In low power systems, the entire policy is defined using fixed $E_{\text{min},k}$ and fixed $S_k$. Dedicated channel representation is given by finding a single-user channel, which achieves that particular $E_{\text{min},k}$ and $S_k$.

\textsuperscript{\textdagger}Authors of this text would have preferred to dub these channels “Parallel Channels,” rather than “Dedicated Channels;” however, the keyword Parallel Channels is already reserved for Orthogonal Channels - see a classic text from Cover [17]. In order not to deviate from the standard terminology, we use Parallel Channels and Orthogonal Channels as synonyms. The major point of our work is that Dedicated Channels need not be Orthogonal Channels.
Example 7. Consider the multiaccess channel (6.5) from Example 6, which has $K = 2$ users, $N_k = 1$ transmit and $M = 2$ receive antennas. In contrast to Example 6, we now let coarse user coordination be enabled by policies, which are selected from the optimum $\mathcal{E}_{\text{min},k}$ in (5.7) and the point $\mathcal{S}_1 = \mathcal{S}_2$ from MUD$_*$ in (5.8) and (5.9).

By (5.7), the minimum energy per information bit of each users policy equals

$$\mathcal{E}_{\text{min},k} = \frac{\ln 2}{2},$$

(6.9)

which corresponds to $M = 2$ receivers. Next, by setting $\mathcal{S}_1 = \mathcal{S}_2$, the operating point on MUD$_*$ is uniquely determined to be $\mathcal{S}_k = \lambda_{\text{max}}^{-1}(\mathcal{C})$. Using (5.8) and (5.9), it follows that $\lambda_{\text{max}}(\mathcal{C}) = 1$. Each users policy is constructed with the spectral efficiency slope

$$\mathcal{S}_k = \lambda_{\text{max}}^{-1}(\mathcal{C}) = 1,$$

(6.10)

which corresponds to a single Rayleigh faded receive antenna.

Dedicated channel representation for policies which are selected from (6.9) and (6.10) is given in Figure 6.3. Namely, the capacity function for each single-user channel from Figure 6.3 is given as $E \log(1 + 2|h|^2 P_1)$, which achieves the minimum energy per information bit of (6.9), and the spectral efficiency slope of (6.10). This is verified from the single-user slope formula (4.1), and noting that the first derivative of $E \log(1 + 2|h|^2 P_1)$ is simply $2E[|h|^2] = 2$, while the second derivative is $-4E[|h|^4]$, and the kurtosis $\kappa(h) = 2$.

From the perspective of user 1, the multiaccess system which is designed in Example 7 above is no different from the single-user system in Figure 6.3, because a single set of low power codebooks is applicable for both systems. Namely, if coarse user coordination is designed according to Example 7, then each user effectively sees single-user channel from the Figure 6.3.
Figure 6.3 also illustrates how policies from Example 7, which are designed using the $S_1 = S_2$ operating point from MUD$_*$, divide the aggregate receiver space between the two users. Nevertheless, MUD$_*$ is not achieved by straightforward subspace projection receiver from Example 6, which sacrifices users power in order to partition the receiver space as in Figure 6.2. Rather, achieving MUD$_*$ requires more complex multiuser detection.

Next, we demonstrate that the principle of partitioning receiver dimensions by MUD$_*$ extends beyond the Example 7 and Figure 6.3. The following Corollary to Theorem 6 generalizes the Example 7 for $K$ users, with $N_k$ transmit antennas each and a total of $M$ receivers. Current working assumption is that all coefficients are uncorrelated and Rayleigh distributed.

**Corollary 5 (Partitioning Receiver Dimensions).** Consider the $K$-user channel (2.1), where user $k$ is provided with $N_k$ transmit antennas, and the access point with a total of $M$ receivers. Here, the size of $\mathcal{H}_k$ in (2.1) is $M \times N_k$. Furthermore, assume that all fading coefficients inside $\mathcal{H}_k$ are uncorrelated and complex Gaussian, with kurtosis $\kappa_k = 2$ for each user. Let

\[
M = M_1 + M_2 + \ldots + M_K, \tag{6.11}
\]

represent any partition of $M$ receivers into $K$ groups.

The rectangle specified by

\[
\frac{1}{S_k} \geq \frac{1}{2N_k} + \frac{1}{2M_k}, \tag{6.12}
\]

for $k = \{1, \ldots, K\}$, belongs to the robust slope region MUD$_*$. Conversely, the region MUD$_*$ itself is union of all such rectangles, which is evaluated over all possible partitions (6.11).
Figure 6.4: Characteristics Points on the Robust Slope Region for 4 Receive Antennas

Proof. Special case of Corollary 6, whose Proof is in the Appendix.

The spectral efficiency slope (6.12) is achieved in a single-user system with \( N_k \) transmitters and \( M_k \) receivers - see (5.9). Therefore, a particular choice of receiver partition (6.11) provides policy of user \( k \) with a dedicated channel representation, which has a total of \( M_k \) received spatial dimensions. Similarly to Example 7, the \( \varepsilon_{\min,k} \) remains equivalent to entire \( M \) receivers, and dedicated single-user channel representation for policies from \( \text{MUD}_* \) first scales transmit power of each user by \( M/M_k \), and then grants \( M_k \) receive antennas to each user.

Dedicated channel representation for policies of \( K = 2 \) two users, with \( N_k = 2 \), and \( M = 4 \) is illustrated in Figures 6.4-6.7, for various choices of receiver partition (6.11). Point A is equivalent to the extreme case where all receive dimensions are
allocated to User 1, and User 2 can have no transmission whatsoever, namely $S_2 = 0$. If policies, i.e., spectral efficiency slopes, are divided according to Point B, User 1 is given slope $S_1$ equivalent to three receive dimensions, whereas User 2 is granted slope corresponding to only one receive dimensions. Point C is the symmetric and fair case of equal slopes.

To summarize, Corollary 5 represents parametrization of the robust multiaccess slope region $\text{MUD}_*$, by partitioning the aggregate receiver space dimension $M$. Here, (6.12) is represented by $K$ dedicated channels, where each channel has the spectral efficiency slope $S_k$, which is equivalent to $N_k$ transmitters and $M_k$ receivers. On the other hand, the channel gain on each channel is equivalent to whole $M$ receivers, because $E_{\text{min},k}$ is not affected by multiple access.

\footnote{It is interesting to note that $\text{MUD}_*$ also represents users “freedom” to select their own data rate without coordinating, which establishes a philosophical link with the degrees of freedom concept.}
Figure 6.6: Dedicated Channels - Point B on MUD$_s$ from Figure 6.4

One similarity between transmit/receive antenna arrays and parallel channels was first pointed out in seminal works of Telatar [32] and Foschini [33], where authors established that antenna array transmission and reception has aggregate (sum) capacity, which equals the sum capacity of parallel channels. Corollary 5 extends this similarity, between antenna arrays and parallel channels, one step beyond the sum capacity, because parallel channels can be dedicated, and do not require user coordination between multiple users (for example, frequency division). Here, Corollary 5 demonstrates how to optimally use antenna arrays to establish dedicated channels, in low power multiuser systems.
Correlated Antenna Arrays

The working assumption for dedicated channel representation of users policies, in Corollary 5, was that all antenna coefficients are uncorrelated and Rayleigh distributed. Here, we remove this assumption and define the notion of effective receiver dimension for correlated antenna arrays, which are not necessarily Rayleigh faded. The receive antenna correlation $\Phi_k$ is assumed to be the same for all users $\Phi = \Phi_k$.

We will show that $\omega$, which is defined as

$$\omega \triangleq \frac{(\text{tr} \, \Phi)^2}{\text{tr} (\Phi^2)} \quad \text{[Uncorrelated Receivers]},$$

(6.13)

represents the effective number of spatial dimensions, which are available at the receiver. Naturally, the unit of $\omega$ [Uncorrelated Receivers] in (6.13), is chosen to reflect this property.
To illustrate this point, we first note that spectral efficiency slope of a single-user channel, whose inverse is on the diagonal of the Wideband Matrix $\mathcal{C}$ in (5.6) (5.12) (5.16), only depends on the receive correlation $\Phi_k = \Phi$ through $\omega$ in (6.13). Following proposition addresses the range of $\omega$.

**Proposition 3.** Let $\omega$ be defined as in (6.13).

The following tight upper and lower bounds hold

$$1 \leq \omega \leq M. \quad (6.14)$$

Furthermore, $\omega = M$ achieves the maximum receiver dimension, if and only if all receive antennas are fully uncorrelated, namely $\Phi = I$. From the other side, $\omega = 1$ is the least if all receive antennas are perfectly correlated, when which holds when all entries of $\Phi$ are identical.

**Proof.** See Appendix. \hfill \square

All relevant receiver array characteristics, which are needed to evaluate the single-user spectral efficiency slope, are captured by the real number $\omega$, which ranges as $1 \leq \omega \leq M$. Here, $\omega$ represents the effective number of uncorrelated antennas, which is the aggregate receiver space dimension; this interpretation will extend for multiple access systems as well.

From the other side, if transmitter array of user $k$ is correlated as $\Theta_k$, and if users do not exploit the transmit correlation (5.12), then we define the effective transmit array dimension

$$\vartheta_k \triangleq \frac{(\text{tr} \, \Theta_k)^2}{\text{tr} \, (\Theta_k^2)} \quad \text{[Uncorrelated Transmitters].} \quad (6.15)$$

Namely, a single real number $\vartheta_k$ captures the entire transmitter array contribution to the spectral efficiency slope, and furthermore $1 \leq \vartheta_k \leq N_k$. 

To address multiuser channels, we will demonstrate that robust multiaccess slope region \( \mathsf{MUD}_\kappa \) is parameterized by a partition of the aggregate receiver dimension \( \omega \). First, we note the structure of the Wideband Matrix (5.5) when receive antennas are equally correlated \( \Phi_k = \Phi \) for each user. Namely, \( \mathcal{C} \) is given as

\[
\mathcal{C}_{k,l} = \frac{1}{2\omega} \quad \text{for} \quad k \neq l. \tag{6.16}
\]

\[
\mathcal{C}_{k,k} = \frac{1}{\mathcal{S}_{k,\text{su}}} = \frac{\kappa_k - 2}{2\omega \vartheta_k} + \frac{1}{2\omega} + \frac{1}{2\vartheta_k}. \tag{6.17}
\]

There are three distinct physical quantities which define the Wideband Matrix \( \mathcal{C} \) in (6.16) and (6.17). First, off-diagonal terms are entirely captured by the aggregate receive dimension \( \omega \), which is purely the receiver characteristic. Second, besides \( \omega \), diagonal terms also depend on the aggregate transmit dimension of user \( k \), namely \( \vartheta_k \), which is purely a transmitter characteristic. Third, fading nature of the communication channel contributes to diagonal terms, with the kurtosis \( \kappa_k \).

Robust multiaccess slope region \( \mathsf{MUD}_\kappa \) is parameterized by dividing the aggregate receiver dimension \( \omega \).

**Corollary 6 (Dividing Receiver Dimension).** Suppose \( \mathcal{C} \) from (6.16) and (6.17).

The robust slope region can be parameterized by partitioning \( \omega \) into \( \omega_k \) as follows

\[
\mathsf{MUD}_\kappa = \bigcup_{\omega = \sum_k \omega_k} \left\{ (S_1, \ldots, S_K) : \frac{1}{S_k} - \frac{1}{\mathcal{S}_{k,\text{su}}} \geq \frac{1}{2\omega_k} - \frac{1}{2\omega} \right\}. \tag{6.18}
\]

**Proof.** See Appendix. \( \square \)

In systems which have large effective transmit and receive array dimensions, contribution of the communication channel in terms of kurtosis \( \kappa_k \) becomes negligible. Namely, we note that if both \( \omega \to \infty \) and \( \vartheta_k \to \infty \), then the cross term which is associated with \( \kappa_k - 2 \) in (6.17) vanishes with respect to the two other terms.
The MUD$_*^*$ slope region grows as

$$\text{MUD}_*^* \rightarrow \bigcup_{\omega = \sum_k \omega_k} \left\{ (S_1, \ldots, S_K) : \frac{1}{S_k} \geq \frac{1}{2\omega_k} + \frac{1}{2\vartheta_k} \right\}, \quad (6.19)$$

as transmit and receive antenna dimensions grow, namely $\omega \rightarrow \infty$ and $\vartheta_k \rightarrow \infty$. Here, dedicated single-user channel representation of each users policy is much similar to the one from Corollary 5 and Figures 6.4-6.7. Namely, spectral efficiency slopes from (6.19) are achieved in a single-user channel, where user $k$ communicates having effectively $\vartheta_k$ [Uncorrelated Transmitters], to the receiver array of $\omega_k$ [Uncorrelated Receivers].

6.3 Coarsely Coordinated Users with Single User Detection Receiver

Here, we evaluate the superposition multiple access strategy, assuming that the receiver is composed of a bank of simple single user detectors, and we investigate whether this receiver structure will enable coarse user coordination. We adopt the same approach as before, where we first define, and then evaluate the robust multi-access slope region.

Definition 5. Robust Slope Region of Single User Detectors is defined as intersection of slope regions $\text{SUD}(\mathcal{R})$, which is evaluated over all possible direction vectors $\mathcal{R} = [\mathcal{R}_1 \ldots \mathcal{R}_k]$, namely

$$\text{SUD}_*^* = \bigcap_{\mathcal{R}} \text{SUD}(\mathcal{R}). \quad (6.20)$$

In most cases, $\text{SUD}_*^*$ is a set which contains only the zero element, and we first illustrate this point, by re-visiting the Example 4. Consider the first constraint (4.15)
from the Example 4, which we re-write as

$$\frac{1}{S_1} \geq \frac{1}{S_{1, su}} + \frac{R_2}{R_1}. \quad (6.21)$$

$S_1^{-1}$ from SUD$_*$ must satisfy (6.21) for every ratio of $R_2$ to $R_1$, and users are allowed to select $R_2$ and $R_1$, so that $R_2/R_1$ is arbitrary large. It follows that $S_1^{-1}$ must be greater than any real number, and therefore $S_1 = 0$. By symmetry, this argument holds for $S_2$ as well. For Example 4, we conclude that SUD$_*$ contains only the $(0, 0)$ element.

The same conclusion extends to $K$ users, and one may observe this by considering Theorem 4: as long as a particular row, for example row one, of $C$ has even a single off-diagonal non-zero entry, this restricts $S_1 = 0$. This was demonstrated for two-user case, but additional interference from other $K - 2$ users, certainly can't improve $S_1$.

