UNIVERSITY COLLEGE LONDON

Resource Allocation for Delay Constrained Wireless Communications

by

Jia Chen

A thesis submitted to

Department of Electronic and Electrical Engineering

in partial fulfillment for the degree of Doctor of Philosophy

Email: {j.chen}@ee.ucl.ac.uk

April 2010

Declaration of Authorship

I, Jia Chen, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Signed:

Date:

Abstract

The ultimate goal of future generation wireless communications is to provide ubiquitous seamless connections between mobile terminals such as mobile phones and computers so that users can enjoy high-quality services at anytime anywhere without wires. The feature to provide a wide range of delay constrained applications with diverse quality of service (QoS) requirements, such as delay and data rate requirements, will require QoS-driven wireless resource allocation mechanisms to efficiently allocate wireless resources, such as transmission power, time slots and spectrum, for accommodating heterogeneous mobile data. In addition, multiple-input-multiple-output (MIMO) antenna technique, which uses multiple antennas at the transmitter and receiver, can improve the transmission data rate significantly and is of particular interests for future high speed wireless communications.

In the thesis, we develop smart energy efficient scheduling algorithms for delay constrained communications for single user and multi-user single-input-single-output (SISO) and MIMO transmission systems. Specifically, the algorithms are designed to minimize the total transmission power while satisfying individual user's QoS constraints, such as rate, delay and rate or delay violation. Statistical channel information (SCI) and instantaneous channel state information (CSI) at the transmitter side are considered respectively, and the proposed design can be applied for either uplink or downlink. We propose to jointly deal with scheduling of the users that access to the channel for each

frame time (or available spectrum) and how much power is allocated when accessing to the channel. In addition, the algorithms are applied with modifications for uplink scheduling in IEEE 802.16 Worldwide Interoperability for Microwave Access (WiMAX). The success of the proposed research will significantly improve the ways to design wireless resource allocation for delay constrained communications.

Contents

Acknowledgements

First and foremost, I would like to thank my supervisor Dr. Kai-Kit Wong for his supervision on my completing the thesis. He guided me into this challenging and interesting research area and inspired me to explore the frontier in this area. Without his deep insight, consistent encouragement and tremendous support, the thesis is impossible to complete. Working with him is a precious experience.

I would like to express my sincere thanks to my second supervisor Dr. Yang Yang for his valuable time reading my thesis and taking part in my PhD transfer committee. He guided my first step as a researcher and provided me many valuable suggestions and comments.

It has been a great pleasure to work along with my colleagues from Wireless Communication Research Group at UCL Adastral Park. Special thanks go to research fellows Dr. Jin Shi, Dr. Min Xie, Dr. Gan Zheng and Dr. Yangyang Zhang for their incisive comments and suggestions. Also I would like to thank my fellow PhD colleagues Caijun Zhong, Elsheikh Elsheikh and Li-Chia Choo for the fruitful discussions and for creating such a cheerful group environment.

I had the great opportunity to participate in a three-month internship at British Telecom (BT). I would like to thank members from BT design of mobility and solutions for providing me a welcoming environment and their warm hearted help on my research. Very special and deep thanks go to my mentor Dr. Nicholas Edwards from BT for his

incisive guidance and informative discussions. In addition, I appreciate very much for the enormous help from Pi Huang, Yu Zhang, Maurice Gifford, Milan Lalovic, Garath Evans, Neil Morley, Ji Zhang, Luan Huang, Robert Glassford, Dave Townend, Tim Costello and Ian Rose at BT.

Finally, I would like to dedicate the thesis to my parents and grandparents for their continuous support and love.

I acknowledge the sponsorship for my studies by Engineering and Physical Sciences Research Council. Special thanks for Chinese Government, Institution of Engineering and Technology and UCL for providing the awards to encourage my PhD study.

Abbreviation

AWGN	additive white Gaussian noise
BF	block fading
c.d.f.	cumulative distribution function
CDMA	code division multiple access
CSI	channel state information
CTC	convolutional Turbo coding
DIRECT	DIviding RECTangles
DP	dynamic programming
EC	effective capacity
EOPPB	equal-outage-probability per block
FDMA	frequency division multiple access
FDPB	forward decision per block
FTP	file transfer protocol
HDP	hyper dynamic programming
IEEE	institute of electrical and electronics engineers
i.i.d.	independent and identical distributed
IMI	instantaneous mutual information
ISI	inter-symbol interference
KKT	Karush-Kuhn-Tucker
LTE	long term evolution
MIMO	multiple-input-multiple-output
MPE	minimum power equation
MRC	maximum ratio combining

OFDM	orthogonal frequency division multiplexing
OFDMA	orthogonal frequency division multiplexing access
OLLP	one-step limited look-ahead policy
OSI Model	open systems interconnection reference model
p.d.f.	probability density function
PSK	phase shift keying
PUSC	partial usage of subcarriers
QAM	quadrature amplitude modulation
QoS	quality of service
QPSK	quadrature phase shift keying
SCI	statistical channel information
SINR	signal-to-interference-and-noise ratio
SISO	single-input-single-output
SM	spatial multiplexing
SNR	signal-to-noise ratio
SQP	sequential quadratic programming
STC	space time coding
SVD	singular value decomposition
TDMA	time division multiple access
UMTS	universal mobile telecommunications system
VoIP	voice over internet protocol
WiMAX	worldwide interoperability for microwave access
WLAN	wireless local area network

List of Figures

List of Tables

Chapter 1

Introduction

The past decade has seen the tremendous growth of wireless communications. The prevalent success of wireless cellular systems and wireless local area networks (WLANs) advances the future generation wireless communications to support a wide range of wireless applications beyond voice, such as on-line games and video conferencing. The increasing types of delay sensitive applications supported over wireless link will require different wireless Quality of Service (QoS) guarantees to satisfy diverse requirements for transmission speed and delay constraints, which inevitably leads to the rising demand for wireless resources, such as the scarce spectrum and the limited transmit power. This drives the focus of our research to investigate mechanisms to efficiently allocate the available wireless resource to meet the QoS required by delay sensitive applications.

In Chapter 1 of the thesis, we will introduce the motivation, objectives and contributions of the research. The organization of the chapter is as follows. Section 1.1 introduces the motivation and objectives of the research. The contributions of the research are summarized in Section 1.2. The thesis outline is presented in Section 1.3.

1.1 Motivation

1.1.1 Wireless Communications

Wireless communications have grown tremendously in the past decade. The concept of wireless cellular communications was first developed by Bell laboratories during 1960's to provide wide area long distance wireless connections [1]. They have experienced rapid growth from analog cellular systems in the first generation wireless cellular network providing low speed voice transmission to digital cellular systems in the later generation wireless cellular networks offering higher speed voice and data transmission [2]. Cellular networks provide the entire area coverage by dividing the total area into multiple cells, with one base station in each cell providing wireless communications to multiple users using different multiple access methods [3]. Cellular networks have become critical tools for both business and everyday life in most countries and cellular wireless communications are gradually supplanting communications over wired systems [2, 4]. On the other side, WLANs, which were first proposed in 1979 to provide local area short distance wireless connections, are gradually replacing wired networks in local areas, such as homes, businesses and campuses [5]. Furthermore, the increasing demand for various new wireless applications, such as video transmissions, mobile entertainment, mobile healthcare and mobile remote education, has driven the further development of the next generation wireless networks providing wireless connections with better QoS, such as higher data rate and smaller delay [6].

However, wireless channels raise severe challenges for reliable communications, due to the constantly varying channel strength over time in unpredictable ways caused by fading [2, 3]. At a slow scale, channel quality varies due to large-scale fading effects, for which the received signal strength over distance changes because of the path loss and shadowing. At a fast scale, channel quality varies due to small-scale fading caused by multipath effects, in which the received signal strength changes

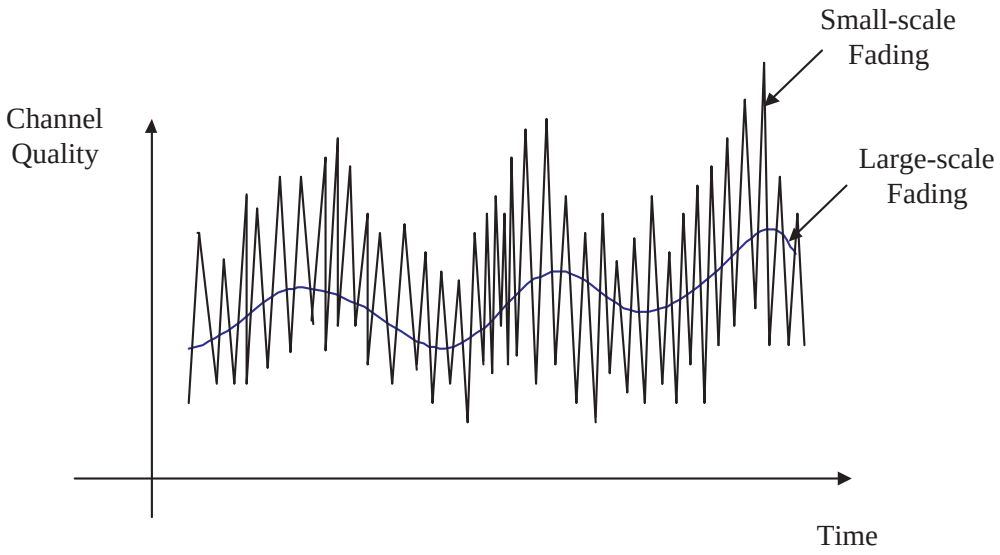

Figure 1.1: Variations of channel quality over time caused by fading.

because of the constructive and destructive interference of multiple signal paths due to reflection and refraction. The variations of channel quality over time are illustrated in Fig. 1.1.

Another challenge of wireless channels is interference. As signals from wireless users are transmitted over the air, a significant interference exists between the users due to the disruptive signal strength using the same frequency. Interference exists in different multiple access cellular systems. In code division multiple access (CDMA) cellular systems, where all users within the same cell access to the base station using the same frequency and time but different codewords, interference exists within the same cell (intra-cell interference) [7, 8]. The interference exists in the uplink, when multiple transmitters are communicating to a common receiver. And it exists in the downlink, when signals are transmitted from a single transmitter to multiple receivers. In addition, interference may exist between users in different cells using the same frequency (inter-cell interference) in CDMA system. Such kind of interference also exists in time division multiple access (TDMA) or frequency division multiple access (FDMA) systems, where users within the same cell access the base station in a time division or frequency division manner and

the same frequency are reused in different cells.

1.1.2 Wireless Resource Allocation

To overcome the existing challenges of wireless channels, a general wireless resource allocation strategy has been proposed to dynamically allocate wireless resources, such as transmit power, based on the channel conditions [9]. Power control schemes were designed to overcome fading and intra-cell interference by improving the signal-to-interference-and-noise ratios (SINRs) of the weakest users in CDMA systems. Since the SINR reflects the user's data communication quality and larger SINR results in higher transmission data rate, it is an important design consideration in such schemes. In order to achieve this, the stronger users which are closer to the base station are limited to a lower transmit signal power and the weaker users in the edge of the cells are allowed to transmit at a higher signal power [3]. In addition, power control schemes are designed in [10, 11, 12, 13, 14, 15] to overcome large-scale fading and inter-cell interference in different cellular systems.

Although power control schemes can improve the overall system performance, they have not considered individual user's QoS requirements, such as data rate. Furthermore, supporting users' QoS becomes essential in the design of wireless resource allocation strategies [16]. To achieve this goal, wireless resource allocation strategies have been further developed to satisfy individual user's reliable wireless communications [17, 18]. In such strategies, wireless transmit power is adapted to combat channel conditions including small-scale fading in order to maximize the average transmission rate that a physical wireless channel can support in a single user system.

For applications in multi-user systems, scheduling, which is the allocation of wireless resources including transmit power, time slots and frequencies, has been proposed to further improve the users' performance [19, 20]. The simplest scheduling

algorithm is round-robin scheduling, where each user is served in turn regardless of the channel conditions. Such scheduling may result in poor system throughput performance because users may transmit data when their channels are in severe fading [21]. When multiple users are communicating to a common base station, they have time varying wireless channels and multi-user diversity is provided in multi-user systems [20]. By tracking the channel variations between each user and the base station, base stations can exploit the multi-user diversity and scheduling transmissions to users with near maximum instantaneous channel quality.

The above scheduling algorithms considering multi-user diversity are called opportunistic scheduling algorithms [3]. This kind of scheduling algorithms have been developed in [22, 23, 24, 25, 26, 27, 28, 29, 30, 31] for achieving different system performance parameters, such as fairness which measures the distribution of available resources across all users, system throughput, and channel utilization efficiency which indicates whether the channel has been efficiently used by transmitting at a better channel quality. For example, user with the best channel quality is allocated for each time slot in maximum signal-to-noise ratio (SNR) scheduling [22, 23, 24]. The system throughput and link utilization are maximized for such algorithms but the users' fairness is poor. Another example is proportional fair scheduling algorithm, where users are scheduled based on the combined considerations of instantaneous SNR and average measured throughout [25]. In such scheduling algorithm, fairness, system throughput and channel utilization are balanced.

On the other hand, in response to the existing challenges of wireless channels and rapidly increasing demand for higher data rate, multiple-input-multiple-output (MIMO) technology, which is a configuration that has multiple antennas at both transmitter and receiver, has been introduced. Fig. 1.2 illustrates a MIMO channel with three transmit and receive antennas. MIMO channels offer larger channel capacity (or support higher transmission data rate) than single-input-single-output

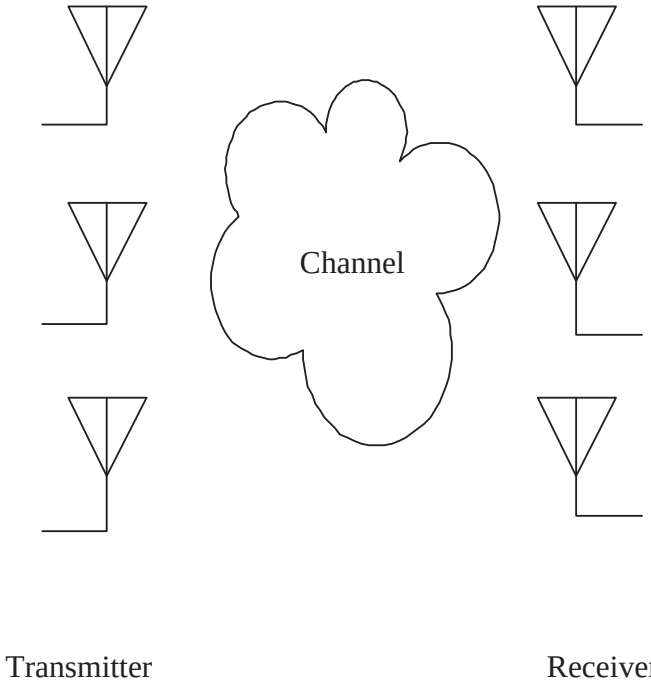

Figure 1.2: An illustration of a MIMO channel.

(SISO) channels by transmitting parallel data streams in space [32, 33, 34, 35]. Due to the great benefits MIMO brings about, it has been included in the next generation wireless network standards, such as Worldwide Interoperability for Microwave Access (WiMAX) and Long Term Evolution (LTE) [36, 37, 38].

In single user multiple antenna system, wireless resource allocation strategies have been proposed to achieve a maximum average transmission data rate by adjusting transmit power to the channel quality in [33, 39, 40, 41]. In multi-user MIMO systems, because of the increased number of transmit antennas, the fluctuation of the channel over time is not obvious. This limits the exploitation of multi-user diversity [20]. Accordingly, opportunistic beamforming was developed to induce larger and faster channel fluctuations and thus maximize the receive SNR at the mobile terminals [25]. Furthermore, the authors analyze the performance of opportunistic beamforming with proportionally fair scheduling in [25, 42]. Opportunistic beamforming is further developed to improve the system performance for larger user number and different fading channels in [43, 44, 45, 46]. Besides, multiple antennas

can separate the signals from different users spatially, which provides spatial multiple access. Such access method is called space division multiple access (SDMA) and it has driven the research work for scheduling algorithms combining SDMA and TDMA [47, 48, 49]. In such schemes, for each time slot, the best set of users with the least interference is allocated resources instead of a single user.

1.1.3 Delay Constrained Communications

An important problem that the traditional resource allocation strategies have not addressed is wireless resource allocation design for delay sensitive communications. Recently, the types of delay sensitive applications supported over wireless link are increasing rapidly, which require wireless resource allocation design to support users' QoS guarantees in terms of transmission data rate and delay constraint [21]. Different delay constrained applications have different QoS requirements. For example, Voice over IP (VoIP) calls have stringent requirements on delay and delay jitter but can cope with reasonable transmission rate. On-line video and games have the features of high speed and low delay requirements.

To address the diverse QoS requirements for accommodating heterogeneous applications, wireless standard bodies have accordingly defined QoS classes based on the delay sensitivities and data rates of wireless traffic for the next generation wireless networks. For instance, four QoS classes have been defined in LTE, including conversational class, streaming class, interactive class and background class [50]. The conversational class represents traffic which is very delay sensitive and usually used for real-time traffic flows, while the background class is the least delay sensitive traffic class applied to applications, such as email and webpage browsing. WiMAX, which is developed to support wireless broadband metropolitan area networks, has defined five QoS classes. These are unsolicited grant service

for VoIP, real-time polling service for streaming audio and video, extended real-time polling service for voice with activity detection, non-real-time polling service for file transfer protocol (FTP), and best-effort service for web browsing [51].

Delay has different interpretations from the viewpoint of different layers in Open Systems Interconnection Reference Model (OSI Model). From the viewpoint of physical layer, delay largely occurs during the process of coding and decoding. Longer transmit codeword improves the transmission performance but introduces larger delay [3]. From the viewpoint of data link layer, delay is referred to as buffering delay, which is caused by the queueing time when a packet arrives at the buffer, until it is being served by wireless link. The data link layer delay is partly affected by physical layer data transmission, which decides the packet service rate.

For delay constrained communications, QoS is traditionally considered as the guarantee of a fixed data rate under a decoding delay constraint, or a specified queueing delay constraint under a given packet arrival rate. In other words, it is regarded as no lost of information within a certain length of codeword or no violation of certain queueing delay bound. This type of guarantee is called deterministic QoS guarantee.

In traditional scheduling algorithms, such as opportunistic scheduling, no delay is considered and data are transmitted at the maximum allowable power, since lowering transmit power results in lowering transmission rates and hence degradation in QoS [52]. However, transmit power is an important and essential factor of wireless systems. As transmit power accounts for a dominant portion of the total energy consumption, reducing transmit power can greatly reduce energy waste and prolong battery life, especially for mobile terminals. Moreover, lowering transmit power can reduce the interference caused by same-frequency users. On the other hand, it is observed in [53] that transmitting data over wireless channel in reduced power can significantly result in larger delay. The delay-power tradeoff in wireless channels drives the development of energy efficient scheduling to minimize

transmit power subject to delay constraints for delay sensitive communications.

There are extensive research work focusing on energy efficient scheduling algorithms design to support deterministic QoS. In [53, 54, 55, 56], scheduling algorithms have been designed to meet the requirement of deterministic data rate under decoding delay constraints in single and multi-user SISO systems. Scheduling algorithms have been developed to satisfy users' deterministic queueing delay constraints in single and multi-user SISO systems [57, 58, 59, 60, 61, 62, 63].

However, wireless channels can change rapidly due to fading and interference. Such unstable natures of the wireless channel make it very difficult to provide wireless service to support deterministic QoS. Besides, some wireless applications, such as FTP data transmission, may require an average data rate under certain decoding delay constraint or average queueing delay constraint. This drives energy efficient scheduling algorithms design to satisfy an average data rate transmission under certain decoding delay constraint in single user SISO systems [64] and single user MIMO systems [65, 66, 67]. Energy efficient scheduling algorithms have been designed to support an average queueing delay requirement in single and multi-user SISO systems [68, 69, 70, 71, 72].

Furthermore, most wireless applications supported by the next generation networks, such as video conference or on-line games, can tolerate a small value of information loss probability or delay jitter. For this type of applications, deterministic QoS or average data rate and queueing delay are not accurate or specific enough to measure the QoS. This drives the development of statistical QoS guarantees to measure the quality of communications over wireless channel [73].

From the viewpoint of physical layer, under a predefined decoding delay constraint and transmission code rate, statistical QoS is measured as a violation probability that a certain transmission code rate is not achieved. Such violation probability is measured by information outage probability [32, 74, 75].

When considering from the viewpoint of data link layer, because of the instant changing characteristic of wireless channels, the amount of information being transmitted by wireless channels varies constantly. This results in an unpredictable departure time for the packets in the buffer, and subsequent random queueing delay. Therefore, a deterministic delay bound cannot be guaranteed. Consequently, a delay violation probability is defined as the probability that a predefined queueing delay bound is exceeded [76]. In this case, statistical QoS is measured as the maximum supported packet arrival rate under such delay violation probability constraint. The supported arrival rate under the queueing delay violation constraint is referred to as effective capacity (EC) [77].

Previous research works have not addressed the wireless resource allocation strategies design under statistical QoS in terms of information outage probability or EC constraint for delay constrained communications, which is essential for the next generation wireless networks to support users' QoS. In addition, the introduction of MIMO technology further increases the design complexity. In order to address these issues, we aim to develop a novel wireless resource allocation framework to support statistical QoS for delay constrained communications in MIMO systems in the thesis.