**Theorem 7 (Robust Slope Region of Single User Detection).** If each row of $C$ has even a single positive off-diagonal element, then this object is a set which contains only the zero vector

$$\text{SUD}_* = \{0\} \quad (6.22)$$

This pessimistic conclusion is drawn because we allow for a large disproportion between users spectral efficiencies, and hence $R_2/R_1$ is allowed to be any real number. The conclusion of (6.22) may be re-visited by considering a finite number of codebooks, within the policy of each user, where the slope of a particular user from SUD$_*$ becomes defined by the highest data rate codebooks of other users, and the lowest data rate codebook of that user, much like $R_2/R_1$ is the worst-case for user 1, when $R_2$ is the largest, while $R_1$ is the least in (6.21).
6.4 Appendix: Chapter Proofs

Proof of Theorem 6. We will require the following Lemma.

Lemma 4. First partials of $C_A(P_A)$ are non-negative - as any user boosts transmit power, the sum capacity increases. Also, second partials of $C_A(P_A)$ are not positive

$$\frac{\partial^2 C(P)}{\partial P_k \partial P_l} \bigg|_{P=0} \leq 0, \quad (6.23)$$

which makes all entries of the Wideband matrix $\mathbf{C}$ positive.

Proof. Rather than explicit computation, both claims from Lemma 4 are demonstrated using elementary mutual information properties.

It suffices to use $K = 2$, to avoid much unnecessary indexing. The first partial of $C(P_1, P_2)$ with respect to $P_1$, evaluated at $(0, 0)$, may be approximated as

$$\frac{\partial C(P_1, 0)}{\partial P_1} \bigg|_{P_1=0} = \frac{C(P_1, 0) - C(0, 0)}{P_1} + o(P_1). \quad (6.24)$$

First partials must be positive because $C(P_1, 0) \geq 0$ and $C(0, 0) = 0$.

To prove that second partials cannot be positive, we first find the first partial of $C(P_1, P_2)$ with respect to $P_1$, but now evaluated at $(0, P_2)$, which is written similarly to (6.24), namely

$$\frac{\partial C(P_1, P_2)}{\partial P_1} \bigg|_{P_1=0} = \frac{C(P_1, P_2) - C(0, P_2)}{P_1} + o(P_1). \quad (6.25)$$

Cross partial equals the limit of difference between (6.25) and (6.24), normalized by $P_2$. Order in which limits are evaluated is unimportant, and we have

$$\frac{\partial^2 C(0)}{\partial P_1 \partial P_2} = \lim_{P_2 \to 0} \lim_{P_1 \to 0} \frac{C(P_1, P_2) - C(0, P_2)}{P_1} + o(P_1) - \frac{C(P_1, 0) - C(0, 0)}{P_2} + o(P_1). \quad (6.26)$$
Significant terms of the numerator in (6.26) are

\[ C(P_1, P_2) - C(0, P_2) - [C(P_1, 0) - C(0, 0)] \]  
(6.27)

\[ = C(P_1, P_2) - C(0, P_2) - C(P_1, 0) \]  
(6.28)

\[ = I(X_1, X_2; Y) - I(X_2; Y|X_1) - I(X_1; Y|X_2) \]  
(6.29)

\[ = I(X_1; Y) + I(X_2; Y|X_1) - I(X_2; Y|X_1) - I(X_1; Y|X_2) \]  
(6.30)

\[ = I(X_1; Y) - I(X_1; Y|X_2) \leq 0. \]  
(6.31)

Here (6.29) is simply because \( C(P_1, P_2) \) is the mutual information - see Notation, and (6.30) is by the chain rule for mutual information [17]. Inequality in (6.31) is because conditioning does not reduce mutual information [17].

With (6.31), it becomes apparent that the limit in (6.26) can not be positive. \( \square \)

Now we proceed to prove the Theorem 6, first achievability and then converse.

**Proof of Achievability.** Any diagonal matrix of spectral efficiency slopes \( \Sigma \), which belongs inside (6.2) satisfies the following: for an arbitrary direction vector \( R \), and an arbitrary \( A \subseteq \{1, ..., K\} \), we have that

\[ \begin{bmatrix} R_A^T \Sigma_A^{-\frac{1}{2}} \end{bmatrix} \begin{bmatrix} \Sigma_A^{-\frac{1}{2}} \mathcal{C}_A \Sigma_A^{-\frac{1}{2}} \end{bmatrix} \begin{bmatrix} \Sigma_A^{-\frac{1}{2}} R_A \end{bmatrix} \leq \begin{bmatrix} R_A^T \Sigma_A^{-\frac{1}{2}} \end{bmatrix} \begin{bmatrix} \Sigma_A^{-\frac{1}{2}} R_A \end{bmatrix}. \]  
(6.32)

Claim (6.32) is true because maximum eigenvalue of \( \Sigma_A^{\frac{1}{2}} \mathcal{C}_A \Sigma_A^{\frac{1}{2}} \) is dominated by 1, which follows from the fact that the maximum eigenvalue of \( \Sigma_A^{\frac{1}{2}} \mathcal{C} \Sigma_A^{\frac{1}{2}} \) is dominated by 1. Note the indexing technicality in (6.32). Because \( \Sigma^{\frac{1}{2}} \) is diagonal, we have used \( \begin{bmatrix} \Sigma_A^{\frac{1}{2}} \mathcal{C}_A \Sigma_A^{\frac{1}{2}} \end{bmatrix} \) instead of \( \begin{bmatrix} \Sigma^{\frac{1}{2}} \mathcal{C} \Sigma^{\frac{1}{2}} \end{bmatrix}_A \). Both are the same because multiplication by diagonal matrix scales either rows or columns, depending on whether it is from right or from left.

From (6.32), it algebraically follows that

\[ R_A^T \mathcal{C}_A R_A \leq R_A^T \Sigma_A R_A, \]  
(6.33)
which means that $\Sigma$ belongs to MUD($R$), for an arbitrary $R$. This statement is made in reference to MUD($R$) from (4.4). Now we are left to prove the inclusion in the reverse direction.

Proof of Converse. This side shows that we were not too conservative in specifying MUD$_*$ with (6.2). Suppose that some $\Sigma$ violates the constraint in (6.2), and we show that there exists an $R$, so that particular $\Sigma$ doesn’t belong to MUD($R$).

Because $\Sigma$ violates (6.2), we have that for some $z$, possibly with negative entries

$$ z^T \left( \Sigma^{\frac{1}{2}} \mathcal{C} \Sigma^{\frac{1}{2}} \right) z \geq z^T z \tag{6.34} $$

$$ \Rightarrow \quad \langle z \rangle^T \left( \Sigma^{\frac{1}{2}} \mathcal{C} \Sigma^{\frac{1}{2}} \right) \langle z \rangle \geq \langle z \rangle^T \langle z \rangle \tag{6.35} $$

We denoted $\langle z \rangle$ to be the vector of absolute values of $z$. The reason for (6.35) is that the right hand side of (6.35) sums all positive entries, and hence dominates (6.34), because the Wideband Matrix $\mathcal{C}$ has all non-negative entries. No harm is therefore done if we constrain $z$ to be all positive.

We have found $R$, which is given with $z = \Sigma^{-\frac{1}{2}}R$, and for that particular $R$, we have that $\Sigma$ is not inside MUD($R$). Demanding positive semi-definiteness is therefore not only sufficient, but necessary as well.

This completes the proof of Theorem 6.

Proof of Example 5. One method is straightforward computation, namely

$$ 2 \dot{R}_1(0) = \frac{d}{dP_1} \text{E} \ln \left[ 1 + (|h_1|^2 + |h_2|^2)P_1 \right] \bigg|_{P_1=0} \quad \text{E} \mid_{P_1=0} $$

$$ = \text{E} \left[ |h_1|^2 + |h_2|^2 \right] = \text{E} \left[ |h|^2 \right] \tag{6.36} $$

$$ \ddot{R}_1(0) = \frac{d^2}{dP_1^2} \frac{1}{2} \text{E} \ln \left[ 1 + (|h_1|^2 + |h_2|^2)P_1 \right] \bigg|_{P_1=0} \quad \text{E} \mid_{P_1=0} $$

$$ = -\frac{1}{2} \text{E} \left[ |h_1|^4 + |h_2|^4 + 2|h_1|^2|h_2|^2 \right] = -\text{E} \left[ |h|^4 \right] - E^2 \left[ |h|^2 \right]. \tag{6.37} $$
Alternatively, one could observe that the payoff function of each policy is represented as a capacity function of a channel with two transmit antennas, and half of the bandwidth (or time). From here one could use standard expressions for slope. □

Proof of Corollary 6. We will require the following Lemma.

**Lemma 5.** If \( a_k \) and \( b_k \) are two sets of real numbers, then

\[
\sum_{k=1}^{K} \frac{b_k^2 a_k^{-1}}{\left[ \sum_{k=1}^{K} b_k \right]^2} \geq \left[ \sum_{k=1}^{K} a_k \right]^{-1}.
\]  

(6.38)

Furthermore, the minimum of the left hand side of (6.38) over all \( b_k \), is equal to the right hand side of (6.38). This minimum is achieved when vector \([b_k]_{k=1}^{K}\) is directly proportional to the vector \([a_k]_{k=1}^{K}\).

**Proof.** We have that

\[
\begin{bmatrix}
\sum_{k=1}^{K} b_k^2 a_k^{-1} \\
\sum_{l=1}^{K} a_l
\end{bmatrix}
= \sum_{k=1}^{K} b_k^2 + \sum_{k \neq l} b_k^2 a_k^{-1} a_l
\]

(6.39)

\[
= \sum_{k=1}^{K} b_k^2 + 2 \sum_{k < l} b_k^2 a_k^{-1} a_l + b_l^2 a_l^{-1} a_k.
\]

(6.40)

The left hand side of (6.39) is in fact double summation over \( \sum_k \sum_l \). The right hand side of (6.39) simply makes a distinction between diagonal terms from the double summation \( \sum_k \sum_l \), which are expressed as \( l = k \), and off-diagonal terms which are expressed as \( l \neq k \). To get from (6.39) to (6.40) we simply noted that off-diagonal terms in (6.39) duplicate themselves when \( l \) and \( k \) trade places, and the summation \( k < l \) in (6.40) is simply over the lower diagonal.

To proceed from (6.40) we use the inequality \( x + y \geq 2\sqrt{xy} \iff (\sqrt{x} - \sqrt{y})^2 \geq 0 \), which holds with equality if and only if both \( x \) and \( y \) are equal. Using \( x + y \geq 2\sqrt{xy} \).
for the off-diagonal terms $\sum_{k<l}$ in (6.40), we have that
\begin{equation}
\begin{bmatrix}
\sum_{k=1}^{K} b_k^2 a_k^{-1}
\end{bmatrix} \begin{bmatrix}
\sum_{l=1}^{K} a_l
\end{bmatrix} \geq \sum_{k} b_k^2 + 2 \sum_{k<l} b_k b_l \geq \left( \sum_{k=1}^{K} b_k \right)^2.
\end{equation}
(6.41)
(6.42)

The chain (6.41) - (6.42) is equivalent to the inequality (6.38). Finally, the equality holds between (6.40) and (6.41) if and only if $b_k^2 a_k^{-1} a_l = b_l^2 a_l^{-1} a_k$, for all $k$ and $l$. This is true if and only if vector $[b_k]_{k=1}^{K}$ is directly proportional to the vector $[a_k]_{k=1}^{K}$. \(\Box\)

For this case, we note that all off-diagonal entries of the Wideband matrix $\mathcal{C}$ are all the same, namely they are all equal to the receive array correlation $\omega$. Using this fact, the Wideband matrix can be expressed as follows
\begin{equation}
\mathcal{C} = \Sigma^{-1}_{su} - (2\omega)^{-1} I_{K \times K} + (2\omega)^{-1} 11^T,
\end{equation}
(6.43)
where $1$ is the column vector of all ones, and $\Sigma_{su}$ is a diagonal matrix of spectral efficiency slopes, in the single-user communication mode.

Using $\mathcal{C}$ from (6.43), the constraint which is given by MUD($\mathcal{R}$) becomes
\begin{equation}
\mathcal{R} \left[ \Sigma^{-1} - \Sigma^{-1}_{su} + (2\omega)^{-1} I_{K \times K} - (2\omega)^{-1} 11^T \right] \mathcal{R}^T \geq 0.
\end{equation}
(6.44)

Here, the left hand side of (6.44) is simply an inner product.

We express (6.44) by expanding the terms
\begin{equation}
\sum_{k=1}^{K} \mathcal{R}_k^2 \left[ S_k^{-1} - S_{k,su}^{-1} + (2\omega)^{-1} \right] \geq 0.
\end{equation}
(6.45)
\begin{equation}
\frac{\sum_{k=1}^{K} \mathcal{R}_k^2 \left[ S_k^{-1} - S_{k,su}^{-1} + (2\omega)^{-1} \right]}{\left[ \sum_{k=1}^{K} \mathcal{R}_k \right]^2} \geq (2\omega)^{-1}.
\end{equation}
(6.46)

We note that the term $S_k^{-1} - S_{k,su}^{-1} + (2\omega)^{-1} \geq 0$ is positive because $S_k \leq S_{k,su}$ and because $\omega \geq 0$. In order to belong to the robust multiaccess slope region, the set
of spectral efficiency slopes $S_k$ must satisfy (6.46) for all direction vectors $R$, and therefore (6.46) must also hold for the particular direction vector $R$, which minimizes the left hand side of (6.46). According to the Lemma 5, this occurs when $R_k^{-1}$ is proportional to $S_k^{-1} - S_{k, su}^{-1} + (2\omega)^{-1}$, and the minimum is given by the right hand side of the Lemma 5.

To utilize Lemma 5, we put $a_k^{-1} = S_k^{-1} - S_{k, su}^{-1} + (2\omega)^{-1}$, and $b_k = R_k$ inside (6.38). Evaluating minimum over all $R_k$ is given on the right hand side of (6.38), because $S_k$ must belong to the robust multiaccess slope region. It follows that

$$\left[ \sum_{k=1}^{K} \left[ S_k^{-1} - S_{k, su}^{-1} + (2\omega)^{-1} \right]^{-1} \right]^{-1} \geq (2\omega)^{-1}$$

$$\Rightarrow \sum_{k=1}^{K} \left[ S_k^{-1} - S_{k, su}^{-1} + (2\omega)^{-1} \right]^{-1} \geq (2\omega)$$

(6.47)

(6.48)

The last inequality can be parameterized by breaking $\omega$ into parts as $\sum_{k=1}^{K} \omega_k = \omega$. Here, each $\omega_k$ will satisfy the following inequality

$$S_k^{-1} - S_{k, su}^{-1} + (2\omega)^{-1} \geq (2\omega_k)^{-1}.$$  

(6.49)

This finishes the proof of Corollary 6.