Specifically, our aim is to design wireless resource allocation algorithms to minimize the total transmit power while satisfying individual user's statistical QoS constraints in terms of information outage probability and EC constraint, respectively. We aim to design the algorithms in both single user and multi-user TDMA (or FDMA) systems, as TDMA and FDMA are essential multiple access methods for the next generation networks [8, 36, 37, 38]. In multi-user systems, we propose to jointly deal with scheduling of the users that access to the channel for each frame time (or available spectrum) and how much power is allocated when accessing to the channel.

We consider designing wireless resource allocation algorithms with statistical channel information (SCI) and instantaneous channel state information (CSI) at the transmitter side, respectively, in response to different channel feedback situations. The reasons are stated as follows. In a wireless communication system, a receiver can obtain instantaneous CSI through training sequence sent by a transmitter, and use such information for decoding. The knowledge of CSI at the transmitter is useful to improve the performance of wireless communications by reducing information outage probability and delay violation probability [78, 79]. However, it is difficult for the transmitter to trace the wireless channel. Firstly, feedback of instantaneous CSI will increase the overhead occupying more wireless resources, such as transmit power and bandwidth. This leads to transmission inefficiency. Secondly, small scale fading can change very fast, which makes it impossible for the transmitter to get the CSI feedback. Therefore, in some cases, it is more reasonable to send back SCI. However, this may result in some severe performance loss due to the lack of channel information. To address the above concerns, we consider CSI and SCI feedback in the thesis, respectively.

In the end, we aim to apply the developed wireless resource allocation framework to WiMAX, which is considered as a prominent standard for the next generation wireless networks supporting data services with high level of QoS [51, 80].

To summarize, the measurable objectives of the project are listed in detail as follows:

1. To develop power allocation strategies in delay constrained single user system with transmitter side SCI while satisfying user's QoS requirement in terms of information outage probability under certain decoding delay constraint.

2. To develop power allocation strategies in delay constrained single user system with transmitter side CSI to satisfy user's QoS requirement in terms of information outage probability under certain decoding delay constraint.

3. To further explore joint power and time slots scheduling algorithms in delay constrained multi-user systems with transmitter SCI while satisfying individual user's QoS requirement in terms of information outage probability under certain decoding delay constraint in TDMA systems.

4. To further explore joint power allocation and scheduling algorithms in delay constrained multi-user systems with transmitter CSI while satisfying individual user's QoS requirement in terms of information outage probability under certain decoding delay constraint in TDMA systems.

5. To devise joint power allocation and scheduling algorithms in delay constrained single and multi-user MIMO systems with individual user's QoS constraint, in terms of EC constraint. To evaluate the performance of the proposed scheduling technique with data of various QoS requirements under simulated wireless channels. Both TDMA and FDMA systems are to be considered respectively.

6. To further apply the proposed energy efficient scheduling algorithms to WiMAX systems, where the designed scheduling algorithms are applied with modifications for WiMAX scheduling.

1.2 Contributions and List of Publications

The contribution of the thesis is the consideration of statistical QoS requirements in terms of information outage probability and EC constraint for delay constrained communications when designing wireless resource allocation mechanisms. The QoS measurements for delay constrained communications used in the thesis are precise and effective to reflect those defined in many wireless standards, such as LTE and WiMAX. The consideration of MIMO in physical layer is another contribution of the research.

Specifically, the contribution of the thesis can be elaborated as follows.

Firstly, for a single user MIMO system, we have considered power minimization problem under QoS constraint, i.e., an information outage probability with decoding delay constraint, with transmitter side SCI. We have derived the closed-form of mean and variance of MIMO channel capacity, based on which we obtain the power allocation strategy to minimize transmit power while satisfying QoS requirement.

Secondly, for single user SISO and MIMO systems, we have solved a similar problem but considering transmitter side CSI instead. The problem can be solved optimally by using Dynamic Programming (DP), which searches the optimal solution numerically with great complexity. We have also proposed an efficient near-optimal solution involving much less complexity.

Thirdly, for multi-user TDMA MIMO systems with only SCI at the transmitter side, we have addressed the joint power and time slot allocation problem with individual user's QoS constraint in terms of an information outage probability. We propose a near-optimal method by jointly considering each user's power and time slot allocation.

Fourthly, for multi-user TDMA SISO systems with CSI at the transmitter side, we have solved the joint power and time slot allocation problem with individual user's information outage probability constraint.

Fifthly, we have investigated the power minimization problem for multi-user TDMA or FDMA MIMO systems with individual user's QoS constraint in terms of EC constraints. By jointly optimizing power allocation and the number of time-slots (or subcarriers) allocation for all of the users with the aid of SCI at the transmitter, we have shown that a significant power saving can be achieved.

Finally, we have applied the designed scheduling algorithm to IEEE 802.16 WiMAX system. We have modified the designed scheduling algorithms from previous chapters to meet the specifications of WiMAX standard. Through simulation, we have

observed significant advantages, such as power reduction and users' QoS satisfaction, by using the designed algorithm.

The contributions of the thesis have led to the following publications:

1. J. Chen and K. K. Wong, "Communication with causal CSI and controlled information outage", in *IEEE Transactions on Wireless Communications*, Vol. 8, No. 5, pp. 2221-2229, May 2009.

2. J. Chen and K. K. Wong, "Power minimization of central Wishart MIMO block-fading channels", in *IEEE Transactions on Communications*, Vol. 57, No. 4, pp. 899-905, April 2009.

3. K. K. Wong and J. Chen, "Time-division multiuser MIMO with statistical feedback", in *EURASIP Journal on Advances in Signal Processing: Special Issue on MIMO Transmission with Limited Feedback*, Vol. 2008.

4. K. K. Wong and J. Chen, "Near-optimal power allocation and multiuser scheduling with outage capacity constraints exploiting only channel statistics", in *IEEE Transactions on Wireless Communications*, Vol. 7, No. 3, pp. 812-818, March 2008.

5. J. Chen and K. K. Wong, "Improving energy efficiency for multiuser MIMO systems with effective capacity constraints", in proceedings of *IEEE Vehicular Technology Conference*, April 2009, Barcelona, Spain.

6. J. Chen and K. K. Wong, "An energy-saving QoS-based resource allocation for multiuser TDMA systems with causal CSI", in proceedings of *IEEE Global Communications Conference*, December 2008, New Orleans, USA.

7. J. Chen and K. K. Wong, "Communication with causal CSI and information outage", in proceedings of *International Conference on Mobile Computing and Ubiquitous Networking*, June 2008, Tokyo, Japan.

8. J. Chen and K. K. Wong, "Multiuser MIMO-TDMA with statistical feedback", in proceedings of *International Conference on Information, Communications and Signal Processing*, December 2007, Singapore.

9. J. Chen and K. K. Wong, "Channel-statistics assisted power minimization for central Wishart MIMO block-fading channels with an outage probability constraint", presented in *IEEE Sarnoff Symposium*, May 2007, Princeton, USA.

1.3 Thesis Outline

The thesis is organized as follows:

- Chapter 1, the current chapter, gives the motivation, objectives, contributions, list of publications and outline of the thesis.

- Chapter 2 introduces fundamental background information on wireless channel capacity towards understanding wireless resource allocation strategies in later chapters. Firstly, we review SISO and MIMO channel capacity results and their corresponding power allocation strategies with and without delay constraint. Then we introduce the concept of EC.

- In Chapter 3, we present some mathematical methodologies used in the thesis, including convex optimization theory, nonlinear optimization theory and DP.

- Chapter 4 deals with power minimization problem in single user transmission systems with SCI at the transmitter to achieve a given information outage probability under certain decoding delay constraint. Closed-form analytical expressions for the mean and variance of MIMO channel capacity are derived and Gaussian distribution is used to approximate the probability density

function (p.d.f.) of MIMO channel capacity. We use DIviding RECTangles (DIRECT) algorithm, which is a one-dimensional sampling method to numerically search for the optimal power allocation. SISO channel is considered as a special case of MIMO channel. Numerical results for different MIMO systems with different QoS requirements are compared and analyzed.

- Chapter 5 investigates power minimization problem in single user transmission system with transmitter CSI to achieve a certain information outage probability under certain decoding delay constraint. The problem is solved for both SISO and MIMO systems optimally with high complexity by adapting the power and rate allocation. An efficient near optimal solution at much less complexity is proposed which is asymptotically optimal for small values of information outage probability. A lower bound and a simpler method are developed for comparisons. Numerical results under different simulation settings are compared and discussed.

- In Chapter 6, we extend the resource allocation problem in Chapter 4 to multi-user TDMA MIMO transmission systems. Power minimization and scheduling for a time-division multi-user MIMO system is studied with transmitter SCI under individual user's information outage probability constraint. We apply research results in Chapter 4 to determine the minimum power for attaining a given outage probability constraint under certain decoding delay constraint. On the other hand, we propose an optimization approach to find the suboptimal number of blocks allocated to the users. The two main techniques are combined to obtain a joint solution for both power and time allocation for the users. Simulation results demonstrate that the proposed method achieves near-optimal performance.

- In Chapter 7, resource allocation problem in Chapter 5 is extended to multi-user transmission systems. We devise an energy efficient scheduling algorithm with individual user's QoS constraint in terms of information outage

probability requirements for TDMA SISO systems, by exploiting transmitter CSI. First, a simple closed-form power allocation solution one-step limited look-ahead policy (OLLP) is developed to meet a given outage probability constraint for a single user channel. Then we construct an optimization problem which enables a joint consideration of power consumption and users' QoS constraints to find a suboptimal multi-user time-sharing solution. This time-sharing solution in conjunction with the OLLP forms the joint solution for the multi-user system, which is shown by simulation results to be near-optimal and yields a significant energy saving as compared to those without time-sharing optimization or OLLP power allocation.

- Chapter 8 considers the power minimization problem for multi-user TDMA or FDMA MIMO systems with individual user's QoS constraint in terms of an EC function with the aid of SCI at the transmitter side. We obtain a closed-form expression for the EC of a single-user MIMO system, which is then applied to solve the power minimization problem in multi-user system. In the following, we show that the optimal solutions for multi-user TDMA or FDMA MIMO systems can be obtained by DP. Then, heuristic algorithms achieving near-optimal results at reduced complexity are presented. Simulation results demonstrate the near-optimality of the heuristics approach.

- Chapter 9 applies the designed algorithm to IEEE 802.16 WiMAX for uplink scheduling. The scheduling algorithms are modified to meet the specifications of WiMAX. Simulations are presented under different settings to verify the significant advantages, such as power reduction and users' QoS satisfaction, by using the designed algorithm.

- Finally, Chapter 10 concludes the thesis and discusses some future research directions.

Chapter 2

Wireless Channel Capacity

This chapter provides an essential background of wireless channel capacity towards understanding and solving wireless resource allocation problems in the subsequent chapters. Firstly, wireless channel mathematical models, capacity results and their corresponding power allocation strategy for both SISO and MIMO systems are presented with and without decoding delay constraint, in Sections 2.1 and 2.2. In Section 2.3, we introduce the concept of EC. Finally, Section 2.4 summaries this chapter.

2.1 SISO Channel and its Capacity

2.1.1 SISO Channel Model

A discrete-time SISO channel transmission system sending an input message s from the transmitter to the receiver is represented in Fig. 2.1 [17]. The message is encoded into the codeword x, which is transmitted over time varying channel as x_k at the kth coherence period of T_c symbols. The channel power gain g_k corresponding to the fading process changes independent and identical distributed

(i.i.d.) over the transmission and remains constant for the kth time period. This is defined as *block fading* (BF) channel model [3]. Blocks can be considered as separated in time or separated in frequency, such as multicarrier systems [64]. It is assumed g_k is known to the receiver and a feedback link may exist allowing receiver to feedback instantaneous CSI to the transmitter.

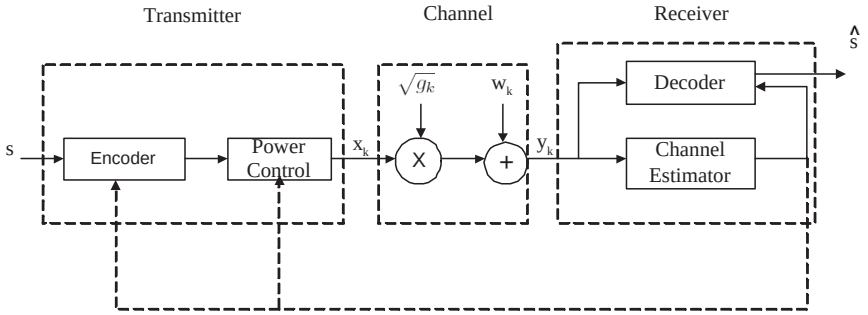

Figure 2.1: An illustration of a SISO channel model.

The received signal y_k of the above SISO channel signal transmission is represented by

$$y_k = \sqrt{g_k} x_k + w_k, \tag{2.1}$$

in which w_k is i.i.d. complex Gaussian noise with normal distribution denoted by $\mathcal{N}(0, N_0)$, where N_0 is the variance or noise power. N_0 is assumed to be 1 for normalization and it is assumed that the expectation over all channel power gain realizations $\mathtt{E}[g_k] = 1$ for normalization. In addition, it is assumed that transmit power p for block k is p_k. Hence, the instantaneous receive SNR for block k is $\gamma_k = g_k p_k$.

2.1.2 Channel Capacity with No Delay Constraint

Without any delay constraint, transmission codeword can be assumed long enough to experience all fading channel states. In such circumstances, channel capacity is measured by ergodic capacity, which is the average of instantaneous channel

capacity over all possible channel realizations. According to Shannon's communication theory, instantaneous channel capacity of a fading channel for time instance k is $C_k = \log_2(1+\gamma_k)$. Let $f(\gamma) = \mathsf{Pr}\{\gamma_i = \gamma\}$ denote the p.d.f. of receive SNR, the ergodic capacity is defined as [17]

$$C \triangleq \mathtt{E}_\gamma \left[\log_2(1+\gamma)\right] = \int_0^\infty \log_2(1+\gamma) f(\gamma) d\gamma \quad \text{in bps/Hz.} \tag{2.2}$$

Under an average transmit power constraint $\mathtt{E}[p] \leq P$, SISO channel capacity is given considering without and with CSI at the transmitter side, respectively.

With no CSI at the transmitter, constant power P is allocated for each block to achieve channel capacity C_{CP}, which is obtained as [17]

$$C_{CP}(P) = \mathtt{E}_g[\log_2(1+gP)] = \int_0^\infty \log_2(1+gP) f(g) dg, \tag{2.3}$$

in which $f(g)$ denotes the p.d.f. of g.

When perfect CSI is available at the transmitter, adaptive power and code rate are allocated for each block and the capacity C_{AP} is obtained by maximizing over the power allocation function $p = p(g)$ with the constraint $\mathtt{E}[p] \leq P$ and is expressed as

$$C_{AP}(P) = \max_{E[p] \leq P} \mathtt{E}_g \left[\log_2(1+gp)\right]. \tag{2.4}$$

The optimal power allocation is given as [3, 17]

$$p(g) = \left[\frac{1}{\xi} - \frac{1}{g}\right]^+, \tag{2.5}$$

where $[z]^+ = \max(z, 0)$. This is referred to the *waterfilling power allocation* strategy [3], in which *waterlevel* ξ is obtained numerically from the power constraint

$$\mathtt{E}_g[p] = \int_\xi^\infty \left(\frac{1}{\xi} - \frac{1}{g}\right) f(g) dg = P. \tag{2.6}$$

Fig. 2.2 illustrates the waterfilling power allocation strategy. The value of $\frac{1}{g}$ is plotted as a function of block, and it is regarded as the bottom of a vessel. If P units of water per block are filled into the vessel and $\frac{1}{\xi}$ is the water surface, the depth of the water at each block equals to the power allocated to that block. No power is allocated when the bottom of the vessel is above the surface, since the channel is too poor to transmit information. Generally, the transmitter allocates more power to the stronger channels and less or zero power to weaker channels.

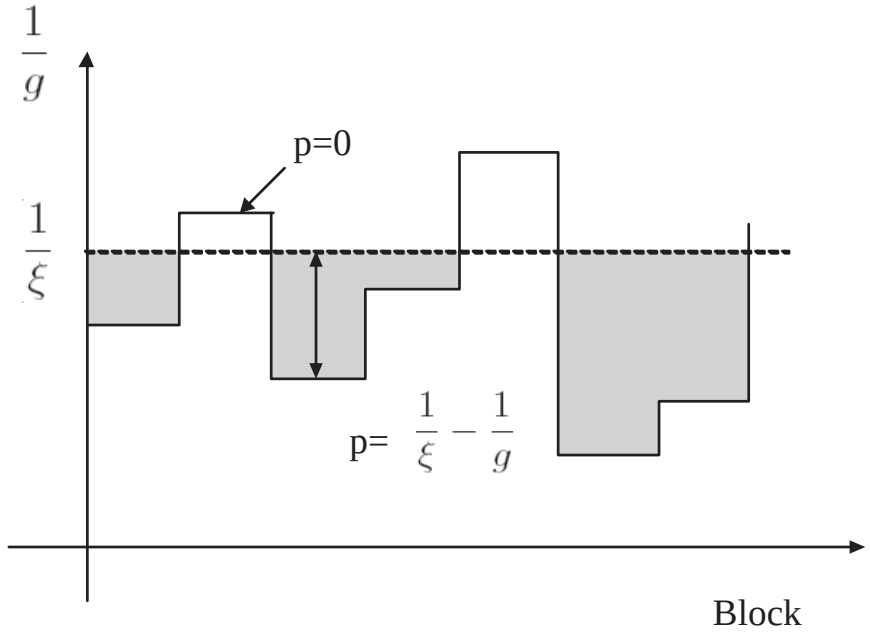

Figure 2.2: An illustration of waterfilling power allocation strategy.

As a result, the expression for $C_{AP}(P)$ is derived as [17]

$$C_{AP}(P) = \int_{\xi}^{\infty} \log_2(g/\xi) f(g) dg. \tag{2.7}$$

For Rayleigh fading channel, $f(g) = e^{-g}$, the channel capacity performance of constant power and waterfilling power allocation are compared in Fig. 2.3. The ergodic capacity under the waterfilling power allocation strategy outperforms the constant power allocation. For higher transmit power, the ergodic capacity results under two settings are closer to each other.

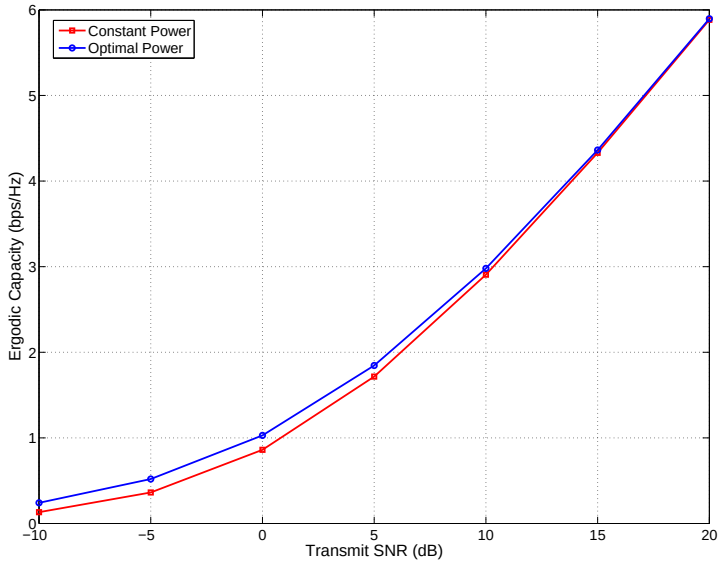

Figure 2.3: Ergodic capacity of SISO channel with no delay constraint.

2.1.3 Channel Capacity with Delay Constraint

When delay constraint is considered, channel capacity is investigated in [64, 81]. Suppose a codeword spans over K blocks. For a fixed T_c, the number of blocks K is related to the system coding and decoding complexity [82]. Therefore, K can be considered as the overall coding/decoding delay constraint.

For delay constrained communications, ergodic capacity is not capable of describing the channel performance. Instead, *information outage probability* is proposed as a performance indicator of wireless channel [64]. First, *instantaneous mutual information* (IMI) of a K-BF channel measuring the instantaneous channel capacity is defined as [64]

$$\mathfrak{I}_K(\boldsymbol{p}, \boldsymbol{g}) \triangleq \frac{1}{K} \sum_{k=1}^{K} \log_2(1 + g_k p_k), \tag{2.8}$$

where $\boldsymbol{p} \triangleq (p_1, p_2, \ldots, p_K)$ and $\boldsymbol{g} \triangleq (g_1, g_2, \ldots, g_K)$. Since $\mathfrak{I}_K(\boldsymbol{p}, \boldsymbol{g})$ is a random variable, the transmission code rate R_0 may or may not be supported by IMI. By

defining short-term power constraint as

$$\frac{1}{K}\sum_{k=1}^{K} p_k \leq P, \tag{2.9}$$

and long-term power constraint as

$$\mathtt{E}\left[\frac{1}{K}\sum_{k=1}^{K} p_k\right] \leq P, \tag{2.10}$$

the *information outage probability* evaluated at R_0 with power constraint P is denoted by [64]

$$\mathcal{P}_{out}(R_0, P) \triangleq \mathsf{Pr}\{\mathcal{I}_K(\boldsymbol{p}, \boldsymbol{g}) < R_0\}. \tag{2.11}$$

The optimization for delay constrained communications can be formulated as

$$\begin{cases} \min\limits_{\boldsymbol{p}} & \mathsf{Pr}\{\mathcal{I}_K(\boldsymbol{p}, \boldsymbol{g}) < R_0\} \\ \text{s.t.} & \frac{1}{K}\sum\limits_{k=1}^{K} p_k \leq P, \\ \text{or} & \mathtt{E}\left[\frac{1}{K}\sum\limits_{k=1}^{K} p_k\right] \leq P. \end{cases} \tag{2.12}$$

The above problem has been solved considering transmitter side CSI only. When CSI of the whole K blocks is available at the transmitter prior to transmission (acausal CSI), such as in multicarrier systems, the optimal power allocation strategies have been solved in [64].