Proof of Proposition 3. First see show the side which states that $(\text{tr } \Phi)^2 \leq M \text{ tr } (\Phi^2)$.

Supposing that $\Phi$ is diagonal, we need to prove that

$$\left[ \sum_{m=1}^{M} \Phi(m, m) \right]^2 \leq M \sum_{m=1}^{M} [\Phi(m, m)]^2.$$  

(6.50)

Here, (6.50) represents inequality between arithmetic and quadratic means of $\Phi(m, m)$, which is in fact a special case of Lemma 5. To see this put $a_k = 1$, also $K = M$, and $b_k = \Phi(m, m)$ in Lemma 5.

This completes the proof for the case of diagonal $\Phi$. Because $\Phi$ must be hermitian, when we add off-diagonal elements to $\Phi$, then tr $(\Phi^2)$ only grows because it is equal
to the right hand side of (6.50) and plus extra terms, which come from off-diagonal elements. While \( \text{tr} (\Phi^2) \) only grows by adding off-diagonal elements, the left hand side \( (\text{tr} \Phi)^2 \) stays the left hand side of (6.50), which means that \( (\text{tr} \Phi)^2 \leq M \text{tr} (\Phi^2) \) only becomes stronger. Finally, equality holds only if \( \Phi \) is exactly diagonal and if and only if all \( \Phi(m,m) \) are equal.

Now we show the other side of the inequality, which states that \( (\text{tr} \Phi)^2 \geq \text{tr} (\Phi^2) \), and to do so, we let \( \Phi = QAQ^\dagger \) be eigenvalue decomposition of \( \Phi \). We have that \( \text{tr} \Phi = \text{tr} \Lambda \), and that \( \text{tr} (\Phi^2) = \text{tr} (QAQ^\dagger QAQ^\dagger) = \text{tr} (\Lambda^2) \). All we need to show is that \( (\text{tr} \Lambda)^2 \geq \text{tr} (\Lambda^2) \), for the diagonal matrix of eigenvalues. The last inequality is obvious, because \( (\text{tr} \Lambda)^2 \) is square of a sum, while \( \text{tr} (\Lambda^2) \) is sum of squares. To conclude, we observe that equality holds in \( (\text{tr} \Phi)^2 \geq \text{tr} (\Phi^2) \) if and only if \( \Phi \) has only one non-zero eigenvalue. \( \square \)
Chapter 7

Comparison of Multiple Access Strategies

In this chapter, we utilize the mathematical toolbox which we have developed in previous chapters, in order to compare various multiple access strategies and receiver structures. This comparison will mostly be done in terms of performance, for the regime of low spectral efficiencies.

7.1 Tightly Coordinated Users: TDMA vs. SUD(\(\mathcal{R}\))

It has already been established in [2, 3, 16, 17], and others, that in tightly coordinated multiple access systems, multiuser detection provides far superior performance, as compared to single-user detection, or time division multiple access. One simple reason is that the achievable rate region of single user detectors, and the rate region of time division, is dominated by the rate region of multiuser detectors.

The same conclusion holds for "slope regions," namely

\[
\text{SUD}(\mathcal{R}) \subseteq \text{MUD}(\mathcal{R})
\]

\[
\text{TDMA} \subseteq \text{MUD}(\mathcal{R}).
\]

(7.1)

(7.2)

For this reason, here we focus on comparison between TDMA and SUD(\(\mathcal{R}\)). Given single-antenna channels with no fading, TDMA achieves superior spectral efficiencies when compared to the single-user detectors approach [17]. Here, we show how opposite occurs in certain multi-antenna channels, with low power transmission, where
the set of spectral efficiency slopes achieved by TDMA is contained in the set of spectral efficiency slopes, which are achieved by SUD(\mathcal{R}).

Before demonstrating superiority of SUD(\mathcal{R}) by inclusion, we provide a visual comparison of SUD(\mathcal{R}) and TDMA in Figure 7.1, which is relevant for two users, with two Rayleigh fading receive antennas. In Figure 7.1, we observe how each point from TDMA slope region, becomes dominated by a point from SUD(\mathcal{R}), for some \mathcal{R}. We generalize the intuition from Figure 7.1 in the following Theorem: it provides a sufficient condition for TDMA \subseteq SUD(\mathcal{R}), for some \mathcal{R}.

**Theorem 8.** Whenever \(2\mathcal{E} - \Sigma_{su}^{-1} \geq 0\), for every \(\Sigma_{TDMA} \in TDMA\), there exists \(\mathcal{R}\) so that slopes \(\Sigma_{SUD(\mathcal{R})}\) are superior to TDMA, namely \(\Sigma_{TDMA} \leq \Sigma_{SUD(\mathcal{R})}\). Hence, if the base-station is allowed to control users data rates, all users can achieve superior
spectral efficiency slopes, than with TDMA.

Proof. See Appendix. \qed

The sufficient condition of Theorem 8 is relatively easy to satisfy. Namely, for $2\mathcal{C} - \Sigma_{su}^{-1} \geq 0$ to hold, all that it takes is two Rayleigh faded receive antennas, with $\kappa = 2$. In this case, the Wideband matrix $\mathcal{C}$ from (5.8) has $3/4$ as diagonal elements, and $1/4$ as off-diagonals, therefore

$$2\mathcal{C} - \Sigma_{su}^{-1} = \frac{1}{4} 1 + \frac{1}{2} 11^T \geq 0,$$

where $1$ is the column vector with all-ones. Sufficient condition (7.3) holds true, because the matrix in (7.3) is a rank-one positive perturbation of the scaled identity matrix. Namely, eigenvectors of (7.3) are all elementary vectors, with only one non-zero entry, and all their corresponding eigenvalues in (7.3) are all positive. We now shift our focus towards coarsely coordinated multiaccess systems.

7.2 Coarsely Coordinated Users: TDMA vs. MUD$_*$

Here we demonstrate how policies, represented with the spectral efficiency slope, from MUD$_*$, are always superior to policies, chosen from TDMA. Again, we first provide the intuition, because we established the fact that MUD$_*$ represents a virtual partition of receiver space dimension: the claim is that TDMA fails to exploit all spatial receiver dimensions.

We start with a heuristic calculation in reference to the two-user system from Fig. 6.5 - 6.7, but assuming that coarse coordination is implemented with time-division, and each user is given equal time slot. Every user has 2 transmit antennas, and receiver is made up of 4 receive elements; therefore, given that either user transmits, its transmission policy is awarded approximately with $\min(2,4) = 2$ [space] dimensions.
Notice that this calculation would be exact in high-power regime, where the capacity grows linearly with the minimum of transmit and receive antennas, but it is only approximate for low power systems, where the effective number of spatial channel dimensions is the harmonic average of transmit and receive antennas, see [2] and also (5.9). Approximately two receive antenna dimensions are clearly wasted because of the fact that min(2,4) = 2. Furthermore, each user transmits only half of the [time], and therefore the aggregate complex dimension of [space] \times [time] \times [frequency], awarded to each user equals [time] \times [frequency], with TDMA as a multiple access strategy.

On the other hand, if policies are divided by partitioning receive antennas virtually, assuming superposition of two users, exactly like in Fig. 6.5 - 6.7, with MUD\_s, then policy of each user is in effect provided with min(2,2) = 2 [space] dimensions, and the receiver spatial dimension 4 is fully exploited. Because of superposition, no [time] or [frequency] dimensions are divided between users, and therefore each policy is awarded with an aggregate complex dimension which equals 2 \times [time] \times [frequency], in effect twice more than with TDMA. Notice that this calculation is based on the robust slope region, and hence receive dimensions are awarded irrespective of other users transmission rates. This intuition is rigorously justified by the inclusion

\[
\text{TDMA} = \left\{ \Sigma : \text{tr} \left( \Sigma^{1/2} \mathcal{C} \Sigma^{1/2} \right) \leq 1 \right\} \quad (7.4)
\]

\[
\subseteq \left\{ \Sigma : \lambda_{\max} \left( \Sigma^{1/2} \mathcal{C} \Sigma^{1/2} \right) \leq 1 \right\} = \text{MUD}_s \quad (7.5)
\]

Here, (7.5) holds because trace equals the sum of eigenvalues, and all eigenvalues of \mathcal{C}, and hence eigenvalues of \Sigma^{1/2} \mathcal{C} \Sigma^{1/2} are non-negative. If trace of \Sigma^{1/2} \mathcal{C} \Sigma^{1/2} is less than 1, so must be \lambda_{\max}; therefore \Sigma which belongs inside (7.4) must also belong inside (7.5).
Necessary and sufficient condition, for equality of TDMA and MUD, is that the matrix $\Sigma^{1/2}C\Sigma^{1/2}$ is rank one and in this case, the Wideband Matrix $C$ has only one eigenvalue. One example is single antenna systems with no spreading or fading to separate users statistically, and one may observe this from (4.5), by putting $\kappa(h) = 1$, which is valid only for a deterministic channel.

### 7.3 Appendix: Chapter Proofs

*Theorem 8.* We first require the following Lemma.

**Lemma 6.** Union over all slope regions $\text{SUD}(\mathcal{R})$, which are achieved with a bank of single-user detectors, is given as follows

$$\text{SUD}^* = \bigcup_{\mathcal{R}} \text{SUD}(\mathcal{R}) = \left\{ \Sigma : \lambda_{\max}(\Sigma^{1/2} \left[ 2C - \Sigma_{\text{su}}^{-1} \right] \Sigma^{1/2}) \leq 1 \right\}. \quad (7.6)$$

Once we prove this lemma, we will show that TDMA is contained inside (7.6).

*Proof of Lemma 6.* As usual, there are two parts to the proof, first we show that every $\text{SUD}(\mathcal{R})$ is contained inside the right hand side of (7.6).

Let $\Sigma$ be achievable for some $\mathcal{R}$ with single user detectors, where operating point is on the boundary of the slope region $\text{SUD}(\mathcal{R})$ in (4.14), and $\geq$ becomes exact equality. From here, we conclude that $\Sigma^{-1/2}\mathcal{R}$ is an eigenvector of $\Sigma^{1/2} \left[ 2C - \Sigma_{\text{su}}^{-1} \right] \Sigma^{1/2}$, with eigenvalue of one - this statement is equivalent to $\text{SUD}(\mathcal{R})$ in (4.14). Because $\Sigma^{1/2} \left[ 2C - \Sigma_{\text{su}}^{-1} \right] \Sigma^{1/2}$ has all non-negative entries, and $\Sigma^{-1/2}\mathcal{R}$ is its eigenvector with all positive entries, then $\Sigma^{-1/2}\mathcal{R}$ must be the principal eigenvector of $\Sigma^{1/2} \left[ 2C - \Sigma_{\text{su}}^{-1} \right] \Sigma^{1/2}$. So, if $\Sigma$ is achieved by some $\mathcal{R}$, then the principal eigenvalue of $\Sigma^{1/2} \left[ 2C - \Sigma_{\text{su}}^{-1} \right] \Sigma^{1/2}$ cannot exceed one. This finishes the $\subseteq$ part of the Lemma 6.

Now we show $\supseteq$ part, which claims that any $\Sigma$ from the right hand side of (7.6) is contained in some $\text{SUD}(\mathcal{R})$. 

Given $\Sigma$ from the right hand side of (7.6), we must find a vector of data rates $\mathcal{R}$, for which $\Sigma$ is achieved by $\text{SUD}(\mathcal{R})$. Let (7.6) hold with equality for $\Sigma$, and let $z$ be the eigenvector which has the eigenvalue equal to unity. We claim that we have found $\mathcal{R}$ and that it satisfies $z = \Sigma^{-1/2}\mathcal{R}$.

To see this we write the eigen-equation with $z$ as the eigenvector of unit eigenvalue

$$\Sigma^{1/2} \left[ 2\mathbf{c} - \Sigma_{\text{su}}^{-1} \right] \Sigma^{1/2} \mathcal{R} = \Sigma^{-1/2}\mathcal{R} \quad (7.7)$$

$$\left[ 2\mathbf{c} - \Sigma_{\text{su}}^{-1} \right] \mathcal{R} = \Sigma^{-1}\mathcal{R}. \quad (7.8)$$

From (7.8), and the generic expression for $\text{SUD}(\mathcal{R})$, we conclude that $\Sigma$ is achievable for this particular $\mathcal{R}$. Namely, if $\Sigma$ is inside the right hand side of (7.6), then we found $\mathcal{R}$ for which it is also in $\text{SUD}(\mathcal{R})$. □

With Lemma 6, we are ready to finish the proof of this theorem.

Fact one: if $2\mathbf{c} - \Sigma_{\text{su}}^{-1} \geq 0$, then $\Sigma^{1/2} \left[ 2\mathbf{c} - \Sigma_{\text{su}}^{-1} \right] \Sigma^{1/2} \geq 0$ and all its eigenvalues are non-negative. Fact two: diagonal elements of $\Sigma^{1/2} \left[ 2\mathbf{c} - \Sigma_{\text{su}}^{-1} \right] \Sigma^{1/2}$ are $\mathcal{S}_k\mathcal{S}_{k,\text{su}}^{-1}$.

The following therefore holds

$$\text{TDMA} = \left\{ \Sigma : \sum_k \lambda_k \left( \Sigma^{1/2} \left[ 2\mathbf{c} - \Sigma_{\text{su}}^{-1} \right] \Sigma^{1/2} \right) \leq 1 \right\} \quad (7.9)$$

$$\subseteq \left\{ \Sigma : \lambda_{\text{max}} \left( \Sigma^{1/2} \left[ 2\mathbf{c} - \Sigma_{\text{su}}^{-1} \right] \Sigma^{1/2} \right) \leq 1 \right\} = \text{SUD}^\ast. \quad (7.10)$$

Here, (7.9) holds because of the fact two, and because sum of diagonal elements equals the sum of eigenvalues, namely $\sum_k \lambda_k \left( \Sigma^{1/2} \left[ 2\mathbf{c} - \Sigma_{\text{su}}^{-1} \right] \Sigma^{1/2} \right) = \sum_k \mathcal{S}_k\mathcal{S}_{k,\text{su}}^{-1}$.