Under short-term power constraint, achieving maximum IMI with the available power P over K blocks for each transmission results in the minimum outage.

Therefore, the optimization problem is equivalent to [64]

$$\begin{cases} \max\limits_{\boldsymbol{p}} & \mathfrak{I}_K(\boldsymbol{p},\boldsymbol{g}) \\ \text{s.t.} & \dfrac{1}{K}\sum\limits_{k=1}^{K} p_k \le P. \end{cases} \tag{2.13}$$

The optimal power allocation for block k under short-term power constraint p_k^{st} is given by waterfilling strategy as

$$p_k^{st} = \left[\frac{1}{\xi^{st}} - \frac{1}{g_k}\right]^+, \tag{2.14}$$

in which the waterlevel ξ^{st} can be numerically found by solving the short-term power constraint equation

$$\frac{1}{K}\sum_{k=1}^{K}\left[\frac{1}{\xi^{st}} - \frac{1}{g_k}\right]^+ = P. \tag{2.15}$$

Under long-term power constraint, the problem is considered in two situations. When the channel is bad, or in outage, the transmitter should allocate zero power. On the other hand, the transmitter should allocate the least amount of power to meet the exact target code rate R_0. Consequently, the optimization problem is equivelant to [64]

$$\begin{cases} \min\limits_{\boldsymbol{p}} & \dfrac{1}{K}\sum\limits_{k=1}^{K} p_k \\ \text{s.t.} & \mathfrak{I}_K(\boldsymbol{p},\boldsymbol{g}) = R_0. \end{cases} \tag{2.16}$$

Similar to (2.13) and (2.14), the optimal power allocation for block k under long-term power constraint p_k^{lt} is given by waterfilling strategy as

$$p_k^{lt} = \left[\frac{1}{\xi^{lt}} - \frac{1}{g_k}\right]^+, \tag{2.17}$$

substituting (2.17) into (2.8), the waterlevel ξ^{lt} can be numerically found by solving the rate constraint

$$\frac{1}{K}\sum_{k=1}^{K}\log_2\left(\frac{g_k}{\xi^{lt}}\right) = R_0. \tag{2.18}$$

Determining whether the transmission is in outage needs further calculation. In [64], the authors give a method to define the boundary P^B. Denoting the required sum power for each transmission as $p_{\mathsf{sum}} = \sum_{k=1}^{K} p_k^{lt}$, the power transmission boundary P^B is calculated by solving the following equation satisfying long-term power constraint

$$\int_0^{P^B} p_{\mathsf{sum}}(\boldsymbol{g})d\mathcal{F}(\boldsymbol{g}) = P, \tag{2.19}$$

where $\mathcal{F}(\boldsymbol{g})$ is the cumulative distribution function (c.d.f.) of channel vector $\boldsymbol{g}$. For each transmission, if the required sum power for achieving code rate $p_{\mathsf{sum}} > P^B$, the transmission is in outage and no power is allocated. The above equation involves high complexity and hence, can only be solved numerically when block number $K > 2$ [64].

When CSI for the current block k and previous blocks $1, 2, \cdots, k-1$ are available at the transmitter (causal CSI), the transmitter will adapt power and rate transmission for each block based on the current and previous CSI and the statistics of future channel [81]. The optimal solution to such problem can be solved numerically by applying DP, which deals with optimization problems where decisions are made in stages with the objective to minimize a certain cost [83].

2.2 MIMO Channel and its Capacity

2.2.1 Channel Model

A discrete-time MIMO channel transmission system sending an input message s from the transmitter to the receiver is represented in Fig. 2.4. The MIMO transmission system has n_t transmit antennas and n_r receive antennas. It is assumed that $n_t \geq n_r$ and the system performance can be analyzed in a similar way when $n_t < n_r$. Channel matrix for the kth time period (or block) $\mathbf{H}_k$ is an $n_r \times n_t$ complex matrix $\mathbf{H}_k = \left[h_{i,j}^{(k)}\right] \in \mathbb{C}^{n_r \times n_t}$, in which $h_{i,j}^{(k)}$ denotes the channel gain from the j^{th} transmit antenna to the i^{th} receive antenna for block k. As a normalization, it is assumed that $\mathtt{E}[|h_{i,j}|^2] = 1$. In addition, $\mathbf{H}_k$ is known to the receiver and a feedback link may exist allowing receiver to feedback instantaneous CSI to the transmitter.

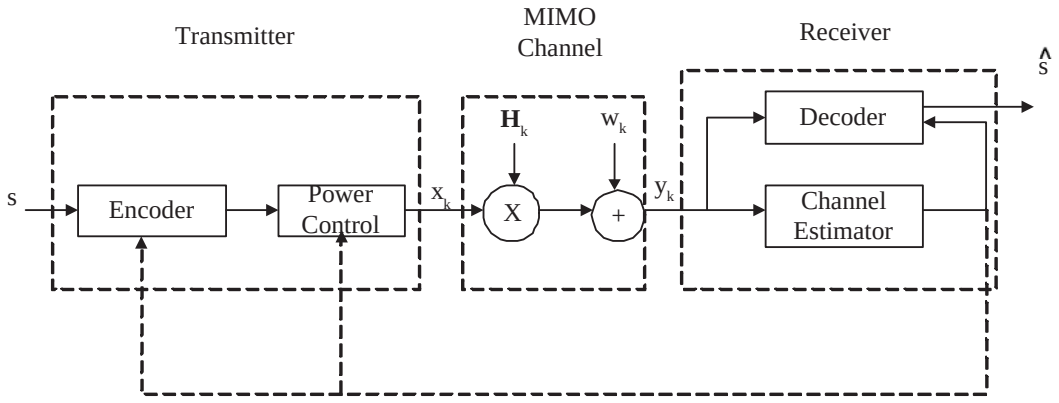

Figure 2.4: An illustration of a MIMO channel model.

The above MIMO channel signal transmission at block k is represented by

$$\boldsymbol{y}_k = \mathbf{H}_k \boldsymbol{x}_k + \boldsymbol{w}_k, \tag{2.20}$$

in which $\boldsymbol{x}_k$ is an $n_t \times 1$ vector denoting the transmit symbols. $\boldsymbol{w}_k$ is an $n_r \times 1$ noise vector and the entries of each element is i.i.d. complex Gaussian distribution

$\mathcal{N}(0, N_0)$. N_0 is assumed to be 1 for normalization. In addition, it is assumed that transmit power for block k $tr(\boldsymbol{x}_k \boldsymbol{x}_k^\dagger)$ is p_k, where $\mathbf{Z}^\dagger$ denotes the conjugate transpose of the matrix $\mathbf{Z}$ and $tr(\mathbf{Z})$ denotes the trace of the matrix $\mathbf{Z}$.

2.2.2 Channel Capacity with No Delay Constraint

With no delay constraint, ergodic capacity is used to measure the performance of MIMO system [33, 39, 40, 41]. Under a transmit power constraint $tr(\boldsymbol{x}_k \boldsymbol{x}_k^\dagger) \leq P$, the capacity of MIMO channels is derived without and with transmitter side CSI respectively.

With no CSI at the transmitter, equal power P/n_t is allocated for each antenna and the channel capacity is given as [33]

$$C_{CP}(P) = \mathtt{E}_{\mathbf{H}} \left[\log_2 \left| \mathbf{I}_{n_r} + \frac{P}{n_t} \mathbf{H}\mathbf{H}^\dagger \right| \right] \quad \text{in bps/Hz}, \tag{2.21}$$

where $|\mathbf{Z}|$ denotes the determinant of matrix $\mathbf{Z}$.

The matrix $\mathbf{H}$ has a *singular value decomposition* (SVD) as

$$\mathbf{H} = \mathbf{U}\mathbf{D}\mathbf{V}^\dagger, \tag{2.22}$$

where $\mathbf{U} \in \mathbb{C}^{n_r \times n_t}$ and $\mathbf{V} \in \mathbb{C}^{\mathbf{n_t} \times \mathbf{n_r}}$ are unitary matrix and $\mathbf{D} \in \mathbb{R}^{n_r \times n_t}$ is a real diagonal matrix. The real diagonal elements $d_1 \geq d_2 \cdots \geq d_{n_r}$ are the ordered *singular values* of the matrix $\mathbf{H}$. Since

$$\mathbf{H}\mathbf{H}^\dagger = \mathbf{U}\mathbf{D}\mathbf{D}^T\mathbf{U}^\dagger, \tag{2.23}$$

where $\mathbf{Z}^T$ is the transpose of the matrix $\mathbf{Z}$, the eigenvalues of the matrix $\mathbf{H}\mathbf{H}^\dagger$ known as $(\lambda_1, \lambda_2, \cdots, \lambda_{n_r})$ are the squared singular values $(d_1^2, d_2^2, \cdots, d_{n_r}^2)$.

Therefore, the capacity can be further evaluated as [33]

$$C_{CP}(P) = \mathtt{E}_{\boldsymbol{\lambda}} \left[\sum_{i=1}^{n_r} \log_2 \left(1 + \frac{P}{n_t} \lambda_i \right) \right]. \tag{2.24}$$

When perfect CSI is available at the transmitter, adaptive power p_i is allocated along each *eigenchannel*, which corresponds to each λ_i of the channel, to support each data stream transmission. The capacity C_{AP} is obtained by maximizing over power allocation constraint $\sum_{i=1}^{n_r} p_i \leq P$ given as [3, 33]

$$C_{AP}(P) = \max_{\sum_{i=1}^{n_r} p_i \leq P} \mathtt{E}_{\boldsymbol{\lambda}} \left[\sum_{i=1}^{n_r} \log_2 \left(1 + p_i \lambda_i \right) \right]. \tag{2.25}$$

The above power allocation solution is given by waterfilling solution as [3, 33]

$$p_i = \left[\frac{1}{\delta} - \frac{1}{\lambda_i} \right]^+, \tag{2.26}$$

where waterlevel δ is obtained by solving the power constraint equation

$$\frac{1}{K} \sum_{i=1}^{n_r} \left[\frac{1}{\delta} - \frac{1}{\lambda_i} \right]^+ = P. \tag{2.27}$$

2.2.3 Channel Capacity with Delay Constraint

When delay is considered, information outage probability is used as a performance indicator for MIMO channels [65].

For a K-BF channel, let $\boldsymbol{\lambda}_k = (\lambda_{k,1}, \cdots, \lambda_{k,n_r})$ be the eigenvalues for block k, and $\boldsymbol{p}_k = (p_{k,1}, \cdots, p_{k,n_r})$ be the power allocated for block k along eigenchannels, the IMI of the K-block MIMO fading channel is defined as [65]

$$\mathcal{I}_K(\mathbf{P}, \boldsymbol{\Lambda}) \triangleq \frac{1}{K} \sum_{k=1}^{K} \sum_{i=1}^{n_r} \log_2(1 + \lambda_{k,i} p_{k,i}), \tag{2.28}$$

where $\mathbf{P} \triangleq (\boldsymbol{p}_1, \boldsymbol{p}_2, \cdots, \boldsymbol{p}_K)$ and $\boldsymbol{\Lambda} \triangleq (\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2, \cdots, \boldsymbol{\lambda}_K)$.

Defining short-term power constraint as

$$\frac{1}{K}\sum_{k=1}^{K}\sum_{i=1}^{n_r} p_{k,i} \leq P, \tag{2.29}$$

and long-term power constraint as

$$\mathtt{E}\left[\frac{1}{K}\sum_{k=1}^{K}\sum_{i=1}^{n_r} p_{k,i}\right] \leq P, \tag{2.30}$$

information outage probability evaluated at a target code rate R_0 with short-term or long-term power constraint P is defined by [65]

$$\mathcal{P}_{out}(R_0, P) \triangleq \mathsf{Pr}\{\mathcal{I}_K(\mathbf{P}, \boldsymbol{\Lambda}) < R_0\}. \tag{2.31}$$

Accordingly, the problem for delay constrained communications is formulated as [65]

$$\begin{cases} \min\limits_{\mathbf{P}} & \mathsf{Pr}\{\mathcal{I}_K(\mathbf{P}, \boldsymbol{\Lambda}) < R_0\} \\ \text{s.t.} & \frac{1}{K}\sum\limits_{k=1}^{K}\sum\limits_{i=1}^{n_r} p_{k,i} \leq P, \\ \text{or} & \mathtt{E}\left[\frac{1}{K}\sum\limits_{k=1}^{K}\sum\limits_{i=1}^{n_r} p_{k,i}\right] \leq P. \end{cases} \tag{2.32}$$

The above problem has been solved considering transmitter side acausal CSI only [65]. Under short-term power constraint, achieving maximum IMI with the available power P over K blocks for each transmission leads to the minimum outage. Therefore, the optimization problem is equivalent to [65]

$$\begin{cases} \max\limits_{\mathbf{P}} & \mathcal{I}_K(\mathbf{P}, \boldsymbol{\Lambda}) \\ \text{s.t.} & \frac{1}{K}\sum\limits_{k=1}^{K}\sum\limits_{i=1}^{n_r} p_{k,i} \leq P. \end{cases} \tag{2.33}$$

The optimal power allocation for the kth block and ith eigenchannel under short-term power constraint denoted as $p_{k,i}^{st}$ is given by waterfilling strategy as

$$p_{k,i}^{st} = \left[\frac{1}{\delta^{st}} - \frac{1}{\lambda_{k,i}}\right]^{+}, \tag{2.34}$$

where waterlevel δ^{st} is obtained by solving the short-term power constraint equation

$$\frac{1}{K}\sum_{k=1}^{K}\sum_{i=1}^{n_r}\left[\frac{1}{\delta^{st}} - \frac{1}{\lambda_{k,i}}\right]^{+} = P. \tag{2.35}$$

Under long-term power constraint, the problem is considered in two situations. When the channel is bad, or in outage, the transmitter should allocate zero power. On the other hand, the transmitter should allocate the least amount of power to meet the exact target code rate R_0. Consequently, the optimization problem is equivalent to [65]

$$\begin{cases} \min\limits_{\mathbf{P}} & \dfrac{1}{K}\displaystyle\sum_{k=1}^{K}\sum_{i=1}^{n_r} p_{k,i} \\ \text{s.t.} & \mathfrak{I}_K(\mathbf{P}, \mathbf{\Lambda}) = R_0. \end{cases} \tag{2.36}$$

The optimal power allocation for the kth block and ith eigenchannel under long-term power constraint denoted as $p_{k,i}^{lt}$ is given by waterfilling strategy as

$$p_{k,i}^{lt} = \left[\frac{1}{\delta^{lt}} - \frac{1}{\lambda_{k,i}}\right]^{+}, \tag{2.37}$$

substituting (2.37) into (2.28), the waterlevel δ^{lt} can be obtained by solving the transmission rate constraint equation

$$\frac{1}{K}\sum_{k=1}^{K}\sum_{i=1}^{n_r}\log_2\left(\frac{\lambda_{k,i}}{\delta^{lt}}\right) = R_0 \tag{2.38}$$

Determing whether the transmission is in outage needs further calculation. In [65], the authors present a method to define the boundary. Denoting the required

sum power for each transmission for achieving R_0 as $p_{\mathsf{sum}} = \sum_{k=1}^{K} \sum_{i=1}^{n_r} p_{k,i}^{lt}$, the power transmission boundary P^B is calculated by solving the following equation satisfying long-term power constraint

$$\int_0^{P^B} p_{\mathsf{sum}}(\mathbf{\Lambda}) d\mathcal{F}(\mathbf{\Lambda}) = P, \tag{2.39}$$

where $\mathcal{F}(\mathbf{\Lambda})$ is the c.d.f. of channel eigenvalue matrix $\mathbf{\Lambda}$. For each transmission, if the required sum power $p_{\mathsf{sum}} > P^B$, the transmission is in outage and no power is allocated. The above equation involves high complexity and hence, can only be solved numerically [65].

2.3 EC

The concept of EC was first proposed in [77]. Different from the capacity results considered in Sections 2.1 and 2.2, EC measures wireless capacity from data link layer instead of physical layer. As shown in Fig.2.5, data are transmitted to data link layer with a constant rate (source rate) and stored in the buffer before transmitted over wireless channel. Due to the instant changing characteristic of wireless channel, the amount of information being transmitted by wireless channel (service rate) varies constantly. This results in an unpredictable queueing delay for the packets in the buffer.

Consequently, a delay violation probability is defined as the probability that a predefined queueing delay bound is exceeded. Let $D(t)$ be the queueing delay at time t. The delay bound violation probability, or the probability of $D(t)$ exceeding a delay bound D_{max} is expressed as

$$\mathsf{Pr}\{D(t) \geq D_{\mathsf{max}}\} \approx e^{-\theta D_{\mathsf{max}}}, \tag{2.40}$$

where θ is the delay exponent, indicating the exponential decay rate of delay. Apparently, a larger θ will result in a stringent delay constraint, whereas a smaller θ indicates a loose delay constraint.

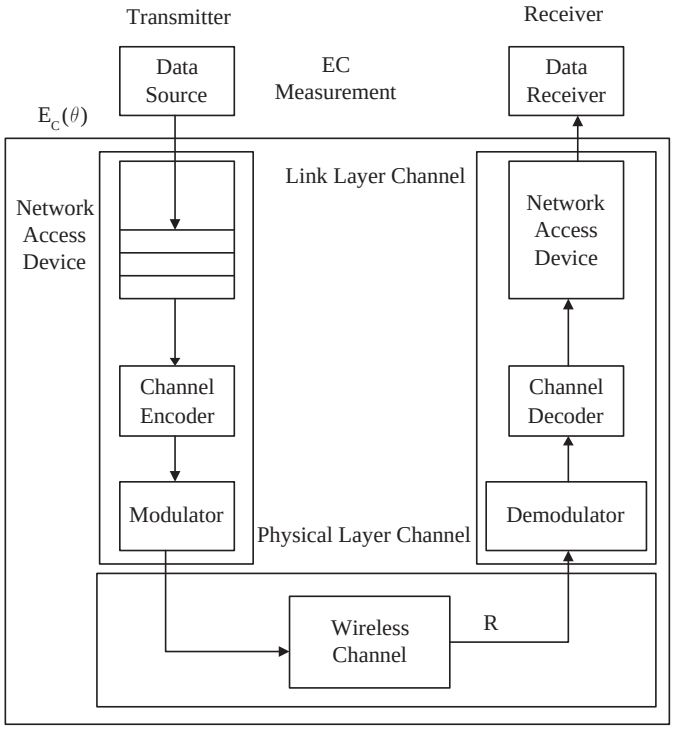

Figure 2.5: An illustration of EC model.

Under the above delay bound violation probability constraint, EC measures the maximum achievable source rate that the wireless channel can support. With an i.i.d. BF channel model, the EC under the above delay constraint $E_C(\theta)$ has been obtained as [77, 79]

$$E_C(\theta) = -\frac{1}{\theta}\log_2 \mathtt{E}\left[e^{-\theta R}\right], \tag{2.41}$$

where R is the instantaneous channel capacity. An example of EC for SISO channel with constant power P allocated for each block and channel power gain g is given by [77, 79]

$$E_C(\theta) = -\frac{1}{\theta}\log_2 \mathtt{E}\left[e^{-\theta \log_2(1+Pg)}\right] = -\frac{1}{\theta}\log_2 \int_0^\infty e^{-\theta \log_2(1+Pg)} dg. \tag{2.42}$$

2.4 Summary

In this chapter, we have presented background information of wireless channel, wireless capacity and EC for understanding the research work in the following chapters. From the viewpoint of physical layer, without decoding delay constraint, ergodic capacity has been used to measure the performance of wireless channel. When delay is considered, the concept of information outage probability has been introduced to measure wireless channel performance. From the viewpoint of data link layer, EC is used to measure the performance of wireless channel under statistical queueing delay constraint. In the following chapters, we will consider wireless resource allocation design under information outage probability constraint and EC constraint respectively.

Chapter 3

Mathematical Preliminaries

In this chapter, we present mathematical theories used in the thesis, including convex optimization theory in Section 3.1, nonlinear optimization theory in Section 3.2 and DP and optimal control theory in Section 3.3.

3.1 Convex Optimization Theory

In some cases, after some mathematical transformation, the optimization problems can be possibly expressed in convex form. For these types of problems, there is a well developed theory to solve the problem optimally and efficiently.