Also, (7.10) holds because of the fact one, namely because $\Sigma^{1/2} \left[ 2\mathbf{c} - \Sigma_{\text{su}}^{-1} \right] \Sigma^{1/2}$ must have all non-negative eigenvalues. □
Chapter 8

System Design

Finally, we provide several design guidelines for practical low power systems. First, we address the issue of receiver antenna positioning and packing inside a limited space. Second, we revisit the classic problem of spreading code design, but now from the perspective of coarse user coordination. Third, we demonstrate how to solve for the minimum bandwidth requirements of a practical system, when transmit powers and data rates are fixed.

8.1 Receiver Antenna Packing: Fundamental Limits for the Effective Degrees of Freedom

Based on Corollary (6) quantity $\omega$ which depend only on physical channel properties, represents the effective number of degrees of freedom for both for single and multiuser systems, and we already established the unit of $\omega$ as [Uncorrelated Receivers]. In any case, tightly or coarsely managed, receivers ought to be designed so that $\omega$ is as large as possible, in order to provide the reception with the maximum amount of dimensionality.

An important problem of $M$ receive antenna packing, inside a finite space or geometry $G$, may be addressed using $\omega$ both as a tool and the target metric to be maximized. Specifically, suppose we are given number of antennas $M$ and the space $G$ that they must be placed. Space $G$ could be a persons laptop computer,
belt or anything else that the imaginative marketing branch suggests. The optimum positioning of M antennas in this space maximizes $\omega$, and solutions are quite intuitive given enough symmetry.

For example, circular geometry, such as belt, suggests uniform antenna distribution, and we are able to prove this (below), in the limit as the number of receive antennas grows. On the other hand, positioning antennas on a square laptop shouldn’t be uniform, because less correlated edges ought to contain more antenna elements than the interior. A simple computer search may be performed to maximize $\omega$ over a grid, for a relatively small M. Finally, placing antennas in a finite size linear array ought not be uniform either, and in fact computer simulations based on maximizing $\omega$ tend to cluster antennas in groups.

As a reminder, for correlated multi-antenna systems, entry $\Theta(n', n)$ is a function of transmit antenna separation in wavelengths, as $\Theta(n', n) = J_0 (2\pi ||r_{n'} - r_n||/\lambda)$. Here, $r_n$ is the three-dimensional position vector of $n$-th transmit antenna, $\lambda$ is the carrier wavelength, and $J_0$ is the Bessel function of zeroth order. Reference origin for the position vector $r_n$ of $n$-th antenna is irrelevant because separation between
antennas \( n \) and \( n' \) is given as \( ||r_{n'} - r_n|| \), which only depends on the relative antenna spacing.

Receive correlation \( \Phi \) is usually modeled analogous to \( \Theta \). However, we note that authors which consider dense multiantenna receivers [37, 39, 40] also adopt the passive* channel model, which is to set \( \Phi(m', m) = M^{-1} J_0(2\pi ||r_{m'} - r_m||/\lambda) \), to keep the average received power normalized to unity. By considering a generic \( \Phi \), we will provide results for both conventions.

### 8.2 Limiting Degrees of Freedom

Dense multi-antenna transceivers contain an abundance of antenna elements which are tightly packed inside a given aperture \( G \subseteq \mathbb{R}^3 \). In order to evaluate how close to optimal is any given antenna configuration, we introduce the following notion of \textit{limiting degrees of freedom}, which we define as the maximum achievable value for the effective degrees of freedom.

**Definition 6.** \textbf{Limiting degrees of freedom} characteristic of a given aperture \( G \subseteq \mathbb{R}^3 \) is defined as follows

\[
\omega(G) \triangleq \sup_{|A| < \infty} \frac{(\text{tr} \Phi_A)^2}{\text{tr} (\Phi_A^2)}. \tag{8.1}
\]

Supremum (8.1) is taken over all possible configurations of finite number \( |A| < \infty \) of antennas inside \( G \), namely \( A \subseteq G \). In (8.1), \( \Phi_A \) denotes the correlation between antennas which are placed inside the finite set of discrete points \( A \).

*In fact [40] considers “normalized capacity” which is mathematically equivalent to the passive channel model.
Definition of limiting degrees of freedom (8.1) is applicable to both receive and transmit antenna arrays, where for the transmit antenna array $\Phi_A$ is replaced by $\Theta_A$ in (8.1). For any antenna configuration $A$, correlation matrix $\Phi_A$ depends only on the ratio of antenna spacing to $\lambda$, which means that limiting degrees of freedom $\omega(G)$ depend only on size and shape of $G$ with respect to the carrier wavelength $\lambda$.

8.2.1 Limiting Degrees of Freedom: Properties

Given an arbitrary transceiver aperture $G$, it is feasible to either compute $\omega(G)$ in a closed form, or alternatively, to find $\omega(G)$ numerically, by performing a search over antenna configurations. These two approaches are covered in Sections 8.2.2 and 8.3 below. By showing that $\omega(G)$ is finite, in this section we ensure that limiting degrees of freedom can be closely achieved with a finite and hence realistic antenna configuration. First, we elaborate on basic Properties (1-3) of limiting degrees of freedom, which are then used to show the finitness of $\omega(G)$ in Property 4.

**Property 1.** Limiting degrees of freedom characteristic $\omega(G)$ is a monotonically non-decreasing set function. Specifically, we mean that $G_1 \subseteq G_2$ necessitates that $\omega(G_1) \leq \omega(G_2)$.

**Property 2.** For a non-empty set $G \subseteq \mathbb{R}^3$, we have $\omega(G) \geq 1$.

Proofs of Properties 1 and 2 are simple observations. Property 1 is true because supremum (8.1) over a superset $G_2$ cannot be smaller. Property 2 is true because positioning a single antenna inside $G$ results in the effective degrees of freedom of one. Less obvious but quite intuitive Property of limiting degrees of freedom is finite subadditivity. Proof is given in the Appendix.
Property 3. Limiting degrees of freedom characteristic is finitely sub-additive, namely

$$\omega (G_1 \cup G_2) \leq \omega(G_1) + \omega(G_2). \quad (8.2)$$

The basic utility of Property 2 is twofold. First, we will use it as a building block to prove finiteness of $\omega$ for any bounded aperture $G$. Second, when apertures $G_1$ and $G_2$ are very far apart, so that each point of $G_1$ is practically uncorrelated with each point of $G_2$, then (8.2) is becomes equality. In that case, we simply evaluate the individual limiting degrees of freedom for $G_1$ and $G_2$ separately to obtain $\omega(G_1 \cup G_2)$.

Property 4. For bounded set $G$, the limiting degrees of freedom characteristic is finite, namely $\omega(G) < \infty$.

The basic claim of Property 4 is that no antenna packing inside bounded space $G$ will yield infinite effective degrees of freedom $\omega$. Because $\omega$ exhibits this saturation effect, so will the spectral efficiency. In essence, even if we had an infinite number of antennas which are packed inside $G$, the spectral efficiency will be finite, which is a fundamental property that distinguishes between spatially dense and sparse multi-antenna systems. We provide a practical example of this phenomenon, where we evaluate the limiting degrees of freedom for a circle.

8.2.2 Limiting Degrees of Freedom of Circular Array

When computing limiting degrees of freedom for a circle, we are faced with a very peculiar phenomenon that uniform circular array is not necessarily optimum. For example, consider placing $M = 2$ antennas on a circle whose radius satisfies $J_0 \left( 2\pi \left[ \frac{2\pi}{\lambda} \right] \right) \neq 0$ and $2\pi \left[ \frac{2\pi}{\lambda} \right] \geq 2.404$, where 2.404 is the first zero of the Bessel function. Here, positioning two antennas diametrically opposite is strictly suboptimum, because they will be correlated. Superior configuration is the one for which distance
between antenna locations corresponds to the first zero of $J_0$, which produces an array of two perfectly uncorrelated antennas.

This phenomenon is also reflected in Figure 8.2, which depicts effective degrees of freedom $\omega$ versus the number of antennas that are uniformly located on the circle. Curve labeled as $r = \lambda$ in Figure 8.2 is not monotonically increasing, which means that increasing number of uniform antennas on the circle can actually reduce $\omega$. In particular, by migrating from the uniform circular array of $M = 8$ antennas to the uniform circular array of $M = 10$ antennas, $\omega$ reduces from $\omega = 7$ to $\omega = 6$. When these antennas transmit independent streams of data, the spectral efficiency actually reduces by increasing the number of antennas! The same conclusion holds for the
receive mode, when we adopt the passive channel model where aggregate receive power is kept constant because of dense array packing.

In Theorem 9, we show that uniform antenna configuration on a circle becomes optimum only in the limit as $M \rightarrow \infty$. Limiting degrees of freedom characteristic for a circle is also given in Theorem 3.

**Theorem 9.** If $G$ is a circle of radius $r$, and if the carrier wavelength is $\lambda$, then

\[
\omega^{-1}(G) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \left| J_0 \left( \frac{2\pi r}{\lambda} \sqrt{2-2\cos \varphi} \right) \right|^2 d\varphi.
\] (8.3)

For circles of various $r/\lambda$, the Figure 8.2 illustrates convergence of $\omega$ to $\omega(G)$ when antennas are uniformly positioned on the circle boundary. For example, given a circle with $r = \lambda$, limiting degrees of freedom characteristic is close to 11, which is closely achieved by uniformly positioning 14 antennas along the boundary.

One practical implication of Figure 8.2 is that uniformly positioning 14 or more antennas on the circle of radius $r = \lambda$ is fundamentally a near-optimum configuration. Positioning more than 14 antennas on the circle boundary will produce no effect on the spectral efficiency slope $\mathcal{S}$. If antennas are receive antennas, and if we do not adopt the passive channel model from [37,39,40] then more antennas will reduce $\mathcal{E}_{\text{min}}$. In this case, we still conclude that for a fixed number of antennas beyond 14, uniform configuration is very close to optimum. Figure 8.2 is interpreted in a similar fashion for other values of $r/\lambda$. For example, 4 or more uniformly placed antennas is a near-optimum configuration for $r = 0.25\lambda$ and 8 or more uniformly positioned antennas is a near-optimum configuration for $r = 0.5\lambda$. 
8.2.3 Small-Sized Circular Arrays

Small-sized antenna arrays whose radius is on the order of few centimeters are gaining importance because of their portability and practicality for mobile applications. To address these, we next simplify the formula (8.3) under the assumption that \( r/\lambda \) is small. The following Corollary to Theorem 9 is based on the expansion of Bessel function \( J_0 \) for the small values of the argument. Proof is given in the Appendix.

**Corollary 7.** [to Theorem 9] For small-sized circular arrays where \( r/\lambda \to 0 \), we have

\[
\omega(G) = 1 + 4\pi^2 \left[ \frac{r}{\lambda} \right]^2 + o \left( \left[ \frac{r}{\lambda} \right]^2 \right). \tag{8.4}
\]

Numerical results indicate that the approximation based on Corollary 7 remains valid for up to \( r = 0.25\lambda \), which may be verified from the corresponding curve in Figure 8.2. It is commonly understood that in order to obtain two uncorrelated spatial degrees of freedom, the separation between two antennas ought to be approximately \( 2r = 0.5\lambda \) which provides fully uncorrelated array. However, based on the approximation from Corollary 7 it suffices to use a dense multi-antenna transceiver of size \( 4\pi^2 \left[ r/\lambda \right]^2 = 1 \), in order for \( \omega(G) = 2 \). Namely, a dense multi-antenna transceiver which satisfies \( r < \lambda/6 \) will provide two effective degrees of freedom. This translates to more than 50% savings in the array diameter which enhances the array portability for mobile applications.

8.2.4 Approximation for Moderate-Sized Circular Arrays

Another important study is provided by considering (8.3), but now for *large and moderate* values of the circle radius \( r \). In Figure 8.3, we plot the effective degrees of freedom as a function of circle radius (8.3), which remains fairly linear up to \( r = 100\lambda \). Figure 8.3 also depicts a heuristic linear approximation \( \omega(G) \approx 2\pi \left[ \frac{r}{\lambda} \right] \),
which closely approximates the effective degrees of freedom. A simple mnemonic for heuristic linear approximation from Figure 8.3 is that limiting degrees of freedom for a circle approximately equal the circumference in wavelengths. Explanation of this phenomenon is an open question, but consistent with a high-power study in [8].

Taking the signal-space approach, [8] considers the antenna array response to be the communication bottleneck, and shows that its basis contains approximately $4\pi \left[\frac{r}{\lambda}\right]$ functions, for large circle radius $r \gg \lambda$. This implies that for very high signal to noise ratios, the multi-antenna channel capacity is approximately $4\pi \left[\frac{r}{\lambda}\right]$ greater than the capacity of single-input-single-output (SISO) channels. The approximation $\omega(G) \approx 2\pi \left[\frac{r}{\lambda}\right]$ from Figure 8.3 hints at a similar conclusion for the low power regime. Namely, the spectral efficiency of single-antenna Rayleigh fading channels is

\textbf{Figure 8.3:} Limiting Degrees of Freedom and the Linear Approximation.
approximated with (3.9) and \( S = 2\kappa^{-1}(h) = 1 \). Given a communication system with large circular receiver aperture \( G \), and with large effective degrees of freedom of the transmitter \( \omega_T \to \infty \), the spectral efficiency is approximated using \( S = 2\omega_R(G) \approx 4\pi \left[ \frac{r}{\lambda} \right] \). Therefore, the spectral efficiency which is achievable using circular \( G \) of radius \( r \) is approximately \( 4\pi \left[ \frac{r}{\lambda} \right] \) multiple of spectral efficiency of SISO Rayleigh faded channels. However, we emphasize that the high-power analysis in [8], which is based on dimensionality of large signal-spaces, fundamentally differs from the low-power analysis herein, which is based on statistical multi-antenna channel, and that formula (8.3) is exact, whereas \( 4\pi \left[ \frac{r}{\lambda} \right] \) is approximate. Having presented analytical results for the circle packing, we now shift the attention towards geometries with less symmetry, where we present numerical results.

### 8.3 Antenna Array Design Examples

#### 8.3.1 Numerical Procedures

This section focuses on simple numerical procedures for maximizing the effective degrees of freedom \( \omega \), by searching over various antenna configurations. Here, we consider two methods: the exhaustive search, which is optimal but computationally prohibitive, and the greedy search, which is suboptimal but with linear complexity.

First, we consider the option of exhaustive search among all antenna configurations on the coordinate grid which covers \( G \). Even for small and fixed number of antennas, numerical complexity of the exhaustive search is exponential, because \( L \) hypothesis locations for \( M \) antennas results in \( \binom{N}{L} \approx N^L \) computations of \( \omega \). For example, a relatively coarse 2–dimensional search over a \( 10 \times 10 \) grid with \( M = 10 \) antennas results in approximately \( N^L = 100^{10} \) computations of \( \omega \), which is too in-
tensive for current generation computers.

Greedy search is a suboptimal but simple alternative to the exhaustive search. The main idea behind the greedy search is to start with an arbitrary antenna configuration, and successively increase $\omega$ by re-positioning a single antenna at a time. Pseudo-code for one iteration step of the greedy search is given below.

For antenna $m = \{1, .., M\}$ do {

- Remove the $m-$th antenna from the current configuration.

- Fix the remaining $(M - 1)$ antennas.

- Consider all $L - (M - 1)$ empty locations for the $m-$th antenna. Select the one which results in the overall configuration with maximum $\omega$.
}

It usually takes several iteration steps for the greedy configuration search to converge to a local maximum. Each iteration step of the greedy search consists of $M(L - M + 1) < ML$ evaluations of $\omega$; hence, the aggregate complexity is approximately $MLQ$ evaluations of $\omega$, where $Q$ is the number of iteration steps.

As an important observation, we point out that the greedy search for antenna configurations is suboptimum. The objective function $\omega$ has multiple extreme points, because all rotations of a certain antenna configuration result in the same effective antenna-space. Since the objective function has multiple extreme points, it must be non-convex, and the greedy search algorithm may converge to a local maximum, depending on the initial antenna configuration. We therefore recommend running several trials with different initial configurations. Greedy search is tested on linear and square arrays below.
**Figure 8.4**: Effective Degrees of Freedom vs. Number of Antennas in a Uniform Linear Array of Length $d$

### 8.3.2 Linear Arrays

Uniform linear arrays are the common benchmark case study for antenna configurations which are positioned on a straight line of length $d$. Figure 8.4 illustrates the effective degrees of freedom for uniform linear arrays of various size and number of antennas. As predicted by Property 4, $\omega$ remains bounded, regardless of how many antennas are packed in a fixed-size array. We make several observations based on Figure 8.4.

Unlike for circular arrays, effective degrees of freedom of uniform linear arrays attain a peak at a finite number of antennas, which depends on the array size.
<table>
<thead>
<tr>
<th>Size</th>
<th>Num. of Ant.</th>
<th>ω for Uniform</th>
<th>ω for Non-Uniform</th>
<th>Spec. Eff. Savings</th>
</tr>
</thead>
<tbody>
<tr>
<td>$d = 4\lambda$</td>
<td>6</td>
<td>5.10</td>
<td>5.72 [Figure 8.5]</td>
<td>12%</td>
</tr>
<tr>
<td>$d = 4\lambda$</td>
<td>7</td>
<td>5.20</td>
<td>6.60 [Figure 8.6]</td>
<td>27%</td>
</tr>
<tr>
<td>$d = 4\lambda$</td>
<td>8</td>
<td>5.30</td>
<td>7.26 [Figure 8.7]</td>
<td>37%</td>
</tr>
<tr>
<td>$d = 4\lambda$</td>
<td>9</td>
<td>6.80</td>
<td>7.93 [Figure 8.8]</td>
<td>17%</td>
</tr>
</tbody>
</table>

**Table 8.1:** Comparison of Uniform and Non-Uniform Linear Arrays.

<table>
<thead>
<tr>
<th>Num. of Ant.</th>
<th>Locations</th>
</tr>
</thead>
<tbody>
<tr>
<td>6</td>
<td>[0.06 0.16 0.40 0.64 0.88 0.98] × 4\lambda</td>
</tr>
<tr>
<td>7</td>
<td>[0.02 0.13 0.24 0.49 0.73 0.84 0.95] × 4\lambda</td>
</tr>
<tr>
<td>8</td>
<td>[0.05 0.16 0.27 0.38 0.63 0.74 0.85 0.96] × 4\lambda</td>
</tr>
<tr>
<td>9</td>
<td>[0.04 0.15 0.27 0.38 0.50 0.62 0.73 0.85 0.96] × 4\lambda</td>
</tr>
</tbody>
</table>

**Table 8.2:** Non-Uniform Antenna Locations in a Linear Array of Size $d = 4\lambda$.

For small-sized arrays where $d \leq 2\lambda$, the peak occurs at mutual antenna spacing of approximately $\lambda/4$, which is therefore recommended spacing for antennas which transmit/receive independent streams of data. For $d \leq 2\lambda$, effective degrees of freedom remain close to the upper bound given by uncorrelated arrays ($d = \infty$), before the peak occurs. Regardless of number of antennas, we report that greedy search algorithm converges to uniform configuration, given a small-sized array with $d \leq 2\lambda$.

The benefit of non-uniform configurations becomes apparent by considering moderate sized apertures, such as $d = 4\lambda$. Here, the number of effective degrees of freedom from Figure 8.4 grows very slow in the range from $M = 6$ to $M = 9$ antennas, which
Figure 8.5: Locations of $M = 6$ Antennas in a Linear Array of Size $d = 4\lambda$.

Figure 8.6: Locations of $M = 7$ Antennas in a Linear Array of Size $d = 4\lambda$.

Figure 8.7: Locations of $M = 8$ Antennas in a Linear Array of Size $d = 4\lambda$.

Figure 8.8: Locations of $M = 9$ Antennas in a Linear Array of Size $d = 4\lambda$.

indicates suboptimality of uniform arrays. For these values of $M$, the greedy search converges to non-uniform antenna configurations, which are given in Figures 8.5 - 8.8. The comparison of effective antenna-space for uniform and non-uniform configurations, along with savings in spectral efficiency are given in Table 8.1. The largest saving in spectral efficiency (37\%) is achieved for $M = 8$ antennas.

Finally, as a conjecture based on data from Figure 8.4, we observe that for both moderate ($d = 4\lambda$) and large ($d = 20\lambda$ and $d = 40\lambda$) uniform linear arrays, the effective number of degrees of freedom attains a peak at $M = 2 \left[ \frac{d}{\lambda} \right] + 2$. For this particular value of $M$, the antenna spacing is $\frac{A}{2+d-1}$. Note that conventional $\frac{1}{2}$ spacing leads to a uniform linear array of $M = 2 \left[ \frac{d}{\lambda} \right] + 1$ antennas, which can incur a significant degradation in spectral efficiency (Figure 8.4), when compared to $M = 2 \left[ \frac{d}{\lambda} \right] + 2$, which requires an additional antenna. As a practical conclusion, when constrained by the number of antennas, non-uniform arrays should be considered, such as the
ones we discovered by the greedy search procedure in Figures 8.5 - 8.8.

8.3.3 Square Arrays

Positioning antennas on a square aperture $G$ is another practical problem, which can be addressed using the greedy procedure for configuration search. Well-established benchmark configurations for square apertures are uniform square arrays, whose effective degrees of freedom we report in Figure 8.9. The greedy search for antenna configurations is performed for a few sample values of aperture size and number of antennas, as reported in Table 8.3. Again, significant spectral efficiency savings are possible by simply considering non-uniform arrays, with no additional cost in number of antennas.

<table>
<thead>
<tr>
<th>Size</th>
<th>Num. of Ant.</th>
<th>$\omega$ for Uniform</th>
<th>$\omega$ for Non-Uniform</th>
<th>Spec. Eff. Savings</th>
</tr>
</thead>
<tbody>
<tr>
<td>$d = \frac{\lambda}{2}$</td>
<td>4</td>
<td>3.05</td>
<td>3.76 [Figure 8.10]</td>
<td>23%</td>
</tr>
<tr>
<td>$d = \lambda$</td>
<td>9</td>
<td>5.40</td>
<td>6.78 [Figure 8.11]</td>
<td>25%</td>
</tr>
</tbody>
</table>

**Table 8.3: A Comparison of Uniform and Non-Uniform Square Arrays**

<table>
<thead>
<tr>
<th>Size</th>
<th>M</th>
<th>Locations $(x, y)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\frac{\lambda}{2}$</td>
<td>4</td>
<td>$\left[ (0.00, 0.00); (1.00, 1.00); (0.74, 0.26); (0.21, 0.79) \right] \times \frac{\lambda}{2}$</td>
</tr>
</tbody>
</table>
| $\lambda$ | 9 | $\left[ (0.00, 0.00)\; (0.00, 0.44)\; (0.00, 0.89)\; (0.33, 0.22)$
| | | $\left. \; (0.33, 0.66)\; (0.67, 1.00)\; (0.78, 0.11)\; (0.78, 0.56)\; (1.00, 0.89) \right] \times \lambda$ |

**Table 8.4: Non-Uniform Antenna Locations on a Square Array**
8.4 Bandwidth Requirements of A Low Power Multiple Access System

In this section, we demonstrate how to analyze bandwidth requirements for a low power multiuser system, when each users transmission rate $R_k$ [bits/sec], and transmit power $P_k$ [Watts] are fixed. Namely, we demonstrate how to solve for the minimum system bandwidth $B$ [Hz], which is required to support all $R_k$ and $P_k$.
Naturally, we have that

$$\mathcal{R}_k = R_k \mathcal{B} \quad \text{[bits]/[sec]}$$  \hspace{1cm} (8.5)

$$\mathcal{P}_k = P_k \mathcal{B} \quad \text{[Watts]}$$ \hspace{1cm} (8.6)

where $R_k$ is the spectrum efficiency in [bits]/[sec]/[Hz], whereas $P_k$ is the power spectrum density in [Watts]/[Hz], for user $k$.

The low power assumption is that all users operate close to their minimum values for energy per information bit, namely $\mathcal{E}_{\text{min},k}$. We initialize the discussion assuming that the multiple access strategy is frequency division, because it the simplest one to analyze, and it points out the main methodology for evaluating the bandwidth requirements of a low power system.

### 8.4.1 Frequency Division

Here, bandwidth requirement $\mathcal{B}_k^{(F)}$, for each user $k$, is evaluated, and all requirements are summed, because users are transmitting over non-overlapping frequency bands.
The problem is therefore reduced to a single-user problem, and the spectral efficiency slope $\mathcal{S}_{k,su}$, in a single user communication mode, will provide us with a bandwidth requirement for user $k$

First, energy per information bit $\mathcal{E}_k$ is found from powers $P_k$ and rates $R_k$, and second, the increment of $\mathcal{E}_k$ from $\mathcal{E}_{min}$ is evaluated as

$$\mathcal{E}_k = \frac{P_k}{R_k} = \frac{P_k}{R_k} \quad \text{[Joules]/[bit]}$$  \hspace{1cm} (8.7)$$

$$\delta \mathcal{E}_k = \log_2 \mathcal{E}_k - \log_2 \mathcal{E}_{k,\text{min}} \quad [3 \text{ dB}].$$  \hspace{1cm} (8.8)$$

Spectral efficiency is approximated using the increment (8.8) along with the spectral efficiency slope $\mathcal{S}_k = \mathcal{S}_{k,\text{su}}$, as $R_k = \mathcal{S}_{k,\text{su}} \delta \mathcal{E}_k$. Bandwidth requirement $\mathcal{B}_k^{(F)}$ for user $k$, will be evaluated directly from relationship between spectral efficiency and data.
rate \( B_k^{(F)} R_k = R_k \), namely

\[
R_k = S_{k, su} \delta \varepsilon_k = \frac{R_k}{B_k^{(F)}}
\]

\[
\Rightarrow B_k^{(F)} = \frac{R_k}{S_{k, su} \delta \varepsilon_k}.
\]

(8.10)

It is instructive to verify that dimensions of (8.10) are \([\text{symbols}] / [\text{sec}] = [\text{Hz}]\).

Finally, all \( B_k^{(F)} \) are added to produce

\[
B^{(F)} = \sum_{k=1}^{K} B_k^{(F)}
\]

(8.11)

\[
B^{(F)} = \sum_{k=1}^{K} \frac{R_k}{S_{k, su} \delta \varepsilon_k}.
\]

(8.12)

Transition (8.12) is provided by (8.10), and (8.12) represents the system aggregate bandwidth requirement, which is evaluated provided that users transmit over non-overlapping frequency bands.

### 8.4.2 Time Division

Naturally, time division will not produce significantly different results from frequency division, but calculation herein will demonstrate the important guiding principle of the “dominant constraint,” which will also be used in superposition.

Namely, for time division multiple access, users operate on the same frequency band \( B^{(T)} \). This band, in turn, needs to be large enough to support the bandwidth requirement for each user. If \( B_k^{(T)} \) is bandwidth requirement for user \( k \), then overall system bandwidth is given as

\[
B^{(T)} = \max_{1 \leq k \leq K} B_k^{(T)}.
\]

(8.13)

Even though guiding principle (8.13) conceptually differs from frequency division (8.11), it will result in the same expression for overall bandwidth requirement, provided that time sharing parameters \( \alpha_k \) are carefully selected.
Here, $\alpha_k$ denotes the percentage of time which is dedicated to user $k$ transmission, and during this period of time, the effective transmission rate of user $k$ is actually $R_k/\alpha_k$, which averages to $R_k$ over all time. Bandwidth requirement for user $k$ is therefore given by analogy with (8.10), and the overall bandwidth requirement is found from (8.13)

$$B_k^{(T)} = \frac{R_k/\alpha_k}{S_{k,su} \delta E_k}$$  \hspace{1cm} (8.14)

$$B^{(T)} = \max_{1 \leq k \leq K} \frac{R_k}{\alpha_k S_{k,su} \delta E_k}$$ \hspace{1cm} (8.15)

One expects $B^{(T)}$ to be equal to $B^{(F)}$, but this is not yet apparent from expressions (8.15) and (8.12).

Relationship between (8.15) and (8.12) is established once we note that the maximum (8.15) of $K$ real numbers, is not less than their convex combination (8.12). Equality holds if and only if all numbers are equal, namely

$$B^{(T)} \geq \sum_{k=1}^{K} \alpha_k \left( \frac{R_k}{\alpha_k S_{k,su} \delta E_k} \right)$$ \hspace{1cm} (8.16)

$$B^{(T)} \geq B^{(F)}.$$ \hspace{1cm} (8.17)

Transition (8.17) is provided by (8.12).