3.1.1 Convex Optimization Problem

A function f: $\mathbb{R}^n \to \mathbb{R}$ is defined as a *convex* function when the domain of f denoted by **dom** f is convex set (which means for all n-dimensional vectors $\boldsymbol{x}$, $\boldsymbol{y} \in$ **dom** f and $\alpha \in [0,1]$, $\alpha\boldsymbol{x} + (1-\alpha)\boldsymbol{y} \in$ **dom** f) and for all $\boldsymbol{x}, \boldsymbol{y} \in$ **dom** f and $\theta \in [0,1]$, the function f satisfies

$$f(\theta\boldsymbol{x} + (1-\theta)\boldsymbol{y}) \leq \theta f(\boldsymbol{x}) + (1-\theta)f(\boldsymbol{y}). \tag{3.1}$$

A convex optimization problem is one of the form [84]

$$\begin{aligned} \text{minimize} \quad & f_0(\boldsymbol{x}) \\ \text{subject to} \quad & f_i(\boldsymbol{x}) \leq 0 \quad i = 1, \cdots, m, \\ & h_i(\boldsymbol{x}) = 0 \quad i = 1, \cdots, p, \end{aligned} \tag{3.2}$$

where the functions $f_0, \cdots, f_m : \mathbb{R}^n \to \mathbb{R}$ are convex and $h_1, \cdots, h_p : \mathbb{R}^n \to \mathbb{R}$ are linear. f_0 is the *cost function* or *objective function*. The inequalities $f_i(\boldsymbol{x}) \leq 0$ are *inequalities constraints* and equalities $h_i(\boldsymbol{x}) = 0$ are *equalities constraints*.

3.1.2 Convex Functions

3.1.2.1 Basic Properties

The basic properties to determine whether function f is convex is **first-order condition** stated as follows [84, 85]:

Suppose f is differentiable, i.e., its gradient ∇f exists at each point in **dom** f, then f is convex if and only if **dom** f is convex and

$$f(\boldsymbol{y}) \geq f(\boldsymbol{x}) + \nabla f(\boldsymbol{x})^T(\boldsymbol{y} - \boldsymbol{x}) \tag{3.3}$$

holds for all $\boldsymbol{x}$, $\boldsymbol{y} \in$ **dom** f.

Moreover, a more commonly used condition to distinguish a convex function is **second-order condition**, which is presented as follows [84, 85]:

Assume that f is twice differentiable, i.e., its Hessian or second derivative $\nabla^2 f$ exists at each point in **dom** f. Then f is convex if and only if **dom** f is convex and its Hessian is positive semidefinite for all $\boldsymbol{x} \in$ **dom** f, or

$$\nabla^2 f(\boldsymbol{x}) \succeq 0. \tag{3.4}$$

In addition, f is concave if $-f$ is convex.

3.1.2.2 Operations that Preserve Convexity

In this section, operations to preserve convexity of convex functions are introduced. This can be used to analyze the convexity of a function [84].

- **Nonnegative weighted sums**. A nonnegative weighted sum of convex functions,

$$f = w_1 f_1 + \cdots + w_m f_m, \tag{3.5}$$

is convex. The property extends to infinite sums and integrals. An example is given as follows. If $f(\boldsymbol{x}, \boldsymbol{y})$ is convex in $\boldsymbol{x}$ for each $\boldsymbol{y} \in \mathcal{A}$, and $w(\boldsymbol{y}) > 0$ for each $\boldsymbol{y} \in \mathcal{A}$, the function $g(\boldsymbol{x})$ defined as

$$g(\boldsymbol{x}) = \int_{\mathcal{A}} w(\boldsymbol{y}) f(\boldsymbol{x}, \boldsymbol{y}) d\boldsymbol{y} \tag{3.6}$$

is convex in $\boldsymbol{x}$.

- **Composition with an affine mapping**. Suppose f: $\mathbb{R}^n \to \mathbb{R}$, $\mathbf{A} \in \mathbb{R}^{n \times m}$, and $b \in \mathbb{R}^n$. Define g: $\mathbb{R}^m \to \mathbb{R}$ by

$$g(\boldsymbol{x}) = f(\mathbf{A}\boldsymbol{x} + \boldsymbol{b}), \tag{3.7}$$

with $\mathbf{dom}\ g = \{\boldsymbol{x} | \mathbf{A}\boldsymbol{x} + \boldsymbol{b} \in \mathbf{dom} f\}$. If f is convex, so is g.

- **Composition**. Define two functions h: $\mathbb{R}^k \to \mathbb{R}$ and g: $\mathbb{R}^n \to \mathbb{R}^k$. The composition function $f = h \circ g$: $\mathbb{R}^n \to \mathbb{R}$ is defined by

$$f(\boldsymbol{x}) = h(g(\boldsymbol{x})), \qquad \mathbf{dom}\ f = \{\boldsymbol{x} \in \mathbf{dom}\ g | g(\boldsymbol{x}) \in \mathbf{dom}\ h\}. \tag{3.8}$$

Consider a scalar composition, where $k = n = 1$, assume h and g are twice differentiable, the second derivative of the composition function denoted as

$f''(x)$ is obtained as

$$f''(x) = h''(g(x))g'(x)^2 + h'(g(x))g''(x), \tag{3.9}$$

where l' denotes the first derivative of the function l.

According to the second-order condition, f is convex if $f'' \geq 0$. Therefore, the following rules can be used to distinguish the convexity of f:

f is convex if h is convex and nondecreasing, and g is convex,

f is convex if h is convex and nonincreasing, and g is concave,

f is concave if h is concave and nondecreasing, and g is concave,

f is concave if h is concave and nonincreasing, and g is convex.

3.1.3 Solutions to Convex Optimization

3.1.3.1 Lagrangian Duality Function and KKT Condition

For some cases, convex problems can be analytically solved using Lagrangian duality and Karush-Kuhn-Tucker (KKT) condition [84]. The basic idea of Lagrangian duality is to take the constraints in (3.2) into account by augmenting the objective function with a weighted sum of the constraint functions. We define the *Lagrangian* associated with the problem (3.2) as

$$L(\boldsymbol{x}, \boldsymbol{\phi}, \boldsymbol{\upsilon}) = f_0(\boldsymbol{x}) + \sum_{i=1}^{m} \phi_i f_i(\boldsymbol{x}) + \sum_{i=1}^{p} \upsilon_i h_i(\boldsymbol{x}), \tag{3.10}$$

in which ϕ_i is referred as *Lagrangian multiplier* associated with inequality constraints $f_i(\boldsymbol{x}) \leq 0$ and υ_i is the Lagrangian multiplier associated with equality constraints $h_i(\boldsymbol{x}) = 0$. The vectors $\boldsymbol{\phi} \in \mathbb{R}^m$ and $\boldsymbol{\upsilon} \in \mathbb{R}^p$ are called *Lagrangian*

multiplier vectors. The *Lagrangian dual function* $g(\boldsymbol{\phi}, \boldsymbol{\upsilon})$ is defined as the minimum value of the Lagrangian over $\boldsymbol{x}$, i.e.,

$$g(\boldsymbol{\phi}, \boldsymbol{\upsilon}) \triangleq \inf_{\boldsymbol{x}} L(\boldsymbol{x}, \boldsymbol{\phi}, \boldsymbol{\upsilon}). \tag{3.11}$$

The optimal solution for (3.11) $(\boldsymbol{x}^*, \boldsymbol{\phi}^*, \boldsymbol{\upsilon}^*)$ is also the optimal solution for the dual problem (3.2). The solution satisfies KKT conditions stated as [84]

$$\begin{aligned} h_i(\boldsymbol{x}^*) = 0, \qquad & f_i(\boldsymbol{x}^*) \leq 0, \\ & \phi_i^* \geq 0, \\ \nabla f_0(\boldsymbol{x}^*) + \sum_{i=1}^{m} \phi_i^* \nabla f_i(\boldsymbol{x}^*) + \sum_{i=1}^{p} \upsilon_i^* \nabla h_i(\boldsymbol{x}^*) = 0, & \\ & \phi_i^* f_i(\boldsymbol{x}^*) = 0. \end{aligned} \tag{3.12}$$

KKT conditions are very important and useful in solving convex optimization problems analytically. However, for most convex optimization problems, there is no analytical solution. Some numerical methods have been developed to efficiently solve convex optimization. For example, interior-point method has been developed to solve constrained convex optimization problems to a specified accuracy [84]. It uses Newton's method to solve the original convex problem (3.2) or its KKT condition (3.12) by iteratively searching the optimal solution to its descent direction on which the value of the function decreases the most [84, 85].

3.1.4 Jensen's Inequality

A very important theory used in the thesis is Jensen's Inequality, which extends the inequality in (3.1) to convex combinations of more than two points. In other words, if f is convex, $\boldsymbol{x}_1, \cdots, \boldsymbol{x}_k \in \mathbf{dom} f$, and $\theta_1, \cdots, \theta_k \geq 0$, with $\theta_1 + \cdots + \theta_k = 1$, then

$$f(\theta_1 \boldsymbol{x}_1 + \cdots + \theta_k \boldsymbol{x}_k) \leq \theta_1 f(\boldsymbol{x}_1) + \cdots + \theta_k f(\boldsymbol{x}_k). \tag{3.13}$$

This is called *Jensen's Inequality*. When the sums become infinite, for example, if $p(\boldsymbol{x}) \geq 0$ on $\mathcal{S} \subseteq \mathbf{dom}\ f$, $\int_{\mathcal{S}} p(\boldsymbol{x})d\boldsymbol{x} = 1$, then we have

$$f\left(\int_{\mathcal{S}} p(\boldsymbol{x})\boldsymbol{x}d\boldsymbol{x}\right) \leq \int_{\mathcal{S}} f(\boldsymbol{x})p(\boldsymbol{x})d\boldsymbol{x}. \tag{3.14}$$

If $\boldsymbol{x}$ is a random variable such that $\boldsymbol{x} \in \mathbf{dom} f$ with probability one, and f is convex, then we have

$$f(\mathtt{E}[\boldsymbol{x}]) \leq \mathtt{E}[f(\boldsymbol{x})]. \tag{3.15}$$

3.2 Nonlinear Optimization Theory

Nonlinear optimization is described as an optimization problem (3.2) when the objective or constraint functions are not linear and not known to be convex. Generally, there are no effective ways to solve such problems. Several different approaches exist to solve general nonlinear problems. In the following, we list some nonlinear optimization algorithms used in the later chapters in the thesis.

When the objective function is a real value function with high complexity and the constraints are given in bound constraints, DIRECT algorithm has been used in the thesis to solve the problems for global optimization [86]. DIRECT is a sampling algorithm which requires no knowledge of the objective function gradient. The algorithm samples points in the domain, and uses the information it has obtained to decide where to search next. However, for this algorithm the global optimization may come at the expense of a large and exhaustive search over the domain.

Alternatively, a more general method, sequential quadratic programming (SQP), can solve most of the nonlinear optimization at an affordable complexity [87]. However, this method can only guarantee a local optimization solution.

3.3 DP and Optimal Control Theory

Different from general optimization theory, DP and optimal control theory deals with situations where decisions are made in stages with the objective to minimize a certain cost [83]. For each of the stages, the outcome of the decision may not be fully predictable but can be anticipated to some extent before the next decision is made. Decisions must be balanced to consider both the desire for low present cost and the undesirability of high future costs. In order to capture this tradeoff, DP ranks decisions based on the sum of the present cost and the expected future cost, assuming optimal decision making for subsequent stages.

The uniqueness of DP lies in two aspects. First, the optimization problem is solved over time. Second, DP can achieve the optimal solution at a relatively affordable cost of complexity. Thus, it has been used in wireless communication systems to solve the power allocation problem [68, 81].

3.3.1 Basic Problem

The basic problem for DP has two principal features: (1) an underlying discrete-time dynamic system and (2) a cost function that is additive over time. The system has the form

$$x_{i+1} = f_i(x_i, p_i, g_i), \quad \text{for } i = 1, 2, \ldots, K, \tag{3.16}$$

where x_i is the state of the system at stage i, which summarizes past information that is relevant for future optimization, $p_i = \mu_i(x_i)$ is the control or decision variable to be selected at stage i based on the policy function $\mu_i(\cdot)$,[1] g_i is a random parameter (i.e., disturbance or noise depending on the context) and K is the number of times that control is applied. The function $f_i(\cdot)$ describes the mechanism by

[1]There is a main difference between control and policy. Policy is a global solution for every possible system state while control is the use of the policy for a specific initialization or state.

which the state is updated. For a given initial state x_1, DP has the cost function

$$J(x_1) = \mathtt{E}_{\boldsymbol{g}} \left[c_{K+1}(x_{K+1}) + \sum_{i=1}^{K} c_i(x_i, \mu_i(x_i), g_i) \right], \tag{3.17}$$

where c_i is the cost incurred at time i with c_{K+1} being the terminal cost incurred at the end of the process. The optimal policy $\boldsymbol{\mu}^* = \{\mu_1^*, \ldots, \mu_K^*\}$ is one that minimizes this cost (3.17).

3.3.2 Solving DP Problems

The optimal policy DP solving the above problem is by proceeding backwards in time from period K to period 1 as follows [83]

$$J_i(x_i) = \min_{p_i} \mathtt{E}_{\boldsymbol{g}} \left[c_i(x_i, p_i, g_i) + J_{i+1}\left(x_{i+1}\right) \right], \quad \text{for } i = 1, 2, \ldots, K, \tag{3.18}$$

with initialization $J_{K+1}(x_{K+1}) = c_{K+1}(x_{K+1})$ representing the terminal cost and $J_i(x_i)$ denoting the cost of the tail subproblem that starts at time i at state x_i. The optimal cost $J^*(x_1)$ is given by $J_1(x_1)$ at the last step of the algorithm. In other words, DP proceeds sequentially to solve all the tail subproblems of a given time length, by using the solution of the tail subproblems of shorter time length. Note that all the tail subproblems are solved optimally in addition to the original problem and hence lead to the intensive computational requirements.

3.3.3 Problem Reformulations

In some applications the system cost depends not only on the current state x_i and control p_i, but also the earlier states and controls. In other words, states and controls influence future costs with some time lag. This situation can be handled by *state augmentation*, which expands the state to include earlier states

and controls. In this case, the problem can be formulated to the basic format by introducing the augmented state

$$\tilde{\boldsymbol{x}}_i = (x_1, x_2, \ldots, x_i, p_1, p_2, \ldots, p_{i-1}, g_1, g_2, \ldots, g_{i-1}). \tag{3.19}$$

An example of time lags in the cost, which is particularly related in the thesis, occurs in the non-additive form with an initial state x_1 as

$$J(x_1) = \mathtt{E}_{\boldsymbol{g}}\left[c_{K+1}(x_{K+1}, x_K, \cdots, x_1, p_K, p_{K-1} \cdots, p_1, g_K, g_{K-1}, \cdots, g_1)\right]. \tag{3.20}$$

and $\mathtt{E}_{\boldsymbol{g}}\left[c_{K+1}(\tilde{\boldsymbol{x}}_{K+1})\right]$ is considered as the reformulated cost.

Chapter 4

Wireless Resource Allocation in Single User Systems with Transmitter SCI

In Chapter 4 of the thesis, we will discuss the wireless resource allocation problem in single user transmission systems. The aim is to minimize transmit power in single user MIMO systems with transmitter side SCI while satisfying user's QoS requirement in terms of information outage probability under certain decoding delay constraint. The mathematical tools used in this chapter are non-linear optimization and mathematical calculus. Compared with the traditional resource allocation problem, our proposed schemes first consider power minimization problem under users' QoS requirement in terms of information outage probability. Note that wireless resource allocation problem in single user systems with transmitter CSI will be considered in Chapter 5.

This chapter is organized as follows. Introduction and motivation are given in Section 4.1. In Section 4.2, we will present the system model. Section 4.3 further analyzes the problem to be solved. In Section 4.4, closed-form mean and variance

of MIMO channel capacity are derived. In the following, the power minimization problem is solved in Section 4.5. Simulation results are shown in Section 4.6 and we summarize this chapter in Section 4.7.

4.1 Introduction

The aim of this chapter is to solve the transmit power minimization problem in single user MIMO system with transmitter side SCI while satisfying user's QoS requirement in terms of information outage probability under certain decoding delay constraint. There exists limited research work considering power allocation strategies in single user systems for delay constrained communication. In [64, 65], an acausal BF SISO and MIMO channel was considered and the optimal power allocation under short-term or long-term power constraint are obtained to maximize average capacity or minimize the outage probability. More recently, in [88, 89], the power allocation problem in an acausal BF channel, similar to [64], but with a service outage probability constraint (which requires a variable-rate encoder at the transmitter to adapt the transmission service rate) as opposed to an information outage (in which case a fixed-rate encoder is considered) has been addressed. In [66], the authors extend the research in [65] and propose a near-optimal low complexity power allocation strategy for achieving capacity in an acausal BF channel and further apply it in MIMO multicarrier systems [67]. Acausality of CSI may be valid for parallel BF channels (i.e., blocks separated in frequency) but not for serial channels (i.e., blocks separated in time). This has motivated the adaptive transmission research works for causal BF channels.

In [81, 90], the authors exploited the causal CSI of BF channels for maximizing the expected rate and minimizing the outage probability by optimizing the power allocation over the blocks using a DP approach. Simplified, but suboptimal, techniques for the expected rate maximization problems have also been proposed in

[91, 92]. Note that, all of the above-mentioned works rely on the assumption that the instantaneous CSI is known at the transmitter, which requires a delay-free feedback channel, and will be difficult to obtain for most situations. For this reason, this chapter considers only SCI available at the transmitter.

In contrast to the traditional schemes that aim to maximize the expected achievable rate or minimize the outage probability, future wireless communication systems are anticipated to provide a wide range of applications with diverse QoS requirements that will require QoS-driven resource allocation schemes to efficiently allocate the radio resources for accommodating heterogeneous mobile data [19]. In addition, the application of MIMO technology brings another challenge for the design. This motivates us to develop a more appealing approach to minimize the transmission cost (in terms of the average transmitted power per block) for achieving a given QoS in terms of information outage probability in MIMO system.

In the following, we aim to solve the problem of minimizing the transmit power required for achieving a given information outage probability with transmitter SCI under certain decoding delay constraint. We use Gaussian approximation to express the p.d.f. of the IMI over transmission blocks, and derive analytically the mean and variance of the IMI of the MIMO channel for power minimization. It is shown that the optimal power allocation based on the Gaussian approximation can be obtained numerically by using one of the nonlinear optimization method DIRECT. The problem with causal CSI assumed at the transmitter will be solved in Chapter 5.

4.2 System Model

A MIMO BF channel is considered. As introduced in Section 2.2.1, it can be represented by a complex matrix $\mathbf{H}_k = \left[h_{i,j}^{(k)}\right] \in \mathbb{C}^{n_r \times n_t}$, where n_t and n_r antennas are, respectively, located at the transmitter and the receiver. The amplitude square

of each element, $\left|h_{i,j}^{(k)}\right|^2$, has the p.d.f. of (4.1) as that of c_k,

$$f(c_k) = \begin{cases} e^{-c_k} & \text{if } c_k \geq 0, \\ 0 & \text{if } c_k < 0, \end{cases} \tag{4.1}$$

where $\mathtt{E}[c_k] = 1 \ \forall k$. And the elements of $\mathbf{H}_k$ are i.i.d. for different block k and antenna pairs. In addition, it is assumed that the variance of the noise $N_0 = 1$ for normalization.

Accordingly, the IMI for block k can be found as [33]

$$r_k = \log_2 \det \left| \mathbf{I} + \frac{P}{n_t} \mathbf{H}_k \mathbf{H}_k^\dagger \right| \quad \text{in bps/Hz.} \tag{4.2}$$

In the above, we have used the fact that the transmit covariance matrix at time k is $\frac{P}{n_t}\mathbf{I}$. This is because the transmitter does not have the instantaneous CSI and thus transmits the same power across the antennas. And by transmitting power of $\frac{P}{n_t}$ at each antenna, the transmit power at each block is kept as P. In the sequel, we shall assume that $n_t \geq n_r$.The case of $n_t < n_r$ can be treated in a similar way and thus omitted for conciseness.