Equality in (8.17) holds only when all $B_k$ from (8.14) are the same, which means that $\alpha_k$ are selected so that all users have equal bandwidth requirements $B_k^{(T)}$. This is simply another manifest of duality between time and frequency.

### 8.4.3 Superposition

The principle of the dominant constraint will also be used to evaluate the bandwidth requirement of a low power multiuser system, provided that the receiver utilizes the optimum multiuser detection.
The following proposition deals with this scenario.

**Proposition 4.** For multiple access by superposition, the necessary bandwidth is

\[ B^{(M)} = \max_{A \subseteq \{1, 2, \ldots, K\}} \frac{R_A^T \mathcal{C}_A R_A}{R_A^T \delta \mathcal{E}_A}. \]  

(8.18)

**Proof.** Bandwidth requirements for this case are derived from the generic constraint of a multiaccess slope region, which is given in Theorem 2. For a set of small spectral efficiencies \( R_k \), this constrain may be written as

\[ R_A^T \Sigma_A^{-1} R_A - R_A^T \mathcal{C} R_A \geq 0. \]  

(8.19)

Here we note that \( \Sigma_A^{-1} R_A = \delta \mathcal{E}_A \), and that \( R_A B^{(M)} = R_A \).

The generic constraint from (8.19) becomes

\[ B^{(M)} (R_A^T \delta \mathcal{E}_A) - R_A^T \mathcal{C} R_A \geq 0. \]  

\[ \Rightarrow B^{(M)} \geq \frac{R_A^T \mathcal{C} R_A}{R_A^T \delta \mathcal{E}_A}. \]  

(8.21)

System bandwidth must be large enough so that all multiaccess constraints are satisfied, and therefore \( \max_A \) appears in (8.18). This represents the dominant constraint for superposition, and the equality is written in (8.18) since that bandwidth is sufficient to support all \( K \) users.

Format (8.18) is the most one can say. Even though it may intuitively appear that maximum in (8.18) is satisfied when \( A = \{1, \ldots, K\} \), this is simply not true. For example, consider equal rates, equal increments of energies per bit \( \delta \mathcal{E}_k \), and \( S_{1, su}^{-1} = 2 \), while \( S_{2, su}^{-1} = 1/2 \). By hand computation the maximum is satisfied for \( A = \{1\} \).

For the final comparison, (8.18) shouldn’t be greater than (8.15) or (8.12), but this is not apparent just from observing the expressions. Following calculation provides
rigorous justification

\[
\left( \sum_{k \in A} \frac{R_k}{S_{k, su} \delta E_k} \right) \left( \sum_{l \in A} R_l \delta E_l \right) = \sum_{k \in A} \frac{R_k^2}{S_{k, su}} + \sum_{\substack{k, l \in A \\ k < l}} R_k R_l \left( \frac{\delta E_l}{S_{k, su} \delta E_k} + \frac{\delta E_k}{S_{l, su} \delta E_l} \right) 
\]

\[
\geq \sum_{k \in A} \frac{R_k^2}{S_{k, su}} + 2 \sum_{\substack{k, l \in A \\ k < l}} R_k R_l \sqrt{S_{k, su}^{-1} S_{l, su}^{-1}} \quad (8.22) 
\]

\[
\geq \sum_{k \in A} \frac{R_k^2}{S_{k, su}} + 2 \sum_{\substack{k, l \in A \\ k < l}} R_k R_l \mathcal{C}_{k, l} \quad (8.24) 
\]

\[
= \mathcal{R}_A^T \mathcal{C}_A \mathcal{R}_A. \quad (8.25) 
\]

To clarify, (8.22) is by simple multiplication and grouping. Standard inequality between arithmetic and geometric means of two real numbers \( b_1 + b_2 \geq 2\sqrt{b_1 b_2} \) explains (8.23).

To explain (8.24) we note that Wideband Matrix \( \mathcal{C} \) is symmetric and positive definite, and (8.24) is in reference to its submatrix of size \( 2 \times 2 \). This submatrix consists of diagonal entries \( S_{k, su}^{-1} \) and \( S_{l, su}^{-1} \). Of diagonals are the same cross term \( \mathcal{C}_{k, l} \), found on the right of (8.24). Determinant of this four entry submatrix is nonnegative, which justifies transition (8.23) to (8.24). (8.25) is just (8.24) in the matrix notation.

Connecting (8.25) and the left side of (8.22) explains that

\[
\frac{\mathcal{R}_A^T \mathcal{C}_A \mathcal{R}_A}{R_A^T \delta E_A} \leq \sum_{k \in A} \frac{R_k}{S_{k, su} \delta E_k} \quad (8.26) 
\]

\[
\max_{A \subseteq \{1, \ldots, K\}} \frac{\mathcal{R}_A^T \mathcal{C}_A \mathcal{R}_A}{R_A^T \delta E_A} \leq \sum_{k=1}^{K} \frac{R_k}{S_{k, su} \delta E_k} \quad (8.27) 
\]

\[
\mathcal{B}^{(C)} \leq \mathcal{B}^{(F)}. \quad (8.28) 
\]

Here, (8.27) is simply (8.26) with max taken on both sides, which is still a valid
inequality. Furthermore, (8.28) is straightforward from definitions. First equality condition is derived from (8.24), stating that the Wideband Matrix must be rank one, and $S_{k, su} = S_{l, su}$. Second equality condition is derived from transition (8.22) to (8.23) essentially claiming that equality holds only if increments of energies per bit are equal.

Nevertheless, the morale of this section is to demonstrate how low power analysis captures first order bandwidth requirements for a multiple access system, and these requirements are provided in (8.18), (8.14) and (8.12) for various multiaccess strategies.

### 8.5 Optimal Spreading Sequences for Symmetric Policies

The set of WBE (Welch Bound Equality) sequences, having minimum total square correlation, was shown to be optimal in terms of maximizing the sum capacity of multiple access non-fading channels, if and only if all users are received with equal powers [31]. Given disparate received powers, some users must be fully orthogonal, while others are allocated generalized WBE sequences [24]. In either case, the optimum set of spreading sequences is conditioned on a particular set of received powers.

Rather than assuming a particular set of received powers, here we assume variable data rate and power transmissions, and consider the following question: what is the set of spreading sequences, which always guarantees each user equal payoff function of a policy, represented by the spectral efficiency slope $S$, irrespective of data rate or power of other users? Hence, we consider a system where the set of spreading sequences, of fixed length $L$, is not allowed to adjust in accordance to data rates or powers, and users adjust their data rates by choosing alternate channel codes.
We use the following channel model

\[ Y = \sum_{k=1}^{K} \mathcal{H}_k X_k s_k^T + \eta, \]  

(8.29)

where \( \mathcal{H}_k \) is the matrix of user \( k \) antenna coefficients, \( s_k^T \) is the user \( k \) row spreading sequence, and \( X_k \) is the matrix of user \( k \) symbols. Model (8.29) is still linear, but it is convenient to regard \( \mathcal{H}_k \) as a filter which operates along the row dimension - space, whereas \( s_k^T \) operates along the column dimension - time.

Furthermore, in channel model (8.29) we implicitly assumed that channel coherence period lasts for at least \( L \) chips, which is the length of every users spreading sequence. Transmit antenna correlation for each user is again denoted as \( \Theta_k \), and the receive antenna correlation is \( \Phi_k \).

First we find the \( \mathcal{C} \) matrix.

**Proposition 5.** The Wideband Matrix for the channel model (8.29) is given as

\[ \mathcal{C}_{k,l} = \frac{\text{tr} (\Phi_k \Phi_l)}{2 \text{tr} (\Phi_k) \text{tr} (\Phi_l)} |\rho(k,l)|^2, \quad \text{for} \quad k \neq l. \]  

(8.30)

\[ \mathcal{C}_{k,k} = \frac{\text{tr} (\Phi_k^2) \text{tr} [(F_k \Theta_k)^2]}{2 (\text{tr} \Phi_k)^2 [\text{tr} (F_k \Theta_k)]^2} (\kappa_k - 2) + \frac{\text{tr} (\Phi_k^2)}{2 (\text{tr} \Phi_k)^2} + \frac{\text{tr} [(F_k \Theta_k)^2]}{2 [\text{tr} (F_k \Theta_k)]^2}. \]  

(8.31)

Here, \( |\rho(k,l)|^2 \) represents the squared correlation of spreading sequences between users \( k \) and \( l \). Also, \( F_k^{\frac{1}{2}} \) is the pre-coding filter of user \( k \).

**Proof.** See Appendix.

Set of spreading sequences can be designed so that every user is allocated with an equal spectral efficiency slope policy, and we set \( \Sigma = S I \). Here, \( \Sigma \) must belong inside \( \text{MUD}_* \), and therefore (6.2) constrains the spectral efficiency slope \( S \), of each user \( k \), as follows

\[ S \leq \lambda_{\max}^{-1} (\mathcal{C}). \]  

(8.32)
Under following conditions the set of WBE sequences maximizes $S$.

**Theorem 10 (Saddle Point of WBE Sequences).** When all $C_{k,k}$ are equal, and furthermore all $\Theta_k$ are equal, then WBE sequences minimize the maximum eigenvalue of $C$, and hence maximize (8.32). Principal eigenvector of the $C$ matrix becomes the vector of ones $1^T = [1\ldots1]$.

*Proof.* Express maximum eigenvalue using the Rayleigh quotient. First, $1^T C 1$ is minimized when $C$ consists of WBE sequences, because they minimize total square correlation [2, 24], and because all $\Theta_k$ are equal. From the other side, if $C$ consists of WBE sequences, then $1$ is the principal eigenvector, by the “uniformly good” property of WBE sequences [25]. Namely, this property states that total square correlation of any user $k$, with sequences of all other users $l \neq k$ is independent from $k$, which means that $1$ is the eigenvector of the $C$ matrix, if all $\Theta_k$ are equal and all $C_{k,k}$ are equal. Also, it is principal eigenvector because all entries of $C$ are non-negative. \qed

We conclude that if all single-user channels have the same physical characteristics, namely $S_{k,\text{su}}$ are all equal, then the set of WBE sequences “extends” the robust slope region MUD$_k$ maximally in the direction of equal slopes $S = S_k$. Because the robust slope region is achieved irrespective of users data rates, then WBE sequences are appropriate for variable-rate systems as well, when we desire to be fair, and guarantee each user equal policy in coarsely coordinated multiuser systems.

### 8.6 Appendix: Chapter Proofs

*Proof of Property 3.* First we require the following Lemma.
Lemma 7. Let $A_1 \subseteq G_1$ and $A_2 \subseteq G_2$, where both $A_1$ and $A_2$ are finite antenna configurations. Let $\Phi_{A_1}$ denote the correlation matrix for points in $A_1$, and let $\Phi_{A_2}$ denote the correlation matrix for points in $A_2$. Finally, let $\Phi_{A_1,A_2}$ denote the cross-correlation between points from $A_1$ and points from $A_2$. We claim that

$$\frac{[\text{tr}(\Phi_{A_1}) + \text{tr}(\Phi_{A_2})]^2}{\text{tr} \left( \begin{bmatrix} \Phi_{A_1} & \Phi_{A_1,A_2} \\ \Phi_{A_2,A_1} & \Phi_{A_2} \end{bmatrix} \right)^2} \leq \frac{\text{tr}(\Phi_{A_1})^2}{\text{tr}(\Phi_{A_1}^2)} + \frac{\text{tr}(\Phi_{A_2})^2}{\text{tr}(\Phi_{A_2}^2)}. \quad (8.33)$$

Proof. Note that (8.33) is equivalent to

$$\frac{[\text{tr}(\Phi_{A_1}) + \text{tr}(\Phi_{A_2})]^2}{\text{tr}(\Phi_{A_1}^2) + \text{tr}(\Phi_{A_2}^2) + \text{tr}(\Phi_{A_1,A_2}^2) + \text{tr}(\Phi_{A_2,A_1}^2)} \leq \frac{\text{tr}(\Phi_{A_1})^2}{\text{tr}(\Phi_{A_1}^2)} + \frac{\text{tr}(\Phi_{A_2})^2}{\text{tr}(\Phi_{A_2}^2)}. \quad (8.34)$$

By the symmetry property of the cross-correlation matrix $\Phi$, we have that $\Phi_{A_2,A_1} = \Phi_{A_1,A_2}^\dagger$. From here it follows $\text{tr}(\Phi_{A_1,A_2} \Phi_{A_2,A_1}) = \text{tr}(\Phi_{A_1,A_2} \Phi_{A_1,A_2}^\dagger) = \|\Phi_{A_1,A_2}\|_{\text{FRO}}^2 \geq 0$. Similarly, it also holds that $\text{tr}(\Phi_{A_2,A_1} \Phi_{A_1,A_2}) \geq 0$. Hence, to prove (8.34) it suffices to show that

$$\frac{[\text{tr}(\Phi_{A_1}) + \text{tr}(\Phi_{A_2})]^2}{\text{tr}(\Phi_{A_1}^2) + \text{tr}(\Phi_{A_2}^2)} \leq \frac{\text{tr}(\Phi_{A_1})^2}{\text{tr}(\Phi_{A_1}^2)} + \frac{\text{tr}(\Phi_{A_2})^2}{\text{tr}(\Phi_{A_2}^2)}, \quad (8.35)$$

because the left hand side of (8.35) is always greater than the left hand side of (8.34). Upon multiplication of (8.35) by denominators $[\text{tr}(\Phi_{A_1}^2) + \text{tr}(\Phi_{A_2}^2)] \text{tr}(\Phi_{A_1}) \text{tr}(\Phi_{A_2}^2)$ and cancelling common terms, (8.35) reduces to the following

$$2 \text{tr}(\Phi_{A_1}) \text{tr}(\Phi_{A_2}) \text{tr}(\Phi_{A_1}^2) \text{tr}(\Phi_{A_2}^2) \leq \text{tr}(\Phi_{A_1})^2 [\text{tr}(\Phi_{A_1}^2)]^2 + \text{tr}(\Phi_{A_2})^2 [\text{tr}(\Phi_{A_2}^2)]^2 \quad (8.36)$$

$$0 \leq \text{tr}(\Phi_{A_1}) \text{tr}(\Phi_{A_2}^2) - \text{tr}(\Phi_{A_2}) \text{tr}(\Phi_{A_1}^2). \quad (8.37)$$

Since (8.37) is obvious truth, the proof of Lemma 7 is now complete. \qed
Using the main statement (8.33) of Lemma 7, we have that
\[
\frac{|\text{tr}(\Phi_{A_1}) + \text{tr}(\Phi_{A_2})|^2}{\text{tr}\left(\begin{bmatrix} \Phi_{A_1} & \Phi_{A_1,A_2} \\ \Phi_{A_2,A_1} & \Phi_{A_2} \end{bmatrix}^2\right)} \leq \sup_{A_1 \subseteq G_1 \atop |A_1| < \infty} \frac{(\text{tr} \Phi_{A_1})^2}{\text{tr}(\Phi_{A_1}^2)} + \sup_{A_2 \subseteq G_2 \atop |A_2| < \infty} \frac{(\text{tr} \Phi_{A_2})^2}{\text{tr}(\Phi_{A_2}^2)} = \omega(G_1) + \omega(G_2).
\]
(8.38)

To complete the last step of the proof for Property 2, we next take the supremum on the left of (8.38) to arrive at \(\omega(G) \leq \omega(G_1) + \omega(G_2)\).