Denoting the target code rate as R_0, as introduced in Section 2.1.3, an information outage is said to occur if the IMI over K blocks $\frac{1}{K}\sum_{k=1}^{K} r_k$, is less than R_0. As no CSI is available at the transmitter, equal power allocation over K blocks are performed. If we further denote the target information outage probability as ε_0, the problem of interest is to minimize the average transmit power per block subject to a target information outage probability, i.e.,

$$\begin{aligned} \min_{P \geq 0} \quad & P \\ \text{s.t.} \quad & \mathsf{Pr}\left\{ \frac{1}{K} \sum_{k=1}^{K} r_k < R_0 \right\} \leq \varepsilon_0. \end{aligned} \tag{4.3}$$

4.3 Minimum Power for a Given Outage Probability

Given the model described above, our aim is further elaborated as:

$$\begin{aligned} \min_{P\geq 0} \quad & P \\ \text{s.t.} \quad & \mathsf{Pr}\left\{\frac{1}{K}\sum_{k=1}^{K}\log_2\det\left|\mathbf{I}+\frac{P}{n_t}\mathbf{H}_k\mathbf{H}_k^{\dagger}\right| < R_0\right\} \leq \varepsilon_0. \end{aligned} \tag{4.4}$$

To proceed further, we rewrite the outage probability in (4.4) as follows:

$$\begin{aligned} \mathcal{P}_{\mathsf{out}} &= \mathsf{Pr}\left\{\sum_{k=1}^{K}\log_2\det\left|\mathbf{I}+\frac{P}{n_t}\mathbf{H}_k\mathbf{H}_k^{\dagger}\right| < KR_0\right\} \\ &= \mathsf{Pr}\left\{\sum_{k=1}^{K}\log_2\det\left|\mathbf{I}+\frac{P\mathbf{\Lambda}_k}{n_t}\right| < KR_0\right\} \end{aligned} \tag{4.5}$$

where $\mathbf{\Lambda}_k \triangleq \mathrm{diag}\left(\lambda_1^{(k)}, \lambda_2^{(k)}, \ldots, \lambda_{n_r}^{(k)}\right)$ with $\lambda_1^{(k)} \geq \lambda_2^{(k)} \geq \cdots \geq \lambda_{n_r}^{(k)} > 0$ representing the ordered eigenvalues of $\mathbf{H}_k\mathbf{H}_k^{\dagger}$. Note also from our assumption that $n_r = \min\{n_t, n_r\} = \mathrm{rank}(\mathbf{H}_k)\ \forall k$. The random variables of the outage probability are the eigenvalues $\left\{\lambda_j^{(k)}\right\}$ whose joint p.d.f. is [93]

$$f(\mathbf{\Lambda}) = \frac{\left(\prod_{i=1}^{n_r}\lambda_i\right)^{n_t-n_r}e^{-\sum_{i=1}^{n_r}\lambda_i}}{n_r!\prod_{i=1}^{n_r}(n_r-i)!(n_t-i)!} \times \prod_{1\leq i\leq j\leq n_r}(\lambda_i-\lambda_j)^2, \quad \text{for } \lambda_1,\ldots,\lambda_{n_r}>0, \tag{4.6}$$

where the time index k is omitted for conciseness. Evaluation of the outage probability requires knowing the p.d.f. of $\sum_{k=1}^{K} r_k$ with (4.6), and it has unfortunately been unknown so far. Recently, it was found in [94] that the p.d.f. of r (r_k with the subscript k omitted) can be approximated as Gaussian (so does the $\sum_{k=1}^{K} r_k$) with the mean (ρ_r) and variance (σ_r^2) derived, respectively, as [33, 94]

$$\rho_r = \int_0^{\infty}\log_2\left(1+\frac{P\lambda}{n_t}\right)\sum_{n=0}^{n_r-1}\frac{n!\lambda^{n_t-n_r}e^{-\lambda}}{(n+n_t-n_r)!}\left[L_n^{(n_t-n_r)}(\lambda)\right]^2 d\lambda \tag{4.7}$$

where $L_n^{(n_t-n_r)}(x)$ denotes the generalized Laguerre polynomial of order n, and

$$\sigma_r^2 = n_r \int_0^\infty \omega^2(P,\lambda) p(\lambda) d\lambda - \sum_{i=1}^{n_r}\sum_{j=1}^{n_r} \frac{(i-1)!(j-1)!}{(i-1+n_t-n_r)!(j-1+n_t-n_r)!} \times \left[\int_0^\infty \lambda^{n_t-n_r} e^{-\lambda} L_{i-1}^{(n_t-n_r)}(\lambda) L_{j-1}^{(n_t-n_r)}(\lambda) \omega(P,\lambda) d\lambda\right]^2 \tag{4.8}$$

in which $\omega(P,\lambda) \triangleq \log_2\left(1+\frac{P\lambda}{n_t}\right)$ and

$$p(\lambda) \triangleq \frac{1}{n_r}\sum_{i=1}^{n_r} \frac{(i-1)!}{(i-1+n_t-n_r)!} \lambda^{n_t-n_r} e^{-\lambda} \left[L_{i-1}^{(n_t-n_r)}(\lambda)\right]^2. \tag{4.9}$$

In [94], it was demonstrated that using Gaussian approximation on the rate of a MIMO channel is accurate even with small number of antennas and this claim will be substantiated in Section 4.6 where numerical results will be provided to verify its validity. In light of this, we shall use Gaussian approximation on the sum-rate, $\sum_{k=1}^K r_k$ (as this is a sum of independent random variables, clearly, the approximation will further improve if K increases). Consequently, the probability constraint can be expressed as

$$\frac{1}{2}\left\{1+\mathsf{erf}\left[\frac{R_0 - K\rho_r(P)}{\sigma_r(P)\sqrt{2K}}\right]\right\} \le \varepsilon_0 \tag{4.10}$$

where $\mathsf{erf}(x) \triangleq \frac{2}{\sqrt{\pi}}\int_0^x e^{-t^2} dt$. Accordingly, the optimization problem (4.4) can be re-expressed as

$$\begin{aligned} \min_{P\ge 0} \quad & P \\ \text{s.t.} \quad & \frac{1}{2}\left\{1+\mathsf{erf}\left[\frac{R_0 - K\rho_r(P)}{\sigma_r(P)\sqrt{2K}}\right]\right\} \le \varepsilon_0. \end{aligned} \tag{4.11}$$

The remaining challenges are then to derive the closed form expressions for the mean (4.7) and the variance (4.8), and to seek a method that can solve the problem (4.11).

4.4 Derivation of ρ_r and σ_r^2

Here, we shall first derive the mean and then the variance of the rate with $K = 1$. Before we do this, the following useful expansion of the generalized Laguerre polynomial should be noted

$$L_n^v(\lambda) \triangleq \sum_{m=0}^{n} (-1)^m \frac{(n+v)!}{(n-m)!(v+m)!m!} \cdot \lambda^m, \tag{4.12}$$

where $v = n_t - n_r$. To make our notation succinct, we define $a \triangleq \frac{P}{n_t}$, and

$$b_m(n,v) \triangleq (-1)^m \frac{(n+v)!}{(n-m)!(v+m)!m!} = \frac{(-1)^m}{m!} \binom{n+v}{m+v} \tag{4.13}$$

so that

$$L_n^v(\lambda) = \sum_{m=0}^{n} b_m(n,v)\lambda^m. \tag{4.14}$$

As a result, the mean, ρ_r, can be derived as follows:

$$\begin{aligned}
\rho_r &= \int_0^\infty \log_2(1+a\lambda) \sum_{n=0}^{n_r-1} \frac{n!\lambda^{n_t-n_r}e^{-\lambda}}{(n+n_t-n_r)!} \left[L_n^{n_t-n_r}(\lambda)\right]^2 d\lambda \\
&= \sum_{n=0}^{n_r-1} \frac{n!}{(n+v)!} \int_0^\infty \log_2(1+a\lambda)\lambda^v e^{-\lambda} \left[L_n^v(\lambda)\right]^2 d\lambda \\
&= \sum_{n=0}^{n_r-1} \frac{n!}{(n+v)!} \int_0^\infty \log_2(1+a\lambda)\lambda^v e^{-\lambda} \left[\sum_{m=0}^{n} b_m(n,v)\lambda^m\right]^2 d\lambda \\
&= \sum_{n=0}^{n_r-1} \frac{n!}{(n+v)!} \int_0^\infty \log_2(1+a\lambda)\lambda^v e^{-\lambda} \left[\sum_{m=0}^{n} b_m(n,v)^2\lambda^{2m} + \right. \\
&\qquad \left. 2\sum_{i=0}^{n-1}\sum_{j=i+1}^{n} b_i(n,v)b_j(n,v)\lambda^{i+j}\right] d\lambda \\
&= \sum_{n=0}^{n_r-1}\sum_{m=0}^{n} \frac{n!}{(n+v)!} b_m^2(n,v) \int_0^\infty \log_2(1+a\lambda)\lambda^v e^{-\lambda}\lambda^{2m} d\lambda \\
&\quad +2\sum_{n=1}^{n_r-1}\sum_{i=0}^{n-1}\sum_{j=i+1}^{n} \frac{n!}{(n+v)!} b_i(n,v)b_j(n,v) \int_0^\infty \log_2(1+a\lambda)\lambda^v e^{-\lambda}\lambda^{i+j} d\lambda \\
&= \frac{1}{\ln 2}\sum_{n=0}^{n_r-1}\sum_{m=0}^{n} \frac{n!}{(n+v)!} \left[\frac{1}{m!}\binom{n+v}{m+v}\right]^2 \int_0^\infty \ln(1+a\lambda)\lambda^{v+2m}e^{-\lambda} d\lambda \\
&\quad +\frac{2}{\ln 2}\sum_{n=1}^{n_r-1}\sum_{i=0}^{n-1}\sum_{j=i+1}^{n} \frac{n!}{(n+v)!} \frac{(-1)^{i+j}}{i!j!} \binom{n+v}{i+v}\binom{n+v}{j+v} \\
&\qquad \int_0^\infty \ln(1+a\lambda)\lambda^{v+i+j}e^{-\lambda} d\lambda, \qquad (4.15)
\end{aligned}$$

where the integral of the form $\int_0^\infty \lambda^j e^{-\lambda}\ln(1+a\lambda)d\lambda$ is given in Appendix 10.2 as

$$\begin{aligned}
\int_0^\infty \lambda^j e^{-\lambda}\ln(1+a\lambda)d\lambda &= \frac{e^{\frac{1}{a}}}{a^{j-1}} \sum_{n=0}^{j} (-1)^n \binom{j}{n} \frac{(j-n)!}{\left(\frac{1}{n}\right)^{j-n}} E_1\left(\frac{1}{a}\right) \\
&\quad + \frac{1}{a^j}\sum_{n=0}^{j-1}\sum_{p=1}^{j-n} (-1)^n \binom{j}{n} \frac{(j-n)!}{\left(\frac{1}{n}\right)^{j-n}} \cdot \frac{1}{j-n+1-p} \\
&\quad + \frac{1}{a^{j+1}}\sum_{n=0}^{j-2}\sum_{p=2}^{j-n}\sum_{q=1}^{p-1} (-1)^n \binom{j}{n} \frac{(j-n)!}{(j-n-p+1)!} \frac{a^p}{j-n-q+1}, \qquad (4.16)
\end{aligned}$$

in which $E_1(z) = \int_z^\infty \frac{e^{-t}}{t} dt$ denotes the exponential integral.

For the variance, we express it using the standard result as

$$\begin{aligned}\sigma_r^2 &= \int_0^\infty \log_2^2(1+a\lambda)\sum_{n=0}^{n_r-1}\frac{n!\lambda^v e^{-\lambda}}{(n+v)!}\left[L_n^v(\lambda)\right]^2 d\lambda \\ &\quad -\sum_{i=0}^{n_r-1}\sum_{j=0}^{n_r-1}\frac{i!j!}{(i+v)!(j+v)!}\left[\int_0^\infty \lambda^v e^{-\lambda}L_i^v(\lambda)L_j^v(\lambda)\log_2(1+a\lambda)d\lambda\right]^2 \\ &\equiv \mathcal{I}_1-\mathcal{I}_2 \end{aligned} \tag{4.17}$$

which boils down to evaluating the integrals $\mathcal{I}_1$ and $\mathcal{I}_2$. After some manipulations, we have $\mathcal{I}_1$ as

$$\begin{aligned}\mathcal{I}_1 &= \frac{1}{\ln^2 2}\sum_{n=0}^{n_r-1}\sum_{m=0}^{n}\frac{n!}{(n+v)!}\left(\frac{1}{m!}\binom{n+v}{m+v}\right)^2\int_0^\infty \ln^2(1+a\lambda)\lambda^{v+2m}e^{-\lambda}d\lambda \\ &+\frac{2}{\ln^2 2}\sum_{n=1}^{n_r-1}\sum_{i=0}^{n-1}\sum_{j=i+1}^{n}\frac{n!}{(n+v)!}\frac{(-1)^{i+j}}{i!j!}\binom{n+v}{i+v}\binom{n+v}{j+v}\int_0^\infty \ln^2(1+a\lambda)\lambda^{v+i+j}e^{-\lambda}d\lambda \end{aligned} \tag{4.18}$$

where $\int_0^\infty \lambda^j e^{-\lambda}\ln(1+a\lambda)^2 d\lambda$ is derived in Appendix 10.2 as

$$\begin{aligned}&\int_0^\infty \left[\ln(1+a\lambda)\right]^2\lambda^j e^{-\lambda}d\lambda \\ &=\frac{e^{\frac{1}{a}}}{a^{j+1}}\sum_{n=0}^{j}(-1)^n\binom{j}{n}\frac{(j-n)!}{\left(\frac{1}{a}\right)^{j-n}}\left\{a\left[\left(\ln\frac{1}{a}-\gamma_{\mathsf{EM}}\right)^2+\frac{\pi^2}{6}\right]\right. \\ &\quad \left.-2{}_3F_3\left([1,1,1];[2,2,2];-\frac{1}{a}\right)\right\} \\ &\quad +\frac{2e^{\frac{1}{a}}}{a^j}\sum_{n=0}^{j-1}\sum_{p=1}^{j-n}(-1)^n\binom{j}{n}\frac{(j-n)!}{\left(\frac{1}{a}\right)^{j-n}}\cdot\frac{1}{j-n+1-p}E_1\left(\frac{1}{a}\right) \\ &\quad +\frac{2}{a^j}\sum_{n=0}^{j-2}\sum_{p=1}^{j-n-1}\sum_{q=p+1}^{j-n}(-1)^n\binom{j}{n}\frac{(j-n)!}{\left(\frac{1}{a}\right)^{j-n}}\cdot\frac{1}{(j-n+1-p)(j-n+1-q)} \\ &\quad +\frac{2}{a^{j+1}}\sum_{n=0}^{j-3}\sum_{t=3}^{j-n}\sum_{p=1}^{t-2}\sum_{q=p+1}^{t-1}(-1)^n\binom{j}{n}\frac{(j-n)!}{(j-n-t+1)!}\cdot\frac{a^t}{(j-n+1-p)(j-n+1-q)} \end{aligned} \tag{4.19}$$

where ${}_pF_q$ denotes the generalized hypergeometric function and γ_{EM} is the Euler constant [95]. On the other hand, $\mathcal{I}_2$ can be obtained using the following result:

$$\begin{aligned}
&\int_0^\infty \lambda^v e^{-\lambda} L_i^v(\lambda) L_j^v(\lambda) \log_2(1+a\lambda) d\lambda \\
&= \int_0^\infty \lambda^v e^{-\lambda} \log_2(1+a\lambda) \left(\sum_{m=0}^{i} b_m(i,v) \lambda^m \right) \left(\sum_{n=0}^{j} b_n(j,v) \lambda^n \right) d\lambda \\
&= \int_0^\infty \lambda^v e^{-\lambda} \log_2(1+a\lambda) \left(\sum_{m=0}^{i} \sum_{n=0}^{j} b_m(i,v) b_n(j,v) \lambda^{m+n} \right) d\lambda \\
&= \sum_{m=0}^{i} \sum_{n=0}^{j} b_m(i,v) b_n(j,v) \int_0^\infty \lambda^v e^{-\lambda} \log_2(1+a\lambda) \lambda^{m+n} d\lambda \\
&= \sum_{m=0}^{i} \sum_{n=0}^{j} b_m(i,v) b_n(j,v) \int_0^\infty \lambda^{v+m+n} e^{-\lambda} \log_2(1+a\lambda) d\lambda. \qquad (4.20)
\end{aligned}$$

Now, combining the results (4.17), (4.18) and (4.20), we have

$$\begin{aligned}
\sigma_r^2 =& \frac{1}{\ln^2 2} \sum_{n=0}^{n_r-1} \sum_{m=0}^{n} \frac{n!}{(n+v)!} \left[\frac{1}{m!} \binom{n+v}{m+v} \right]^2 \int_0^\infty \ln^2(1+a\lambda) \lambda^{v+2m} e^{-\lambda} d\lambda \\
&+ \frac{2}{\ln^2 2} \sum_{n=1}^{n_r-1} \sum_{i=0}^{n-1} \sum_{j=i+1}^{n} \frac{n!}{(n+v)!} \frac{(-1)^{i+j}}{i!j!} \binom{n+v}{i+v} \binom{n+v}{j+v} \int_0^\infty \ln^2(1+a\lambda) \lambda^{v+i+j} e^{-\lambda} d\lambda \\
&- \frac{1}{\ln^2 2} \sum_{i=0}^{n_r-1} \sum_{j=0}^{n_r-1} \frac{i!j!}{(i+v)!(j+v)!} \times \\
&\left[\sum_{m=0}^{i} \sum_{n=0}^{j} \frac{(-1)^{m+n}}{m!n!} \binom{i+v}{m+v} \binom{j+v}{n+v} \int_0^\infty \lambda^{v+m+n} e^{-\lambda} \ln(1+a\lambda) d\lambda \right]^2 \qquad (4.21)
\end{aligned}$$

where the integrals of the forms $\int_0^\infty \lambda^j e^{-\lambda} \ln(1+a\lambda) d\lambda$ and $\int_0^\infty \lambda^j e^{-\lambda} \ln^2(1+a\lambda) d\lambda$ are, respectively, given by (4.16) and (4.19).

4.5 Finding the Minimum Power Numerically

The optimization problem (4.11) can be considered to find the minimum power that satisfies

$$f(P) \triangleq K\rho_r\left(\frac{P}{n_t}\right) - \sqrt{2K}\sigma_r\left(\frac{P}{n_t}\right)\mathsf{erf}^{-1}(1-2\varepsilon_0) - R_0 \geq 0. \tag{4.22}$$

Obviously, when P increases, the outage probability will decrease. This indicates the left-hand-side of the above equation will be much greater than 0 if too much power P is allocated. However, if P is too small, the outage constraint will not be satisfied. As the left-hand-side of the constraint is evidently an increasing function of P, the minimum P can be obtained by finding the root of $f(P) = 0$. The above equation used to find the minimum power for a MIMO system is named as minimum power equation (MPE). Function f is a complicated non-linear function involving numerical functions, such as exponential integration and hypergeometric function. Therefore, it is extremely difficult to find the closed-form solution using traditional optimization method. However, as f can be written in closed-form with the help of (4.15) and (4.21), and it is a strictly increasing function where there exists a unique solution for $f(P) = 0$, it is possible to find optimal P^* numerically.

In order to find the solution of $f(P) = 0$, one possible method is to use one of the nonlinear optimization method DIRECT algorithm introduced in Section 3.2 [86, 96]. DIRECT can deal with difficult functions and globally converge to the minimal value of the objective function if the number of iterations and function evaluations is sufficient. In our problem, the root of $f(P) = 0$ can be found efficiently by constructing the following problem:

$$P^* = \arg\min_{0\leq P\leq U_P} |f(P)|. \tag{4.23}$$

We set the upper bound U_P to be a value large enough to make sure $P^* \in (0, U_P)$.

4.6 Simulation Results

4.6.1 Setup

Computer simulations are conducted to evaluate the performance of the proposed power minimization techniques based on different information outage probability constraints for various target code rates and MIMO systems. Channels between each transmit and receive antenna are i.i.d. Rayleigh distributed and transmitter SCI is assumed. We define the transmit SNR per block $\mathsf{SNR} \triangleq P$, as our performance measurement.

4.6.2 Benchmarks

We compare our simulation results with power optimization schemes $\mathbb{Q}_1$ and $\mathbb{Q}_2$ which are based on the upper bound outage rate probability from [97]. $\mathbb{Q}_1$ and $\mathbb{Q}_2$ are derived by replacing the outage rate probability (6.15) by an upper bound shown below.

$$\begin{aligned}
\mathcal{P}_{\mathsf{out}} &= \mathsf{Pr}\left(\left\{\frac{1}{K}\sum_{k=1}^{K} r_k < R_0\right\}\right) \\
&\leq \mathsf{Pr}\left(\left\{\frac{1}{K}\sum_{k=1}^{K} \log_2\left(1+\frac{P\lambda_1^{(k)}}{n_t}\right) < R_0\right\}\right) \equiv \mathcal{P}_{\mathsf{out}}^{\mathsf{UB1}} \\
&\leq \mathsf{Pr}\left(\left\{\frac{1}{K}\sum_{k=1}^{K} \log_2\left(\frac{P\lambda_1^{(k)}}{n_t}\right) < R_0\right\}\right) \\
&= \mathsf{Pr}\left(\left\{\alpha \triangleq \sum_{k=1}^{K} \ln \lambda_1^{(k)} < \ln\left[\left(\frac{n_t}{P}\right)^K 2^{KR_0}\right]\right\}\right) \equiv \mathcal{P}_{\mathsf{out}}^{\mathsf{UB2}}.
\end{aligned} \tag{4.24}$$

Here $\lambda_1^{(k)}$ is the largest eigenvalue of the matrix $\mathbf{H}_k\mathbf{H}_k^{\dagger}$ with p.d.f. shown as [98]

$$f(\lambda) = \sum_{i=1}^{n_r} \sum_{j=n_t-n_r}^{(n_t+n_r)i-2i^2} \left(d_{i,j}\cdot\frac{i^{j+1}}{j!}\right)\lambda^j e^{-i\lambda}, \quad \lambda > 0. \tag{4.25}$$

As such, the optimization problem (4.4) denoted as $\mathbb{Q}_0$ is solved sub-optimally by $\mathbb{Q}_1$ and $\mathbb{Q}_2$ summarized below:

$$\mathbb{Q}_1 \mapsto \begin{cases} \min\limits_{P\geq 0} & P \\ \text{s.t.} & \mathcal{P}_{\text{out}}^{\text{UB1}} \leq \varepsilon_0, \end{cases} \quad \text{and} \quad \mathbb{Q}_2 \mapsto \begin{cases} \min\limits_{P\geq 0} & P \\ \text{s.t.} & \mathcal{P}_{\text{out}}^{\text{UB2}} \leq \varepsilon_0. \end{cases} \tag{4.26}$$

Due to the simplification of the constraint, the power minimization in $\mathbb{Q}_1$ or $\mathbb{Q}_2$ is done more conservatively than in $\mathbb{Q}_0$. Therefore,

$$\inf\, \mathbb{Q}_0 \leq \inf\, \mathbb{Q}_1 \leq \inf\, \mathbb{Q}_2. \tag{4.27}$$

4.6.3 Results

The accuracy of the approximation of Gaussian can be observed from Fig. 4.1 where the c.d.f. are plotted for a (3,2) MIMO system with different number of blocks but same average transmit SNR = 10dB. The solid lines representing the actual simulated results of the "sum-rate" KR_0 of K-block MIMO channel (even for $K = 1$) are extremely close to the dashed lines donating the Gaussian approximation, even when $K = 1$. The reason is stated as follows.