Examination of (8.34), (8.35) and (8.37) will produce necessary and sufficient condition for achieving the upper bound (8.2). Left hand sides of (8.34) and (8.35) are equal if and only if \(\text{tr}(\Phi_{A_1,A_2} \Phi_{A_1,A_2}^\dagger) = \|\Phi_{A_1,A_2}\|_{\text{FRO}}^2 = 0\), which means that each point of \(A_1\) is fully uncorrelated with each point of \(A_2\). As an approximation, this occurs when \(A_1\) and \(A_2\) are separated by several wavelengths, which is true when \(G_1\) and \(G_2\) are separated by several wavelengths. The equality in (8.37) is satisfied when \(\text{tr}(\Phi_{A_1}) \text{tr}(\Phi_{A_2}^2) = \text{tr}(\Phi_{A_2}) \text{tr}(\Phi_{A_1}^2)\). To conclude, given \(G_1\) and \(G_2\) which are separated by several wavelengths, the upper bound (8.2) can only be achieved by designing configurations so that \(\text{tr}(\Phi_{A_1}) \text{tr}(\Phi_{A_2}^2) = \text{tr}(\Phi_{A_2}) \text{tr}(\Phi_{A_1}^2)\). \(\square\)

Proof of Property 4. For clarity sake, the proof of this Property will be provided in two steps. In the first step, we address the scenario where the space \(G\) has a certain size limitation (see below); in the second step, we generalize the proof for an arbitrary sized \(G\).

Step 1. Choose any \(\epsilon \in (0, 1)\). Let \(\alpha\) be the smallest non-negative real number, for which it holds that \(J_0(2\pi \left[\frac{\alpha}{a}\right]) > \epsilon\). Since the Bessel function is monotonically decreasing before the first zero, then \(J_0(2\pi \left[\frac{\alpha}{a}\right]) > J_0(2\pi \left[\frac{a}{\lambda}\right])\) for \(x \in (0, a)\). Suppose that space \(G\) is small enough to be covered by a sphere of diameter \(a\). If \(A\) is any finite subset of \(G\), then each individual entry of the correlation matrix \(\Phi_A\) is
greater than \( \epsilon \), because of the condition that \( J_0 \left( 2\pi \left[ \frac{\varphi}{\Lambda} \right] \right) > J_0 \left( 2\pi \left[ \frac{\varphi}{\Lambda} \right] \right) > \epsilon \). Noting \( \text{tr} \left( \Phi_A^2 \right) = \|\Phi_A\|_{FRO}^2 \geq \epsilon^2 |A|^2 \geq \epsilon |A|^2 = \epsilon \text{tr} \left( \Phi_A \right)^2 \), where \( |A| \) is the cardinality of \( A \), it follows that

\[
\frac{\left( \text{tr} \Phi_A \right)^2}{\text{tr} \left( \Phi_A \right)} \leq \epsilon^{-1} < \infty. \tag{8.39}
\]

Since (8.39) is true for an arbitrary configuration \( A \), and \( \epsilon \) is independent from the configuration, taking the supremum over all \( A \subseteq G \) and \( A < \infty \) on the left hand side of (8.39) results in \( \omega(G) \leq \epsilon^{-1} < \infty \). This completes the Step 1.

**Step 2.** Here, we address the case of arbitrary sized \( G \). Main idea is reduction to Step 1, and the use of monotonicity (Property 1) and subadditivity (Property 3) of the antenna-space capacity. Given any bounded aperture \( G \), it can be covered by only finitely many \( G_k \), where \( G_k \) are spheres with diameter \( a \), as specified in the Step 1. We therefore have \( G \subseteq \bigcup_{k=1}^K G_k \). From here, we first use the monotonicity, and then the finite subadditivity to arrive at

\[
\omega(G) \leq \omega \left( \bigcup_{k=1}^K G_k \right) \leq \sum_{k=1}^K \omega(G_k) \leq K \epsilon^{-1} < \infty. \tag{8.40}
\]

This completes the proof of Property 4, for an arbitrary sized \( G \).

*Proof of Theorem 9.* The goal is to prove that \( \omega(G) \) from (8.3) is the least upper bound for the effective antenna-space of an *arbitrary* circular configuration. For simplicity sake, we present the proof in two steps. In the first step, we focus on configurations that are constrained to angle values of \( \varphi \) which are rational multiples of \( 2\pi \). In the second step, we extend the proof for arbitrary circular configurations.

**Step 1.** Consider an arbitrary finite antenna configuration \( A_1 \), with the constraint that antennas are located on angle values of \( \varphi \) which are rational multiples of \( 2\pi \). Let \( A_2 \) be a uniform finite antenna configuration which is superset of the original
one, namely $A_1 \subset A_2$, like illustrated in the Figure 8.12. The finite configuration $A_2$ which is portrayed on the second circle in Figure 8.12, exists only because of the initial constraint that $A_1$ is located on rational multiples of $2\pi$. We next show that $A_2$ renders a bigger effective antenna-space than $A_1$.

To demonstrate this, we will utilize the following simple inequality

$$
\omega^{-1}_{A_1} = \frac{\text{tr} (\Phi^2_{A_1})}{\text{tr} (\Phi_{A_1})^2} = \frac{\text{tr} (\Phi^2_{A_1})}{|A_1|^2} \geq \min_{p_1 | p_1 = 1} p^T \left[ \Phi_{A_2} \odot \Phi_{A_2}^T \right] p,
$$

where $\odot$ is the elementwise (Hadamard) product of two equal sized matrices. Note that entries of $\Phi_{A_2} \odot \Phi_{A_2}^T$ are square antenna correlations for configuration $A_2$. Inequality in (8.41) is trivially true, because by putting $p_{A_1} = 1/|A_1|$ for indexes in $A_1$ and 0 otherwise on the right of (8.41) produces the left of (8.41). Since the right hand side of (8.41) is minimized over all choices for the vector $p$, it must be less or equal than the left side of (8.41).

To prove that antenna configuration $A_2$ produces larger effective antenna-space than configuration $A_1$ (meaning that $\omega_{A_2} \geq \omega_{A_1}$), we show that the right hand side of (8.41) is in fact $\omega^{-1}_{A_2}$. This is the topic of the following lemma.
Lemma 8. The right hand side of (8.41) equals

$$\min_{p^T 1 = 1} p^T \left[ \Phi_{A_2} \odot \Phi_{A_2}^\dagger \right] p = \frac{\text{tr} \left( \Phi_{A_2}^2 \right)}{(\text{tr} \Phi_{A_2})^2} = \omega_{A_2}^{-1}, \tag{8.42}$$

which is achieved only for $p = \left[ \frac{1}{|A_2|} \right] 1$. Here, $1$ is the vector of all ones.

Proof. The matrix $\Phi_{A_2} \odot \Phi_{A_2}^\dagger$ is non-negative, because it is a Hadamard product of two non-negative (correlation) matrices $\Phi_{A_2}$. This fact is a direct consequence of the Hadamard Product Lemma, whose proof is found in [42]. Since the matrix $\Phi_{A_2} \odot \Phi_{A_2}^\dagger$ is non-negative, then $p^T \left[ \Phi_{A_2} \odot \Phi_{A_2}^\dagger \right] p$ is convex (∪-shaped) function of the vector $p$. Therefore, the Lagrange Multiplier strategy is guaranteed to produce the minimum (8.42). For this purpose, we let $\mu$ be the Lagrange multiplier for the constraint $p^T 1 = 1$. Lagrangian function is given in (8.43), its differential is set to zero in (8.44), and the solution for $p$ is given in (8.45), as

$$p^T \left[ \Phi_{A_2} \odot \Phi_{A_2}^\dagger \right] p + \mu \left( p^T 1 - 1 \right) \tag{8.43}$$

$$\left[ \Phi_{A_2} \odot \Phi_{A_2}^\dagger \right] p = \mu 1 \tag{8.44}$$

$$p = \mu \left[ \Phi_{A_2} \odot \Phi_{A_2}^\dagger \right]^{-1} 1. \tag{8.45}$$

In transition (8.43) to (8.44), which is done by differentiation, we used the abbreviated $\mu$ instead of $-2\mu$, to shortcut the calculation. In (8.45), the Lagrange multiplier $\mu$ must be the reciprocal of $1^T [\Phi_{A_2} \odot \Phi_{A_2}^\dagger]^{-1} 1$, because $p$ must sum to unity.

Since $A_2$ is by construction uniform circular array, then $\Phi_{A_2} \odot \Phi_{A_2}^\dagger$ must be Toeplitz matrix, where each row is obtained from the previous one by a circular shift. All rows of $\Phi_{A_2} \odot \Phi_{A_2}^\dagger$ therefore sum equally, which means that $p = \frac{1}{|A_2|} 1$ satisfies (8.44) by inspection. This optimal choice of $p$ produces the equality in (8.42) and completes the proof of this Lemma. \qed
Lemma 8 demonstrates that uniform circular array $A_2$ yields larger effective antenna-space from the original configuration $A_1$. Next, we simply continue the process with $A_3$ which is uniform circular array that contains $A_2$, and in general with $A_k$ which contains $A_{k-1}$, as illustrated in the Figure 8.12. In order to obtain the limit point for the effective antenna-space when $k \to \infty$, we first find $\text{tr} \left( \Phi_{A_k}^2 \right)$ for a generic uniform circular array.

For a generic uniform circular array, the sum square correlation from the antenna which is located at $\varphi = 0$, to all antennas (including itself) is given as

$$
\sum_{n=0}^{\lfloor A_k \rfloor - 1} |n| \left| J_0 \left( \frac{2\pi r}{\lambda} \sqrt{2 - 2 \cos \left( \frac{2\pi n}{|A_k|} \right)} \right) \right|^2
$$

(8.46)

Justification for (8.46) is a simple cosine law: distance between two points on the circle of unit radius, which are $\varphi$ apart is $2 - 2 \cos(\varphi)$. For uniform circular array, we have $\varphi = \frac{2\pi n}{|A_k|}$, and (8.46) follows. To obtain the total square correlation $\text{tr}(\Phi_{A_k}^2)$ between all points of the uniform circular array, we ought multiply (8.46) by the total number of antennas, which is $|A_k|$.

Hence, we have that

$$
\frac{\text{tr} \left( \Phi_{A_k}^2 \right)}{\text{tr} \left( \Phi_{A_k} \right)^2} = \frac{\text{tr} \left( \Phi_{A_k}^2 \right)}{|A_k|^2} = \frac{1}{|A_k|} \sum_{n=0}^{\lfloor A_k \rfloor - 1} \left| J_0 \left( \frac{2\pi r}{\lambda} \sqrt{2 - 2 \cos \left( \frac{2\pi n}{|A_k|} \right)} \right) \right|^2
$$

(8.47)

$$
\downarrow \frac{1}{2\pi} \int_{-\pi}^{\pi} \left| J_0 \left( \frac{2\pi r}{\lambda} \sqrt{2 - 2 \cos \varphi} \right) \right|^2 d\varphi.
$$

(8.48)

Convergence in (8.48) is as $k \to \infty$ and $A_{k-1} \subset A_k$, which is portrayed by the third circle in the Figure 8.12. Thus, beginning with an arbitrary antenna configuration $A_1$, which is located on rational multiples of $2\pi$, we constructed the sequence $A_k$ of antenna configurations, which monotonically increases the effective antenna-space, to become the inverse of (8.48). This completes Step 1.
Step 2. In the last stage of the proof, we consider the case of initial configuration $A$, which is not located on rational multiples of $2\pi$ like $A_1$, and a finite uniformly spaced superset $A_2$ doesn’t exist. Here, we invoke the continuity argument for the total square correlation, and for the function $(\text{tr} \Phi_A)^2/\text{tr}(\Phi_A^2)$, when $A$ is finite. For brevity of exposition we label (8.48) as $\omega^{-1}_{\text{rat}}$.

Given any $\epsilon > 0$, and finite $A$ which is located on the circle, there exists a finite $A'$, which is located on rational multiples of $2\pi$, close to $A$, and it holds that

$$\left| \frac{(\text{tr} \Phi_A)^2}{\text{tr}(\Phi_A^2)} - \frac{(\text{tr} \Phi_{A'})^2}{\text{tr}(\Phi_{A'}^2)} \right| < \epsilon. \quad (8.49)$$

From the first stage of the proof, because $A'$ is positioned on rational multiples of $2\pi$, we have that $(\text{tr} \Phi_{A'})^2/\text{tr}(\Phi_{A'}^2) \leq \omega_{\text{rat}}$. Combining this inequality with (8.49), we have that for any $A$,

$$\frac{(\text{tr} \Phi_A)^2}{\text{tr}(\Phi_A^2)} \leq \frac{(\text{tr} \Phi_{A'})^2}{\text{tr}(\Phi_{A'}^2)} + \frac{(\text{tr} \Phi_A)^2}{\text{tr}(\Phi_A^2)} - \frac{(\text{tr} \Phi_{A'})^2}{\text{tr}(\Phi_{A'}^2)} \leq \omega_{\text{rat}} + \epsilon. \quad (8.50)$$

Because $\epsilon$ is arbitrary small, we conclude that for any $A$, it holds $(\text{tr} \Phi_A)^2/\text{tr}(\Phi_A^2) \leq \omega_{\text{rat}}$. Hence, $\omega_{\text{rat}}$, which is the inverse of (8.48) is the least upper bound even for the effective antenna-space, even for configurations which are not located on rational multiples of $2\pi$. \qed
Proof of Proposition 5. Notice that each users multi-antenna signal $\mathcal{H}_k X_k$, which is a column vector, is further modulated by a spreading sequence $s_k^T$, which is a row vector. Therefore, if $s_k(l)$ represents $l$–th chip of user $k$, then $s_k(l) H_k X_k$ is users $k$ contribution to the received signal at chip timing $l$, which only contributes to $l$–th column, namely $Y(\cdot, l)$, of the received matrix $Y$.