For MIMO channel with the elements of channel matrix $\mathbf{H}$ being i.i.d. Rayleigh fading, the capacity distribution is asymptotically Gaussian if the number of transmit and receive antennas n_t and n_r goes to infinity [94]. This can be proved by a central limit theorem for MIMO channel capacity [99]. Therefore, when n_t and n_r are large enough, Gaussian distribution can be used to approximate MIMO channel distribution. According to [94], when $\min(n_t, n_r) \geq 5$, the Gaussian approximation fits remarkably well with MIMO channel capacity distribution. For the worst case, when n_t and n_r are small, i.e. $n_t = n_r = 1$ for SISO channel, the Gaussian approximation is not very accurate. However, the fitting of Gaussian approximation in this case is still quite respectable [94].

Fig. 4.2 shows the average transmit SNR as a function of outage probability ε_0 in a $(4,3)$ MIMO system when $K = 3$. $\mathbb{Q}_0$ (plotted with triangles) represents the proposed optimal transmit SNR, while $\mathbb{Q}_1$ and $\mathbb{Q}_2$ (plotted with squares and circles respectively) are sub-optimal results from (4.26). From the figure we can observe that with the same outage probability and transmission rate, transmit SNR required in each scheme satisfies $\mathbb{Q}_0 < \mathbb{Q}_1 < \mathbb{Q}_2$. The transmit SNR gap under $\mathbb{Q}_0$ and $\mathbb{Q}_1$ is relatively larger when the transmission rate R_0 is larger. For example, when $\varepsilon_0 = 10^{-4}$, the gap of transmit SNR between $\mathbb{Q}_0$ and $\mathbb{Q}_1$ increases from around 4dB to 8dB when R_0 changes from 3 to 6 bps/Hz. This can be explained by writing MIMO capacity for one block as

$$\log_2\left(1+\frac{P\lambda_1}{n_t}\right)\left(1+\frac{P\lambda_2}{n_t}\right)\cdots\left(1+\frac{P\lambda_{n_r}}{n_t}\right). \tag{4.28}$$

Assuming λ_1 is the largest eigenvalue, $\log_2\left(1+\frac{P\lambda_1}{n_t}\right)$ is the transmission rate proposed in $\mathbb{Q}_1$. When a relatively larger transmit SNR is required, i.e., R_0 increases, the multiplied value $\left(1+\frac{P\lambda_2}{n_t}\right)\cdots\left(1+\frac{P\lambda_{nr}}{n_t}\right)$ becomes larger. This brings more difference in the actual transmission rate between $\mathbb{Q}_0$ and $\mathbb{Q}_1$.

On the other hand, when the outage probability decreases, i.e., from 10^{-1} to 10^{-4}, only a slightly more power (around 2 dB) is required for $\mathbb{Q}_0$ for all different transmission rate R_0. In addition, when R_0 changes, i.e. from 3 to 6 bps/Hz, around 6 dB more power is needed at the transmitter. This indicates that for MIMO system the transmit SNR is mainly determined by transmission rate, rather than outage probability.

The relationship between transmit SNR and transmission rate R_0 can be observed from Figs. 4.3 and 4.4. The figures present the average SNR needed for various transmission rate R_0 for (2,2) and (4,3) MIMO systems with $\varepsilon_0 = 0.01$ and 0.0001 respectively. From the figures we can see that the greater the transmission rate, the larger SNR is required to achieve the same outage probability constraint. In

particular, results show that the increase in SNR is almost log-linear with R_0 for $\mathbb{Q}_0$ and $\mathbb{Q}_1$ and exactly log-linear with R_0 for $\mathbb{Q}_2$.

In Fig. 4.5, results are plotted for different MIMO systems against K with the same R_0 and ε_0 respectively. The figure shows that less transmit SNR is required for the larger size MIMO systems with the same transmission rate and outage probability constraint. In addition, it can be observed that as K increases (loose delay constraint), the required transmit SNR decreases quite significantly when K is small and decreases slightly when K becomes larger. For example, for $\mathbb{Q}_0$ with $\varepsilon_0 = 0.0001$, the transmit SNR reduces from around 7 dB to 3 dB when K increases from 1 to 5, but it only has around 1 dB deduction when K changes from 5 to 15. This illustrates lager K increases the diversity of the transmission. Fig. 4.6 further illustrates the trade off between delay constraint and the transmit power. We can observe that for achieving the same amount of transmit information bits, i.e., $R_T = R_0 \times K = 12bps/Hz$, and outage probability $\varepsilon_0 = 0.0001$, less total power $(P \times K)$ is required at the cost of larger delay (K).

A special case is shown in Fig. 4.7 for SISO system under scheme $\mathbb{Q}_0$, when setting $n_t = n_r = 1$. Similar with MIMO systems, the transmit SNR increases with the increasing requirement of outage probability ε_0 and transmission rate R_0. In addition, it can be observed that the transmit SNR is more sensitive with ε_0 than that for MIMO systems.

4.7 Summary

In this chapter we have proposed to minimize the total transmit power required for achieving the information outage probability constraint based on the SCI at the transmitter in MIMO BF channel. Gaussian approximation with mean and variance derived in closed-form has been used to represent the p.d.f. of MIMO channel capacity. Transmit power is minimized using one of nonlinear optimization

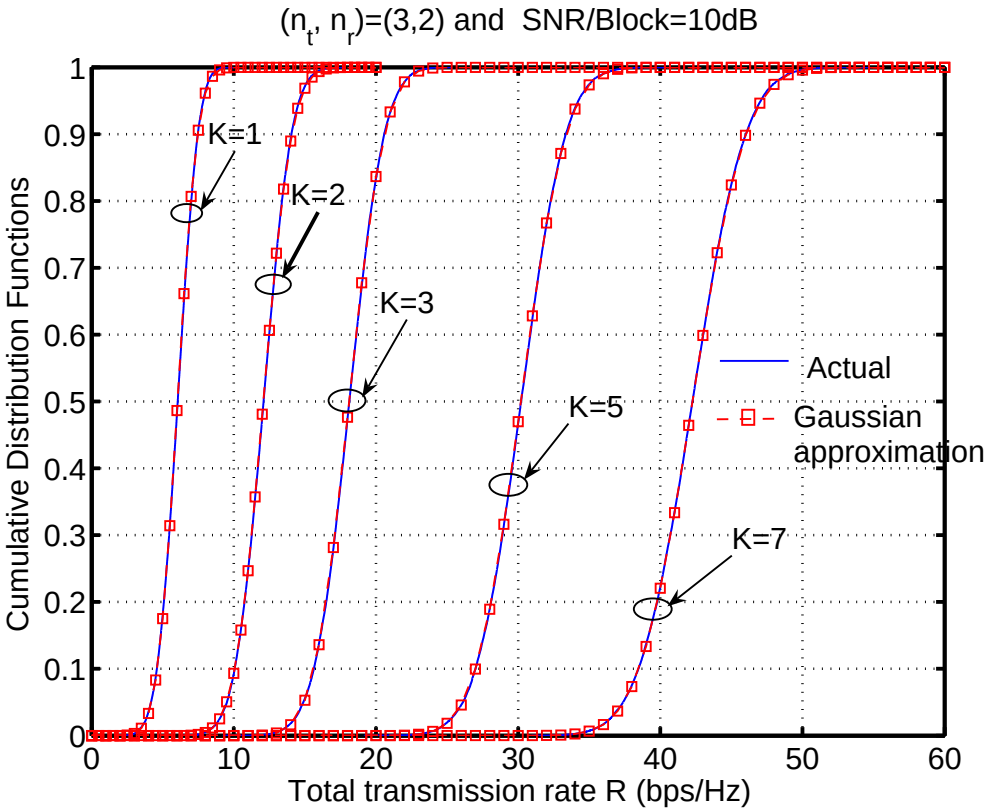

Figure 4.1: Cumulative distribution function for (3,2) MIMO systems with transmit SNR 10dB.

method DIRECT algorithm. Numerical results for different MIMO systems with different information outage probability constraints and target rates have been compared and investigated and the accuracy of Gaussian approximation has been verified.

When transmitter has the perfect knowledge of CSI, power and rate can be dynamically allocated at the transmitter, in which the wireless resource can be used more intelligently. In the following chapter, we will further investigate the scenario when CSI is available at the transmitter in SISO and MIMO systems.

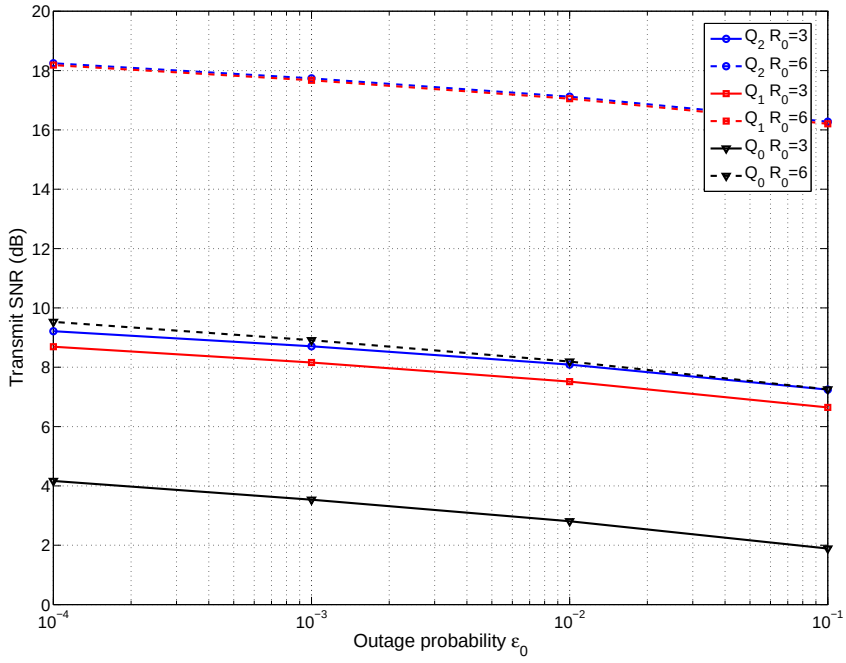

Figure 4.2: Average transmit SNR versus the outage probability for (4,3) systems with block number $K = 3$.

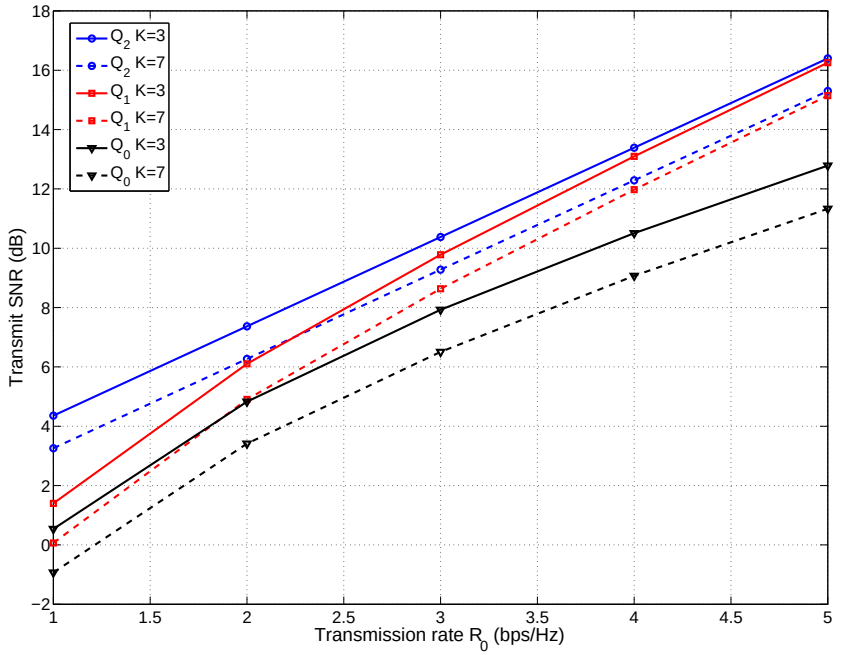

Figure 4.3: Average transmit SNR versus the transmission rate R_0 for a (2,2) MIMO system with ε_0=0.01.

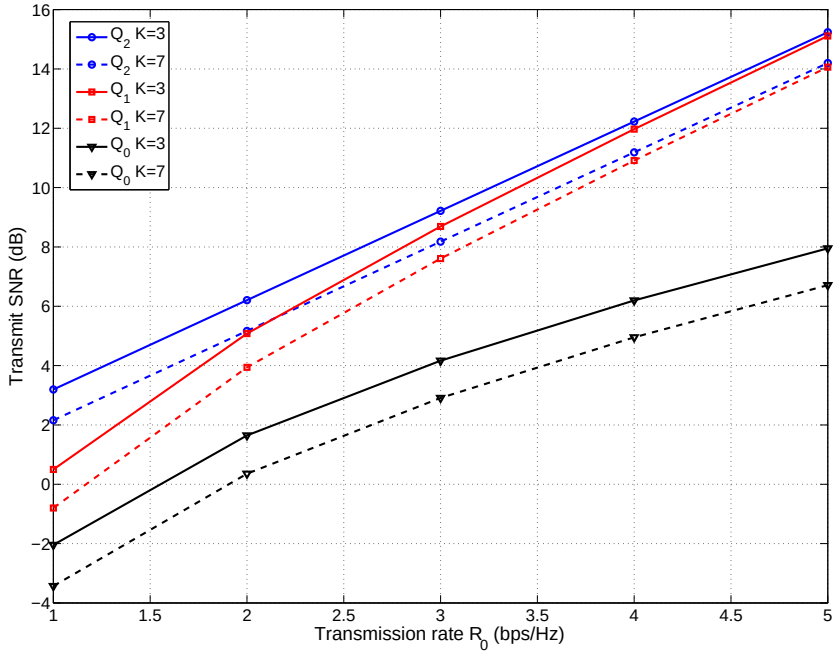

Figure 4.4: Average transmit SNR versus the transmission rate R_0 for a (4,3) MIMO system with ε_0=0.0001.

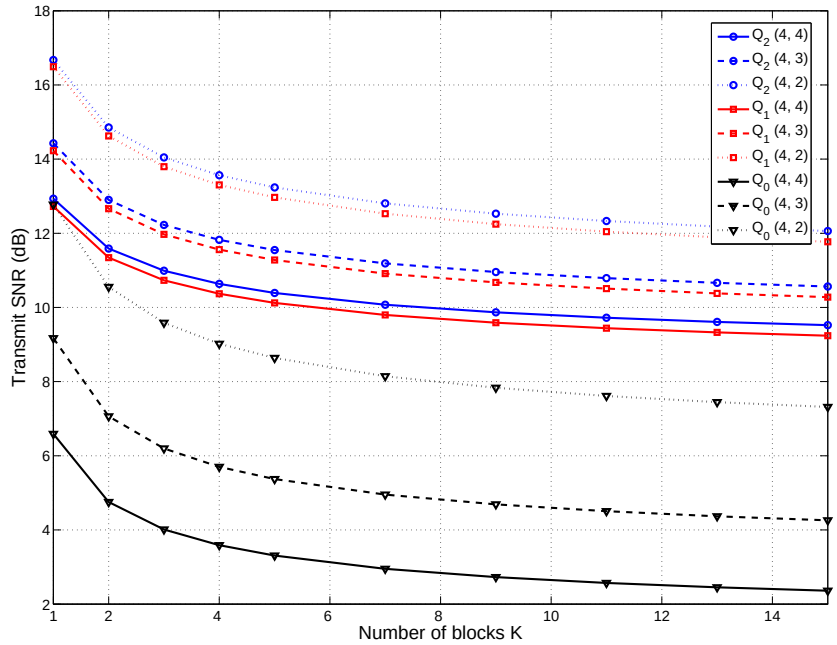

Figure 4.5: Average transmit SNR versus the number of blocks K with $R_0 = 4$ bps/Hz and ε_0=0.0001.

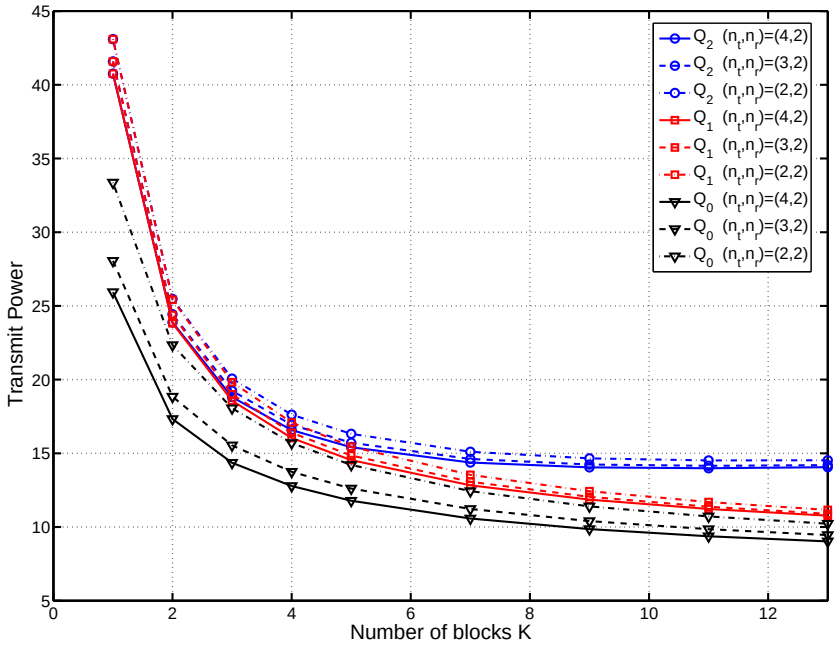

Figure 4.6: Transmit power versus the number of blocks K with $R_T = R_0 \times K = 12$ bps/Hz and ε_0=0.0001.

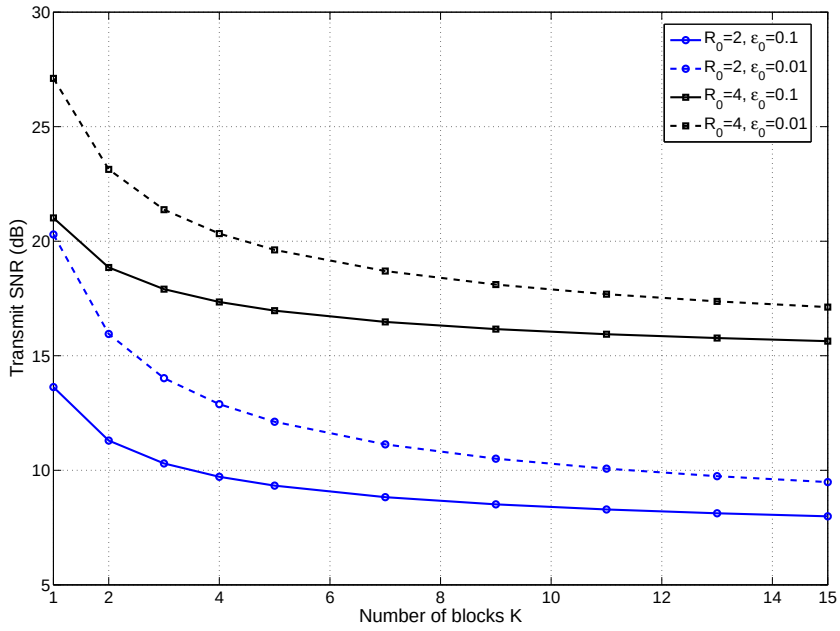

Figure 4.7: Average transmit SNR versus the number of blocks K for SISO systems.

Chapter 5

Wireless Resource Allocation in Single User Systems with Transmitter CSI

In Chapter 5 of the thesis, we will investigate wireless resource allocation problem for satisfying user's information outage probability requirement in single user delay constrained transmission systems with CSI at the transmitter. The mathematical tools used in this chapter are DP and optimal control.

This chapter is organized as follows. Firstly, Section 5.1 briefly overviews the research work in this chapter. Then, system model and problem formulation are presented in Section 5.2. We show that the optimal solution can be obtained by a 3-dimensional DP in Section 5.3. After that, a suboptimal algorithm, which is nearly optimal for small value of outage probability constraints and requires much less complexity, is proposed in Section 5.4. Moreover, both upper and lower performance bounds are developed in Section 5.5. The results are also extended to a MIMO BF channel in Section 5.6. Simulation results are shown in Section 5.7 and Section 5.8 summaries this chapter in the end.

5.1 Introduction

The aim of the chapter is to solve the problem of minimizing the average transmit power of a causal BF channel for meeting a given information outage probability constraint by adapting the power allocation over the blocks with the aid of CSI. The transmitter is assumed to have causal knowledge of CSI, which is exploited to intelligently allocate the rate and power over the blocks to minimize the average transmit power per block for meeting the information outage probability constraint at a given target code rate. The optimal solution to this problem can be obtained by solving the reverse problem of minimizing the information outage probability for a range of long-term power constraints through the use of DP, which is extremely complex. Then, we develop a suboptimal allocation algorithm which still uses DP to exploit the CSI causality but at a much reduced complexity. A performance lower-bound is derived, which permits us to show that the proposed suboptimal algorithm is near-optimal, particularly in the small outage probability regime. A variable-power equal-rate scheme which compromises the performance on the merit of further reducing the complexity is also devised and its analytical performance is derived. The complexities and performance of these algorithms are evaluated and compared, both analytically and numerically. The results are finally generalized to MIMO BF channels.

5.2 System Model and Problem Formulation

A SISO system model is considered (as introduced in Section 2.1), while an extension to MIMO will be discussed in Section 5.6. The channel is assumed to be in BF and fades i.i.d. from one block to another. We use g_k to denote the channel power gain in block k and assume that $\sqrt{g_k}$, is Rayleigh fading so that g_k has the

following p.d.f. shown as

$$f(g_k) = \begin{cases} e^{-g_k} & \text{if } g_k \geq 0, \\ 0 & \text{if } g_k < 0, \end{cases} \tag{5.1}$$

where $\mathtt{E}[g_k] = 1 \ \forall k$.

For a given block, assume the noise power $N_0 = 1$ for normalization, the instantaneous channel capacity of block k allows the following expression

$$r_k = \log_2(1 + p_k g_k) \quad \text{(bps/Hz)}, \tag{5.2}$$

where p_k denotes the transmit power assigned for the current block. Communication is assumed to take place in K blocks at a constant code rate of R_0, so K is regarded as a decoding delay. Hence, the IMI of the K-BF channel is given by

$$\mathcal{I}_K(\boldsymbol{p}, \boldsymbol{g}) = \frac{1}{K} \sum_{k=1}^{K} r_k, \tag{5.3}$$

where $\boldsymbol{p} \triangleq (p_1, p_2, \ldots, p_K)$ and $\boldsymbol{g} \triangleq (g_1, g_2, \ldots, g_K)$.

We assume that both the transmitter and the receiver know perfectly the instantaneous CSI (i.e., g_k at time k), besides the transmitter knows the p.d.f. (5.1). This causal knowledge of CSI is exploited to optimize the power allocation for enhancing the system performance. Consequently, an information outage is declared if $\mathcal{I}_K(\boldsymbol{p}, \boldsymbol{g}) < R_0$.

Under the system model described, we aim at finding a power allocation strategy to minimize the average transmit power per block $P_0 \triangleq \mathtt{E}\left[\frac{1}{K}\sum_{k=1}^{K} p_k\right]$ while controlling the information outage probability to be at most ε_0. That is,

$$\begin{aligned} \min_{\boldsymbol{p} \geq \boldsymbol{0}} \quad & \mathtt{E}\left[\frac{1}{K}\sum_{k=1}^{K} p_k\right] \\ \text{s.t.} \quad & \mathsf{Pr}\left(\mathcal{I}_K(\boldsymbol{p}, \boldsymbol{g}) < R_0\right) \leq \varepsilon_0. \end{aligned} \tag{5.4}$$

Apparently, the minimum P_0 will be achieved when the constraint becomes active, i.e., $\mathsf{Pr}(\mathcal{I}_K(\boldsymbol{p},\boldsymbol{g}) < R_0) = \varepsilon_0$, because $\mathcal{I}_K$ is an increasing function of P_0.

5.3 The Optimal Power Allocation

The causal nature of (5.4) suggests a backward optimization strategy which can be achieved by a 3-dimensional DP [83]. Some of the basics of DP are reviewed in Section 3.3 and we describe the algorithm as follows.

5.3.1 The DP Algorithm for (5.4)

In (5.4), the problem requires to minimize the per-block transmit power averaged over $\boldsymbol{g}$ as cost while the power allocation (or the decision variables $\boldsymbol{p}$) is expected to be done in blocks or stages to balance the present cost and the future expected cost for satisfying the information probability constraint. Clearly, the problem structure fits well with the DP framework, with the only challenge that a probability constraint is involved. To overcome this, we note that the reverse problem of (5.4) is

$$\min_{\boldsymbol{p}\geq\mathbf{0}} \mathtt{E}_{\boldsymbol{g}}\left[1\left(\mathcal{I}(\boldsymbol{p},\boldsymbol{g}) < R_0\right)\right] \quad \text{s.t.} \quad \frac{1}{K}\mathtt{E}\left[\sum_{k=1}^{K} p_k(g_k)\right] \leq P_0, \tag{5.5}$$

where $1(\cdot)$ denotes an indicator function. It is noted that (5.5) has been optimally solved by DP in [90]. For completeness, we here provide the details of the solution.

Clearly, (5.5) can be formulated as a time-lag DP problem (3.20) from Chapter 3 with the augmented system state

$$\tilde{\boldsymbol{x}}_k = \left(R^{(1)}, R^{(2)}, \ldots, R^{(k)}, P^{(1)}, P^{(2)}, \ldots, P^{(k)}, p_1, p_2, \ldots, p_{k-1}, g_1, g_2, \ldots, g_k\right), \tag{5.6}$$

where $R^{(k)}$ denotes the remaining target rate to fulfil from block k to K, and $P^{(k)}$ denotes the average remaining power budget from block k to K. In (5.6), note that g_k is also included in the system state, as it is known and exploited in the optimization. Due to the fact that $(R^{(k)}, P^{(k)})$ contains the information of the previous states $(R^{(1)}, \ldots, R^{(k-1)}, P^{(1)}, \ldots, P^{(k-1)}, p_1, \ldots, p_{k-1}, g_1, \ldots, g_{k-1})$, we can lump the state into the 3-tuple $(g_k, R^{(k)}, P^{(k)})$. The cost (or outage probability) seen at time k is

$$\begin{aligned} c_k(\tilde{x}_k) = \min_{\mu_k(\tilde{x}_k)} \mathtt{E}_{g_k}\left[c_{k+1}(\tilde{x}_{k+1})\right] &= \min_{\mu_k(\tilde{x}_k),\mu_{k+1}(\tilde{x}_{k+1})} \mathtt{E}_{g_k,g_{k+1}}\left[c_{k+2}(\tilde{x}_{k+2})\right] \\ &= \cdots = \min_{\mu_k(\tilde{x}_k),\ldots,\mu_K(\tilde{x}_K)} \mathtt{E}_{g_k,\ldots,g_K}\left[c_{K+1}(\tilde{x}_{K+1})\right] \end{aligned} \quad (5.7)$$

with the initialization $c_{K+1} = 1(R^{(K+1)} > 0)$. The overall cost in (5.5) can then be found as

$$c_1(\tilde{x}_1) = \min_{\boldsymbol{\mu}} \mathtt{E}_{\boldsymbol{g}}[c_{K+1}(\tilde{x}_{K+1})]. \quad (5.8)$$

The optimal policy μ_k^* can be derived by using the principle of optimality (3.18) that proceeds backward from $k = K$ to 1 and solves

$$\begin{aligned} & c_k(\tilde{x}_k) = \min_{\mu_k(\tilde{x}_k)} \mathtt{E}_{g_k}\left[c_{k+1}(\tilde{x}_{k+1})\right] \\ & = \min_{\mu_k(\tilde{x}_k)} \mathtt{E}_{g_k}\left[c_{k+1}(R^{(k+1)}, P^{(k+1)})\right] \\ & = \min_{\substack{\mu_k(g_k,R^{(k)},Q^{(k)})\geq 0 \\ Q^{(k)}\geq 0 \\ E[Q^{(k)}]=P^{(k)}}} \mathtt{E}_{g_k}\left[c_{k+1}\left(R^{(k)} - \log_2\left(1 + \frac{\mu_k(g_k, R^{(k)}, Q^{(k)})g_k}{N_0}\right),\right.\right. \\ & \left.\left. Q^{(k)} - \mu_k(g_k, R^{(k)}, Q^{(k)})\right)\right], \end{aligned} \quad (5.9)$$

where $Q^{(k)}$ denotes the short-term power constraint for stages k to K. In the above, we see that for a given $(g_k, R^{(k)}, Q^{(k)})$, the optimal control policy μ_k^* can be found by

$$\mu_k^*(g_k, R^{(k)}, Q^{(k)}) = \arg\min_{\mu_k} c_{k+1}\left(R^{(k)} - \log_2\left(1 + \frac{\mu_k g_k}{N_0}\right), Q^{(k)} - \mu_k\right). \quad (5.10)$$

Now, in order to solve (5.9), we can then substitute $\mu_k^*(\tilde{x}_k)$ into the equation and give

$$c_k(\tilde{x}_k) = \min_{\substack{Q^{(k)}(g_k)\geq 0 \\ \mathtt{E}_{g_k}[Q^{(k)}]=P^{(k)}}} \mathtt{E}_{g_k}\left[c_{k+1}\left(R^{(k)} - \log_2\left(1+\mu_k^*(\tilde{x}_k)g_k\right), Q^{(k)} - \mu_k^*(\tilde{x}_k)\right)\right]. \tag{5.11}$$

In (5.11), the minimization is done by obtaining the optimal function $Q^{(k)}(g_k)$ for a given pair of $(R^{(k)}, P^{(k)})$. Solving this requires a deterministic DP, which samples the channel instantiation g_k into $\mathcal{G}$ samples (for some large $\mathcal{G}$) and proceeds to derive the required power allocation $Q^{(k)}$ from one sample to the next, subject to

$$\mathtt{E}_{g_k}[Q^{(k)}(g_k)] = \frac{1}{\mathcal{G}}\sum_{j=1}^{\mathcal{G}} Q^{(k)}(g_j) = P^{(k)}. \tag{5.12}$$

This DP aims to find the optimal control policy $\kappa_j^*(s_j)$ according to the system state $s_j = (g_j, u^{(j)})$ which involves g_j (the possible channel realization) and $u^{(j)}$ [the remaining power budget left for $Q^{(k)}(g_j), Q^{(k)}(g_{j+1}), \ldots, Q^{(k)}(g_{\mathcal{G}})$ to fulfill the long-term power constraint (5.12)]. The cost for the jth sample is defined as

$$\delta_j(s_j, \kappa_j(s_j)) = c_{k+1}\left(R^{(k)} - \log_2\left(1 + k\mu_k^*\left(g_j, R^{(k)}, \kappa_j(s_j)\right)g_j\right), \kappa_j(s_j) - \mu_k^*\left(g_j, R^{(k)}, \kappa_j(s_j)\right)\right). \tag{5.13}$$

From the basic DP, we now have the minimum accumulated cost for the tail samples from j to $\mathcal{G}$ as

$$J_j(s_j) = \min_{0\leq\kappa_j(s_j)\leq u^{(j)}} \left[\delta_j(s_j, \kappa_j(s_j)) + J_{j+1}(s_{j+1})\right] \tag{5.14}$$

with the initialization $J_{\mathcal{G}+1} = 0$. Running through the above recursive DP leads to the optimal policy $\kappa_j^*(s_j)$, from which we can obtain the optimal function $Q^{(k)}(g_j)$ by

$$Q^{(k)}(g_j, R^{(k)}, P^{(k)}) = \kappa_j^*(g_j, u^{(j)}) \tag{5.15}$$

in which $u^{(j)}$ is determined by the constraint initialization $u^{(1)} = \mathcal{G}P^{(k)}$ and

$$u^{(j+1)} = u^{(j)} - \kappa_j^*(s_j), \quad \text{for } 1 \leq j < \mathcal{G}. \tag{5.16}$$

As a result, so far, we have covered all the calculations needed to evaluate (5.9) recursively, or the function $c_k(\tilde{x_k})$ can be considered to be known. The DPs that determine $\boldsymbol{\mu}^*$ and $\boldsymbol{\kappa}^*$ need to be run for every possible system states $\tilde{x}_k$, s_j and $Q^{(k)}$, which requires discretization for $0 \leq g_k < \infty$, $0 \leq R^{(k)} \leq KR_0$, $0 \leq P^{(k)} < \infty$, $0 \leq Q^{(k)} < \infty$ and $0 \leq u^{(j)} < \infty$.

Now, let us discuss how (5.4) can be solved based on the results of the DPs. First, it is important to note that $c_1(g_1, R^{(1)}, P^{(1)})$ returns the minimum achievable outage probability of the K-block transmission using the optimal policy $\boldsymbol{\mu}^*$ and is in fact independent of g_1 by definition. With the target rate requirement R_0 (or $R^{(1)} = KR_0$), we can then find the optimal remaining power $P^{(1)}$ from the outage probability requirement ε_0, i.e.,

$$c_1(KR_0, P^{(1)}) = \varepsilon_0. \tag{5.17}$$

During online, the optimal causal power adaptation for (5.4) chooses, for $k = 1$ to K,

$$\text{Online Adaptation} \begin{cases} p_k\left(g_k\right) = \mu_k^*\left(g_k, R^{(k)}, Q^{(k)}\left(g_k, R^{(k)}, P^{(k)}\right)\right), \\ P^{(k)} = Q^{(k-1)}\left(g_{k-1}, R^{(k-1)}, P^{(k-1)}\right) - p_{k-1}\left(g_{k-1}\right), \\ R^{(k)} = KR_0 - \sum_{i=1}^{k-1} \log_2\left(1 + \dfrac{p_i\left(g_i\right) g_i}{N_0}\right), \\ R^{(1)} \equiv KR_0 \qquad \text{(initialization)}, \end{cases} \tag{5.18}$$

where $\mu_k^*(\cdots)$ and $Q^{(k)}(\cdots)$ are directly accessible from the results of the DPs.

5.3.2 Complexity Analysis

The complexity of the optimal DP algorithm is determined by the calculations of $\mu_k^*(g_k, R^{(k)}, Q^{(k)})$, $Q^{(k)}(g_k, R^{(k)}, P^{(k)})$ and $c_k(g_k, R^{(k)}, P^{(k)})$, which depend on the sizes of the discretized spaces for the parameters g_k, $R^{(k)}$, $Q^{(k)}$ and $P^{(k)}$. Let $\mathcal{G}$ and $\mathcal{R}$ denote, respectively, the discretization spaces for g_k and $R^{(k)}$. Likewise, we denote the sizes of the discretization spaces for $P^{(k)}$ and $Q^{(k)}$ as $\mathcal{P}_k$ and $\mathcal{Q}_k$, respectively. In particular, we write $\mathcal{P}_1 = \mathcal{P}$ for the range from 0 to KP_0 and then we have $\mathcal{Q}_1 = \mathcal{GP}$. In what follows, we also have $\mathcal{P}_k = \mathcal{G}^{k-1}\mathcal{P}$ and $\mathcal{Q}_k = \mathcal{G}^k\mathcal{P}$. As a result, the required complexity can be characterized as follows:

- Time complexity—The complexity required for computing $\mu_k^*(\cdots)$, $\mathcal{C}_{\mu_k}$, is given by

$$\mathcal{C}_{\mu_k} = \mathcal{O}\left(\sum_{i=1}^{\mathcal{Q}_k} i\mathcal{GR}\right) = \mathcal{O}\left(\mathcal{GR}\frac{\mathcal{Q}_k(1+\mathcal{Q}_k)}{2}\right) \approx \mathcal{O}(\mathcal{GRQ}_k^2). \tag{5.19}$$

  Also, obtaining $Q^{(k)}(\cdots)$ requires a two-level DP. For each $(R^{(k)}, P^{(k)})$, the inner DP has the complexity of $\mathcal{O}(\mathcal{G}\sum_{i=1}^{\mathcal{GP}_k} i)$. If this calculation is done over the maximum possible range of $P^{(k)}$, then this complexity reflects the overall complexity of the inner DP for each specific $R^{(k)}$. As such, the required complexity for $Q^{(k)}(\cdots)$, $\mathcal{C}_{Q^{(k)}}$, can be found as

$$\mathcal{C}_{Q^{(k)}} = \mathcal{O}\left(\sum_{i=1}^{\mathcal{GP}_k} i\mathcal{RG}\right) = \mathcal{O}\left(\mathcal{GR}\frac{\mathcal{P}_k\mathcal{G}(1+\mathcal{P}_k\mathcal{G})}{2}\right) \approx \mathcal{O}(\mathcal{G}^3\mathcal{RP}_k^2). \tag{5.20}$$

  Finally, the computation of $c_k(\cdots)$ has the complexity of $\mathcal{O}(\mathcal{RP}_k)$. Summarizing this, we have the overall complexity for the optimal DP algorithm as

$$\begin{aligned}\mathcal{C}_{\mathsf{time}}^{\mathsf{Opt-DP}} &= \sum_{k=1}^{K}\left[\mathcal{O}\left(\mathcal{GR}(\mathcal{G}^k\mathcal{P})^2\right) + \mathcal{O}\left(\mathcal{G}^3\mathcal{R}(\mathcal{G}^{n-1}\mathcal{P})^2\right) + \mathcal{O}\left(\mathcal{RG}^{k-1}\mathcal{P}\right)\right] \\ &\approx \mathcal{O}\left(\mathcal{G}^{2K+1}\mathcal{RP}^2\right).\end{aligned} \tag{5.21}$$

- Storage complexity—The algorithm requires storage for the matrices $\mu_k^*(\cdots)$, $Q^{(k)}(\cdots)$ and $c_k(\cdots)$. They are, respectively, $\mathcal{O}(\mathcal{GRQ}_k)$, $\mathcal{O}(\mathcal{GRP}_k)$ and $\mathcal{O}(\mathcal{RP}_k)$. As a result, we have

$$\mathcal{C}_{\mathsf{storage}}^{\mathsf{Opt-DP}} = \sum_{k=1}^{K} \left[\mathcal{O}\left(\mathcal{GRG}^k\mathcal{P}\right) + \mathcal{O}\left(\mathcal{GRG}^{k-1}\mathcal{P}\right) + \mathcal{O}\left(\mathcal{RG}^{k-1}\mathcal{P}\right)\right] \approx \mathcal{O}\left(\mathcal{G}^{K+1}\mathcal{RP}\right). \tag{5.22}$$

5.4 The Proposed Suboptimal Algorithm

5.4.1 Hyper DP with Forward Decision Per Block

Apparently, the required complexity to obtain the optimal solution is prohibitively complex, especially when a high-resolution solution is sought (which requires large spaces for $\mathcal{G}$, $\mathcal{R}$ and $\mathcal{P}$). In this subsection, we develop an algorithm which exploits the use of DP to obtain a near-optimal power allocation solution but at much less complexity. To do so, we first define an outage region $\Omega(\boldsymbol{p}) = \left\{\boldsymbol{g} : \frac{1}{K}\sum_{k=1}^{K} r_k(p_k) < R_0\right\}$ and the non-outage region $\Omega_{\mathsf{c}}(\boldsymbol{p}) = \left\{\boldsymbol{g} : \frac{1}{K}\sum_{k=1}^{K} r_k(p_k) \geq R_0\right\}$, while we have $\Omega^* = \Omega(\boldsymbol{p}^*)$ and $\Omega_{\mathsf{c}}^* = \Omega_{\mathsf{c}}(\boldsymbol{p}^*)$ where $\boldsymbol{p}^*$ denotes the optimal power allocation. Technically, to minimize the average transmit power per block P_0, the algorithm needs to identify whether a particular channel is in Ω^* as soon as possible to avoid further power is wasted. On the other hand, if $\boldsymbol{g} \in \Omega_{\mathsf{c}}^*$, minimum P_0 occurs when $\mathcal{I} = R_0$ is met. To summarize, (5.4) can be solved by knowing Ω^* and Ω_{c}^*. If $\boldsymbol{g} \in \Omega_{\mathsf{c}}^*$, then the optimal causal power allocation can be found by

$$\min_{\boldsymbol{p}} \frac{1}{K}\mathtt{E}\left[\sum_{k=1}^{K} p_k\right] \quad \text{s.t.} \quad \frac{1}{K}\sum_{k=1}^{K} r_k = R_0, \tag{5.23}$$

while if $\boldsymbol{g} \in \Omega^*$, solves

$$\min_{\boldsymbol{p}} \frac{1}{K}\mathtt{E}\left[\sum_{k=1}^{K} p_k\right] \tag{5.24}$$

which gives a zero-power solution when the channel is identified to be in an outage region.

In light of this, we propose an algorithm, termed hyper DP (HDP) with forward decision per block (FDPB). This algorithm combines the optimizations (5.23) and (5.24), and in particular, consists of two sub-algorithms: 1) HDP which allocates minimum power over the blocks for achieving an information rate of R_0 and 2) FDPB which decides, on the fly, whether a given channel is in Ω^*, given limited observations $\boldsymbol{g}^{(k)} \triangleq (g_1, \ldots, g_k)$. If $\boldsymbol{g}^{(k)}$ is identified to be in Ω^*, the transmission is stopped and $\boldsymbol{p} = \mathbf{0}$. On the other hand, if $\boldsymbol{g}^{(k)}$ is not yet known to be in Ω^*, then transmission continues and the power allocation is done as if a rate of R_0 is to be achieved.