If the receiver stacks columns of $Y$ on top of each other, with a stack operator $\cup$, then the channel model may be represented as

$$
\begin{bmatrix}
  Y(\cdot, 1) \\
  \vdots \\
  Y(\cdot, L)
\end{bmatrix} = \sum_{k=1}^K \begin{bmatrix}
  s_k(1) H_k \\
  \vdots \\
  s_k(L) H_k
\end{bmatrix} X + \begin{bmatrix}
  \eta(\cdot, 1) \\
  \vdots \\
  \eta(\cdot, L)
\end{bmatrix} \tag{8.51}
$$

$$
Y^\cup = \sum_{k=1}^K [s_k \otimes \mathcal{H}_k] X_k + \eta^\cup, \tag{8.52}
$$

where $\otimes$ represents the Kronecker matrix product, defined exactly as in (8.51).

Since $X_k$ is independent and identically distributed, for a given $\mathcal{H}_k$, covariance of the signal part from user $k$ is given as

$$
\mathbb{E} \left[ [s_k \otimes \mathcal{H}_k] X X^\dagger [s_k \otimes \mathcal{H}_k]^\dagger \mid \mathcal{H}_k \right] = [s_k \otimes \mathcal{H}_k] [s_k \otimes \mathcal{H}_k]^\dagger \frac{P_k}{N_k} \tag{8.53}
$$

$$
= [s_k \otimes \mathcal{H}_k] \left[ s_k^\dagger \otimes \mathcal{T}_k^\dagger \right] \frac{P_k}{N_k}, \tag{8.54}
$$

which simply describes how Hermitian operator distributes inside the Kronecker product; this is easily verified from the definition (8.51).

To evaluate the aggregate expression for mutual information $C(P)$, we first need the following lemma, which contains one Kronecker product property (8.55), which is necessary to specify the $C(P)$. The other Kronecker product property, namely (8.56), will be exploited later.
Lemma 9. If $A, B, C, D$ are matrices, following identities hold for Kronecker product

$$ (A \otimes B) (C \otimes D) = (AC) \otimes (BD) $$ \hspace{1cm} (8.55)

$$ \text{tr}(A \otimes B) = \text{tr}(A) \text{ tr}(B) $$ \hspace{1cm} (8.56)

Proof. For the sake of clarity, we believe that it is the best to demonstrate the principle of (8.55), first assuming that $A$ and $C$ are $2 \times 2$ matrices. A more general case is straightforward, and it follows the same principle.

$$ (A \otimes B) (C \otimes D) = \begin{bmatrix} A(1,1)B & A(1,2)B \\ A(2,1)B & A(2,2)B \end{bmatrix} \begin{bmatrix} C(1,1)D & C(1,2)D \\ C(2,1)D & C(2,2)D \end{bmatrix} = $$

$$ \begin{bmatrix} A(1,1)C(1,1)BD + A(1,2)C(2,1)BD & A(1,1)C(1,2)BD + A(1,2)C(2,2)BD \\ A(2,1)C(1,1)BD + A(2,2)C(2,1)BD & A(2,1)C(1,2)BD + A(2,2)C(2,2)BD \end{bmatrix} = (AC) \otimes (BD) $$ \hspace{1cm} (8.57)

More general case, when $A$ and $C$ are not $2 \times 2$ is easily verified as well, but we choose not to dwell in the sea of indecent.

Verification of (8.56) is almost transparent too, because if $A$ is say $N \times N$, then entries starting from $A(1,1)B$, through $A(N,N)B$, are blocks on the diagonal of $A \otimes B$. Operator “tr” simply sums diagonal elements of $A \otimes B$ and thereby produces right hand side of (8.56).

With (8.55) we are ready to state the capacity function, which defines the rate region for all $K$ users. First, the covariance of user $k$ signal is given as

$$ [s_k \otimes \mathcal{H}_k] \left[ s_k^\dagger \otimes \mathcal{H}_k^\dagger \right] \frac{P_k}{N_k} = (s_k s_k^\dagger) \otimes (\mathcal{H}_k \mathcal{H}_k^\dagger) \frac{P_k}{N_k}, $$ \hspace{1cm} (8.58)

where users $k$ power $P_k$ is allocated evenly, across all $N_k$ transmit antennas. The noise covariance is identity, namely $I$. Here, (8.58) is justified by (8.55).
From (8.58) we are ready to state the expression for mutual information as

\[
C(P) = \mathbb{E} \ln \det \left[ I + \sum_{k=1}^{K} (s_k s_k^\dagger) \otimes (\mathcal{H}_k \mathcal{H}_k^d) \frac{P_k}{N_k} \right]. \tag{8.59}
\]

Now we are ready to evaluate entries of the wideband matrix \( \mathcal{C} \), by evaluating partials of \( C(P) \). Lemma 1 provided us with necessary tools to evaluate partials of \( C(P) \) from (8.59).

With the aid of (5.20) part of the Lemma 1, and the expression (8.59), we evaluate the first partials as follows

\[
\frac{\partial C(0)}{\partial P_k} = \mathbb{E} \text{tr} \left[ \left( s_k s_k^\dagger \right) \otimes \left( \mathcal{H}_k \mathcal{H}_k^d \right) \right] N_k^{-1} \tag{8.60}
\]

\[
= \mathbb{E} \left[ \text{tr} \left( s_k s_k^\dagger \right) \text{tr} \left( \mathcal{H}_k \mathcal{H}_k^d \right) \right] N_k^{-1}, \tag{8.61}
\]

where transition (8.60) to (8.61) used (8.55) from the Lemma 9. Now we use the fact that spreading codes \( s_k \) are unit energy, and \( \text{tr} \left( s_k s_k^\dagger \right) = \text{tr} \left( s_k^\dagger s_k \right) = 1 \).

From (8.61) we have that

\[
\frac{\partial C_A(0)}{\partial P_k} = \mathbb{E} \text{tr} \left[ \mathcal{H}_k \mathcal{H}_k^d \right] N_k^{-1}. \tag{8.62}
\]

Now we evaluate cross-partial, using the formula (5.22), from the Lemma 1.

\[
-\frac{\partial^2 C_A(0)}{\partial P_k \partial P_l} = \mathbb{E} \text{tr} \left[ \left( s_k s_k^\dagger \right) \otimes \left( \mathcal{H}_k \mathcal{H}_k^d \right) \right] \left[ \left( s_l s_l^\dagger \right) \otimes \left( \mathcal{H}_l \mathcal{H}_l^d \right) \right] N_k^{-2} \tag{8.63}
\]

\[
= \mathbb{E} \text{tr} \left[ \left( s_k s_k^\dagger s_l s_l^\dagger \right) \otimes \left( \mathcal{H}_k \mathcal{H}_k^d \mathcal{H}_l \mathcal{H}_l^d \right) \right] N_k^{-2} \tag{8.64}
\]

\[
= \text{tr} \left( s_k s_k^\dagger s_l s_l^\dagger \right) \mathbb{E} \text{tr} \left[ \mathcal{H}_k \mathcal{H}_k^d \mathcal{H}_l \mathcal{H}_l^d \right] N_k^{-2} \tag{8.65}
\]

Here, (8.63)-(8.65) uses the Kronecker product Lemma (9). To proceed from (8.65), we use \( \text{tr} \left( s_k s_k^\dagger s_l s_l^\dagger \right) = \text{tr} \left( s_k^\dagger s_l s_l^\dagger s_k \right) = \rho(k,l) \rho^*(k,l) = |\rho(k,l)|^2 \). It now follows that

\[
-\frac{\partial^2 C_A(0)}{\partial P_k \partial P_l} = |\rho(k,l)|^2 \mathbb{E} \text{tr} \left[ \mathcal{H}_k \mathcal{H}_k^d \mathcal{H}_l \mathcal{H}_l^d \right] N_k^{-2}. \tag{8.66}
\]
To evaluate entries of the Wideband Matrix $\mathcal{C}$, (8.66) is divided with the product of (8.62) for $k$ and $l$

$$
\mathcal{C}_{k,l} = \frac{\text{Etr} \left[ \mathcal{H}_k \mathcal{H}_k^\dagger \mathcal{H}_l \mathcal{H}_l^\dagger \right]}{2 \text{Etr} \left[ \mathcal{H}_k \mathcal{H}_k^\dagger \right] \text{ETr} \left[ \mathcal{H}_l \mathcal{H}_l^\dagger \right]} |\rho(k,l)|^2.
$$

(8.67)

This is the same as the Wideband Matrix $\mathcal{C}$ where $\mathcal{H}_k$ operates along the row dimensions of $X_k$, but only now each term is scaled by $|\rho(k,l)|^2$. $\square$
Chapter 9

Conclusions and Future Work

Theory of multiuser low power systems is emerging, fueled by wide variety of applications in ultrawideband, satellite and sensor networks. In this thesis, we developed a mathematical framework for multiuser systems with both transmit and receive antenna arrays, which will enable designers to make appropriate choices in low power systems.

The capacity theorem in low power systems is entirely described by the spectral efficiency slope, which is evaluated at the minimum energy per information bit, for single user systems. Naturally, in multiple access systems, this approach extends to become the multiuser spectral efficiency slope region, which describes tradeoffs between data rate and transmit power of each user, and the aggregate system bandwidth. The spectral efficiency slope region is a low power counterpart to the multiuser rate region, which is mathematically intractable for generic fading multiantenna channels. Low power analysis enables us to tackle problems which are not tractable in general and some surprising designs conclusions emerge.

For example, we have seen how the performance of single user detectors is superior to the performance of time division multiple access system, in low power fading multiantenna environments. This conclusion is not true in high power communication regime, where single user detectors force both transmit power and the multiple access interference to grow beyond bounds.

Furthermore, low power analysis enables us to develop the theory of coarsely
coordinated multiuser systems, where we show how multiuser detection can be quite beneficial even without tight user coordination, as demonstrated by comparing users policies within orthogonal and non-orthogonal systems. In essence, polices from MUD\(_s\) fully solve the problem of coarse coordination in low power multiuser systems, and are superior to any classic orthogonal system. Policies from MUD\(_s\) are effectively represented with the concept of virtual antenna partition, which provides the appropriate intuition about MUD\(_s\), and re-defines the notion of degrees of freedom, not in terms of vector spaces, but rather in terms of coarse coordination of multiple users.

Finally, receiver space dimension \(\omega\), which captures correlated antenna arrays, appears fundamental to the study of spatial degrees of freedom, and it is valid for both tightly and coarsely coordinated multiuser system. A natural extension of \(\omega\), for a fixed antenna configuration, considers all possible antenna configurations inside a limited space to maximally expand \(\omega\) and converge to the finite limit of Space Efficiency. The Space Efficiency limit is truly a supremum and not a maximum, because the optimum solution doesn’t exist, but rather only a sequence of approximations which converges to Space Efficiency can be found. Here we have shown how to evaluate the Space Efficiency limit for the boundary of a circle, whereas more asymmetric geometries remain an open problem.
Appendix A

Appendix: Additional Facts and Proofs

Lemma 10. Function $C(P)$, which maps $\mathbb{R}_+^K \to \mathbb{R}$, and which is defined as

$$ C(P) = \text{E} \log \det \left[ I + \sum_{k=1}^{K} \mathcal{H}_k \mathcal{H}_k^\dagger P_k \right], $$

(A.1)

in a concave function of the vector $P = (P_1, \ldots, P_K) \in \mathbb{R}_+^K$.

Following proof can be found in many references, but we include it for the sake of completeness.

Proof. First, it suffices to prove concavity of (A.1) without the expectation operator $\text{E}$, because $\text{E}$ is linear. Second, it is enough to show that $C(P)$ is concave, when restricted to a line in the domain $\mathbb{R}_+^K$. Furthermore, a line in the domain can be expressed as $P = P^{(0)} + tP^{(1)}$, where $t$ is a scalar so that both $P^{(0)} \in \mathbb{R}_+^K$ and $P \in \mathbb{R}_+^K$. Note that $P^{(1)}$ may have negative entries.

With these reductions, the function $C(P)$ becomes a simple function of the scalar argument $t$, namely

$$ C(t) = \log \det [B_0 + tB_1], $$

(A.2)

where $B_0 = I + \sum_{k=1}^{K} \mathcal{H}_k \mathcal{H}_k^\dagger P_k^{(0)}$, and also $B_1 = \sum_{k=1}^{K} \mathcal{H}_k \mathcal{H}_k^\dagger P_k^{(1)}$.

It now follows that

$$ C(t) = \log \det \left[ B_0^{\frac{1}{2}} \right] + \log \det \left[ I + tB_0^{-\frac{1}{2}} B_1 B_0^{-\frac{1}{2}} \right] + \log \det \left[ B_0^{\frac{1}{2}} \right], $$

(A.3)
First, we note that $B_0 \geq 0$ because $P^{(0)} \in \mathbb{R}^{K}_{+}$ and each matrix $H_k H_k^t \geq 0$ is almost surely positive definite. Second, argument $B_0 + tB_1 \geq 0$ because it simply represents the argument of (A.1). It now follows that $I + tB_0^{-\frac{1}{2}} B_1 B_0^{-\frac{1}{2}} \geq 0$, and eigenvalues of $I + tB_0^{-\frac{1}{2}} B_1 B_0^{-\frac{1}{2}}$ are $1 + t\lambda_m \geq 0$, where $\lambda_m$ is an eigenvalue of $B_0^{-\frac{1}{2}} B_1 B_0^{-\frac{1}{2}}$.

It now follows that

$$C(t) = \log \det \left[ B_0^{\frac{1}{2}} \right] + \sum_{m=1}^{M} \log(1 + t\lambda_m) + \log \det \left[ B_0^{\frac{1}{2}} \right], \quad (A.4)$$

which is a concave function of $t$. \qed
Bibliography


[26] 3GPP-TSG RAN, WG1 #3 “Uplink Common Packet Channel” in *http://www.3gpp.org*


[29] D. Forney “Introduction to Digital Communication” “Class Notes”