To solve (5.23) with the casual CSI, HDP is performed by constructing a standard DP, which has the system state $y_k = R^{(k)}$, the (channel) random disturbance g_k, the decision control $v_k(g_k, R^{(k)})$ and the cost at stage k as $\tau_k(y_k) = \frac{1}{g_k}(2^{v_k} - 1)$. The optimal control, or rate adaptation, $v_k^*(g_k, R^{(k)})$, can therefore be obtained by choosing, for $k = K, K-1, \ldots, 1$,

$$v_k^*(g_k, R^{(k)}) = \arg \min_{0 \le v_k(g_k, R^{(k)}) \le R^{(k)}} \left\{ \frac{1}{g_k} (2^{v_k} - 1) + S_{k+1}(R^{(k+1)}) \right\}, \tag{5.25}$$

where $S_k(R^{(k)})$ denotes the total expected cost from block k to K given a remaining rate $R^{(k)}$, with the initialization $S_{K+1}(R^{K+1}) = 0$, and is given by

$$S_k(R^{(k)}) = \mathtt{E}_{g_k} \left[\frac{1}{g_k} (2^{v_k} - 1) + S_{k+1}(R^{(k+1)}) \right]. \tag{5.26}$$

This backward calculation is carried out off-line to obtain $v_k^*(g_k, R^{(k)})$, which is stored for use in the online adaptation. When the system is online, the transmitter

will adapt the power allocation to meet the rate by

$$\text{Online Adaptation} \begin{cases} r_k\left(g_k\right) = v_k^*\left(g_k, R^{(k)}\right), \\ R^{(k)} = KR_0 - \sum_{i=1}^{k-1} r_i\left(g_i\right), \\ R^{(1)} \equiv KR_0, \\ p_k = \dfrac{1}{g_k}\left(2^{r_k} - 1\right). \end{cases} \tag{5.27}$$

The remaining task is to detect if a channel is in Ω^*, which is facilitated by FDPB.[1] In particular, as by-products of the (off-line) DP (5.25), we obtain the following matrices

$$\begin{cases} \chi_k(g_k, R^{(k)}) = 1\left(R - v_k^*(g_k, R^{(k)})\right), \\ \varphi_k(R^{(k)}) = \mathtt{E}_{g_k}[\chi_k], \end{cases} \tag{5.28}$$

where $\varphi_k(R^{(k)})$ estimates the probability that there is still a finite remaining rate at time k to be allocated for the future blocks to fulfill the required rate. In other words, $\varphi_k(R^{(k)})$ indicates the conditional outage probability seen at time k if the transmission has not terminated before block k. As such, if no more power is spent on blocks $k+1$ to K, the outage probability seen at time k is

$$\varepsilon^{(k)} = \varphi_1\left(R^{(1)}\right)\varphi_2\left(R^{(2)}\right)\cdots\varphi_n\left(R^{(k)}\right). \tag{5.29}$$

If $\varepsilon^{(k)} \leq \varepsilon_0$, then there is no need to allocate power in the future blocks because the outage probability requirement has been fulfilled. In fact, if $\varepsilon^{(k)} < \varepsilon_0$, the requirement is over-satisfied and further power reduction can be achieved by revising the probability at the current block k as

$$\tilde{\varphi}_k\left(R^{(k)}\right) = \frac{\varepsilon_0}{\varepsilon^{(k-1)}} \tag{5.30}$$

[1]Note that outage detection for a channel, in theory, requires to use the complex optimal DP strategy. Therefore, the detection by FDPB is suboptimal and incorrect detection may be made.

so that $\varepsilon^{(k)} = \varphi_1 \cdots \varphi_{k-1}\tilde{\varphi}_k = \varepsilon_0$. This probability changes result in the revision of the entries of $\tilde{\chi}_k(g_k, R^{(k)})$ to have more ones than $\chi_k(g_k, R^{(k)})$ (see Fig. 5.1 for details), which provides a new channel threshold g_k^B. If $g_k > g_k^B$, then the channel is considered to be useful and no outage occurs, i.e., $\boldsymbol{g}^{(k)} \in \Omega_{\mathsf{c}}^*$. Otherwise, $\boldsymbol{g}^{(k)} \in \Omega^*$ and $r_k = p_k = 0$. On the other hand, if $\varepsilon^{(k)} > \varepsilon_0$, the power adaptation process continues for the next block. The overall proposed algorithm, which uses HDP and FDPB, is summarized as follows:

1. For $k = 1, \ldots, K-1$, calculate $\varepsilon^{(k)} = \varphi_1\varphi_2 \cdots \varphi_k$ with the initialization $\varepsilon^{(0)} = 1$.

2. If $\varepsilon^{(k)} > \varepsilon_0$, the power for block k is given by (5.27) and will continue in the next block $k+1$.

3. If $\varepsilon^{(k)} = \varepsilon_0$, the power for block k is given by (5.27) and the power adaptation process terminates.

4. If $\varepsilon^{(k)} < \varepsilon_0$, the probability constraint is over-satisfied. Then update $\tilde{\varphi}_k$ to $\frac{\varepsilon_0}{\varepsilon^{(k-1)}}$ to obtain a new channel threshold g_k^*. The power for block n is then found by (5.27) with the threshold g_k^*. The power adaptation process ends as the constraint has been met.

5. If the transmission reaches the last block, i.e., $k = K$, then set $\tilde{\varphi}_K = \frac{\varepsilon_0}{\varepsilon^{(K-1)}}$ and allocate the power using (5.27) accordingly. The power adaptation process has completed.

5.4.2 Complexity Analysis

Given the sizes for the discretization spaces for $R^{(k)}$ and g_k as $\mathcal{R}$ and $\mathcal{G}$, the complexity of the proposed algorithm can be analyzed and summarized as follows.

- Time complexity—The complexities for the required calculations are

$$\begin{cases} \mathcal{C}_{v_k} = \mathcal{O}\left(\mathcal{G}\sum_{i=1}^{\mathcal{R}} i\right) = \mathcal{O}\left(\frac{\mathcal{G}(1+\mathcal{R})\mathcal{R}}{2}\right), \\ \mathcal{C}_{S_k} = \mathcal{O}(\mathcal{R}), \\ \mathcal{C}_{\chi_k} = \mathcal{O}(\mathcal{G}\mathcal{R}), \\ \mathcal{C}_{\varphi_k} = \mathcal{O}(\mathcal{R}). \end{cases} \tag{5.31}$$

  As a result, the overall computational complexity of the proposed method is

$$\mathcal{C}_{\mathsf{time}}^{\mathsf{HDP-FDPB}} = \mathcal{O}\left(K\mathcal{G}\mathcal{R}^2\right). \tag{5.32}$$

- Storage complexity—The algorithm needs the storage complexities for $v_k^*(\cdots)$, $S_k(\cdot)$, $\chi_k(\cdot)$ and $\varphi_k(\cdot)$, which are, respectively, given by $\mathcal{O}(\mathcal{G}\mathcal{R})$, $\mathcal{O}(\mathcal{R})$, $\mathcal{O}(\mathcal{G}\mathcal{R})$, and $\mathcal{O}(\mathcal{R})$. Accordingly,

$$\mathcal{C}_{\mathsf{storage}}^{\mathsf{HDP-FDPB}} = \mathcal{O}(K\mathcal{G}\mathcal{R}). \tag{5.33}$$

As we can see, there is a significant reduction in complexity, as compared to the optimal DP algorithm. In particular, the complexity of HDP-FDPB increases only linearly with K whereas the complexity for the optimal DP algorithm grows exponentially with K (the codeword length).

5.4.3 Asymptotic-Optimality

The proposed algorithm is suboptimal in general because the transmitter may not be able to realize that the channel is in the outage region at the beginning of the first few blocks. In other words, there is a chance that the channel is indeed in an outage region but this is not known until some power is wasted in the previous blocks. In addition, the outage region learned by FDPB is in general different from

that of the optimal DP solution, which contributes to a slight performance loss as well. Arguably however, the performance gap should be insignificant if ε_0 is small. The reason is that when $\varepsilon_0 = 0$, the outage regions for the optimal DP solution and the proposed HDP-FDPB algorithm are empty and hence all the channels are considered to be in non-outage. As the power allocation strategies for non-outage channels for both methods are the same, they will perform equally in this case. As a result, the proposed HDP-FDPB algorithm is asymptotic-optimal. To help demonstrate the effectiveness, some bounds will also be considered in Section 5.5.

5.5 Performance Bounds

5.5.1 Lower Bound

A lower bound for the power minimization problem (5.4) can be easily obtained by the method that always assigns the least power over the blocks to meet the rate of R_0 if the channel is not in outage, and always assigns zero power if the channel belongs to the outage region (i.e., artificially setting $P_\Omega = 0$). Such method may not exist, as in practice, the transmitter will have to learn from the CSI to tell if the channel is in the outage region, and it will take several channel blocks to do so (some power would have to be wasted). Such lower bound can be derived by HDP, which retrospectively assigns zero power to all the blocks if the channel turns out to be in the outage region.

5.5.2 A Simpler Closed-Form Method

To further reduce the complexity of power adaptation, a simpler solution which does not require DP is possible, which we present here in this subsection. The method is referred to as equal-outage-probability per block (EOPPB) and is based

on the idea that converts the outage probability of a K-BF channel, ε_0, into a per-block outage probability constraint $\tilde{\varepsilon}_0$. This is done by setting

$$\tilde{\varepsilon}_0 = 1 - (1 - \varepsilon_0)^{\frac{1}{K}}. \tag{5.34}$$

This probability $\tilde{\varepsilon}_0$ will automatically set the outage region (or effectively the channel threshold) for each block, g^B, by

$$\tilde{\varepsilon}_0 = 1 - e^{-g^B} \quad \Rightarrow \quad g^B = -\ln(1 - \tilde{\varepsilon}_0). \tag{5.35}$$

The power adaptation strategy, in block n, is therefore

$$p_k(g_k, R_0) = \begin{cases} \frac{2^{R_0}-1}{g_k} & \text{if } g_k \geq g^B, \\ 0 & \text{if } g_k < g^B, \end{cases} \tag{5.36}$$

if all the previous blocks $m < k$ are not in outage (or $g_m > g^B \ \forall m < k$). Note also that, in this method, when the channel block is not in outage, the power is allocated such that the instantaneous mutual information for each block meets the target rate R_0.

To find the average transmit power per block of EOPPB, P_0^{EOPPB}, we first obtain

$$\mathtt{E}_{g_k}[p_k] = \int_{g^B}^{\infty} \left(\frac{2^{R_0} - 1}{g_k} \right) df(g_k) = \left(2^{R_0} - 1\right) E_1(-\ln(1 - \tilde{\varepsilon}_0)). \tag{5.37}$$

Consequently, we have

$$P_0^{\mathsf{EOPPB}} = \frac{1}{K} \mathtt{E}_{g_k}[p_k] \left[1 + (1 - \tilde{\varepsilon}_0) + \cdots + (1 - \tilde{\varepsilon}_0)^{K-1}\right], \tag{5.38}$$

which can further be simplified to

$$\begin{aligned} P_0^{\mathsf{EOPPB}} &= \mathrm{E}_{g_k}[p_k]\frac{\varepsilon_0}{K\tilde{\varepsilon}_0} \\ &= \frac{\left(2^{R_0}-1\right)\varepsilon_0}{K\left[1-(1-\varepsilon_0)^{\frac{1}{K}}\right]} E_1\left(-\frac{1}{K}\ln(1-\varepsilon_0)\right). \end{aligned} \tag{5.39}$$

Note that EOPPB offers a closed-form solution (5.36). As a result, it is very computational desirable but numerical results will show that the performance may be severely compromised.

5.5.3 Optimal Allocation with Acausal CSI

In Section 2.1.3, the optimal power allocation policy for an acausal BF channel is introduced, which gives a performance lower bound for P_0 for the causal BF channels. In this case, with a long-term power constraint, P_0^{NC}, the minimum outage probability can be expressed as

$$\varepsilon = \int_{\Omega(R_0,\boldsymbol{g}^B)} d\mathcal{F}(\boldsymbol{g}), \tag{5.40}$$

where $\Omega(R_0,\boldsymbol{g}^B)$ is the outage region and $\mathcal{F}(\boldsymbol{g})$ denotes the joint c.d.f. of $\boldsymbol{g}$ (see [64] for details). In other words, by setting $\varepsilon=\varepsilon_0$, the channel outage region Ω (or $\boldsymbol{g}^B$) can be obtained and hence the minimum power, P_0^{NC}, to achieve ε_0 can be found.

5.6 Extension to MIMO BF Channels

5.6.1 The Power-Rate Relationship

In this section, we extend our results to a MIMO BF channel. In particular, we consider a MIMO channel where the magnitude square of each channel element

follows the p.d.f. of (5.1).

For a particular block k, as the instantaneous CSI is assumed known, the optimal power allocation strategy over the eigenchannelscan be obtained by waterfilling for a given short-term power budget p_k (if p_k is allocated for this block). Thus, the rate achieved for this block is given by

$$r_k = \sum_{\substack{i=1 \\ i:\{\lambda_i > \delta\}}}^{\mathsf{rank}(\mathbf{H}_k)} \log_2 \left(\frac{\lambda_i}{\delta} \right) \tag{5.41}$$

where $\mathbf{H}_k$ denotes the MIMO channel at time k, $\mathsf{rank}(\cdot)$ returns the rank of the input matrix, $\{\lambda_i\}$ denote the eigenvalues of $\mathbf{H}_k^\dagger \mathbf{H}_k$, and the water-level δ is the solution of

$$p_k = \sum_{i=1}^{\mathsf{rank}(\mathbf{H}_k)} \left(\frac{1}{\delta} - \frac{1}{\lambda_i} \right)^+ . \tag{5.42}$$

This result can also be interpreted as the way to find the minimum required transmit power p_k for attaining an instantaneous mutual information of r_k for block k.

Following previous description of DP, we have now the functions $r_k(p_k, \mathbf{H}_k)$ and $p_k(r_k, \mathbf{H}_k)$ for MIMO channels. Adopting the same notation $\boldsymbol{p}$ and also $\boldsymbol{\mathcal{H}} \triangleq (\mathbf{H}_1, \mathbf{H}_2, \ldots, \mathbf{H}_K)$, the aim is to

$$\min_{\boldsymbol{p}} \mathsf{E} \left[\frac{1}{K} \sum_{k=1}^{K} p_k \right] \quad \text{s.t.} \quad \mathsf{Pr} \left(\mathcal{I}_K(\boldsymbol{p}, \boldsymbol{\mathcal{H}}) < R_0 \right) \leq \varepsilon_0, \tag{5.43}$$

where $\mathcal{I}(\boldsymbol{p}, \boldsymbol{\mathcal{H}}) = \frac{1}{K} \sum_{k=1}^{K} r_k$ is evaluated using (5.41). This result generalizes the previously described methods to cope with MIMO channels. However, it is worth mentioning that for SISO channels, the DP algorithm involves comparison between the instantaneous channel gain g_k and the channel threshold g^B (determined by both the off-line and on-line calculations). Yet, for MIMO channels, such comparison cannot be easily made. To achieve an ordering for MIMO channels, we rank a large number of random channel matrices by their power required to achieve a

given rate using (5.41), which then allows the DP to decide a threshold for the required power. During online, the required power of a given MIMO channel will be compared to this threshold for power allocation. In the following, we shall address on the required complexity of the DP for MIMO channels.

5.6.2 Complexity of the DP Solutions

Let $\mathcal{H}$ denote the discretization space for $\mathbf{H}_k$ and $\mathcal{W}$ represent the complexity for getting the water-filling solution. We can then determine the time-complexity for the optimal DP solution of MIMO systems as $\mathcal{O}(\mathcal{H}^{2K+1}\mathcal{R}\mathcal{P}^2\mathcal{W})$ and the storage complexity as $\mathcal{O}(\mathcal{H}^{K+1}\mathcal{R}\mathcal{P})$. Similarly with the proposed HDP-FDPB algorithm, it can be easily shown that the time-complexity can be derived as $\mathcal{O}(K\mathcal{H}\mathcal{R}^2\mathcal{W})$, while the storage complexity can be obtained as $\mathcal{O}(K\mathcal{H}\mathcal{R})$. Again, a significant complexity reduction is observed as compared to the optimal DP solution.

5.7 Simulation Results

Simulation results for different algorithms are presented and compared in this section. The results are evaluated based on the expected SNR ($\frac{\mathrm{E}[\sum_{k=1}^{K} p_k]}{K}$). Results in Figs. 5.2, 5.4 and 5.5 are provided for SISO systems, while Figs. 5.6 and 5.7 investigate the results of MIMO BF channels.

As can be seen in Fig. 5.2, the analytical results (5.39) align well with the Monte-Carlo results of EOPPB, confirming the correctness of the results. Results in this figure also demonstrate that for small ε_0, e.g., 0.01, the lower bound and the proposed algorithm perform nearly the same, although there is a difference of about 1 (dB) between the lower bound and the proposed method for the case $\varepsilon_0 = 0.1$. Moreover, we see that when $K = 2$, EOPPB can be promising and gives similar performance as the lower bound. Besides, we observe that for the

acausal solution, the required SNR decreases with K. This reveals that coding over different blocks in an acausal fashion increases the diversity. However, opposite is observed for EOPPB. The reason is because the calculation of EOPPB decomposes outage probability ε_0 into $\widetilde{\varepsilon}_0$ for each block of which the value is less than ε_0 (see Section 5.5 for details). This inevitably leads to the increase of average required SNR.

For the lower bound and the proposed HDP-FDPB algorithm, it is observed that SNR first increases and then decreases with K. The reason can be explained as follows. When the outage requirement is loose, e.g. $\varepsilon_0 = 0.1$, the SNR required mainly lies in low SNR region. Therefore, small number of blocks, such as $K = 1$, is good enough to be used for power allocation compared with larger number of blocks. However, when the outage requirement is stringent, e.g. $\varepsilon_0 = 0.01$, higher value for transmit SNR is required, which leads to the requirement for larger number of blocks, as smaller number of blocks require more power allocation in this case.

In Fig. 5.3, we further illustrate the delay power trade-off. The total transmit power results are plotted against the number of blocks K with the same amount of transmission information $R_T = R_0 \times K = 8$ (bps/Hz) for different outage requirements. The figure indicates with the reduced requirement of delay (larger K), the required transmit power decreases.

Similar conclusion can also be drawn from the results in Fig. 5.4 where $\varepsilon_0 = 0.01$ and $K = 4$ are set, and the SNR results are plotted against the required rate R_0. Results demonstrate that the proposed HDP-FDPB algorithm performs nearly the same as the lower bound, showing that the proposed algorithm is near-optimal and this is particularly the case for small outage probability constraints. It can also be seen that the required SNR increases near log-linearly with the code rate R_0 for all the solutions. Besides, the gap between HDP-FDPB and EOPPB decreases from 5 to 3 (dB) when R_0 increases from 0.5 to 4 (bps/Hz). This indicates that

with high transmission code rates, there is a tendency to allocate similar rates among the blocks. In Fig. 5.5, the SNR results are plotted against ε_0 when $K = 4$ and $R_0 = 2$ (bps/Hz). Results show that the required SNR increases only mildly with the increasing demand for ε_0 for the acausal solution, lower bound and the proposed HDP-FDPB algorithm, but it increases considerably for EOPPB. This can be explained by having a close look at the proposed HDP-FDPB algorithm that the outage probability requirement, ε_0, mainly affects the stopping criterion for outage channel detection and it does not affect the power allocation for the channel blocks. The exact power allocation for the non-outage channel blocks depends only on the target rate R_0.

Figs. 5.6 and 5.7 provide the SNR results against the code rate R_0 for a (3,2) MIMO system for various K and ε_0. It is observed that the required SNR under different schemes increases near log-linearly with the code rate. In addition, it is seen that the results for the proposed HDP-FDPB algorithm are very close to the lower bounds when the outage probability constraint is small, i.e., 0.01. One final point worth highlighting here is that in contrast to the SISO cases, the SNR gaps among different solutions are much smaller, and that the SNR difference between EOPPB and the lower bound is less than 1 (dB). This can be explained by recognizing that the spatial diversity in MIMO systems reduces the need of dynamically allocating the power over the blocks. On the other hand, the SNR results when only SCI is available at the transmitter are plotted for comparison. It is shown that with CSI at the transmitter, transmit SNR can have a remarkable power saving. For example, the SNR gap for $K = 2$ is around 6dB, while the gap for $K = 8$ is around 4dB. This indicates the advantage in power saving when CSI is available at the transmitter compared with only SCI available.

5.8 Summary

In this chapter, the power minimization problem for delay constrained communications has been addressed with transmitter CSI to achieve a certain information outage probability constraint. The problem has been solved optimally by applying high-complexity DP. An efficient solution at much less complexity has also been proposed which is asymptotically optimal for small values of outage probability. A lower bound and a simpler method termed EOPPB have also been developed. Numerical results under different simulation settings have demonstrated the near-optimality of the proposed suboptimal solution.

The research work in Chapter 5 together with the work in Chapter 4 has enlightened the research in resource allocation strategy design for QoS-orientated delay constrained communication in wireless fading channel. In the following chapters, we will further extend the research into multi-user TDMA scenarios, in which joint optimization for both transmit power and time slots are considered.

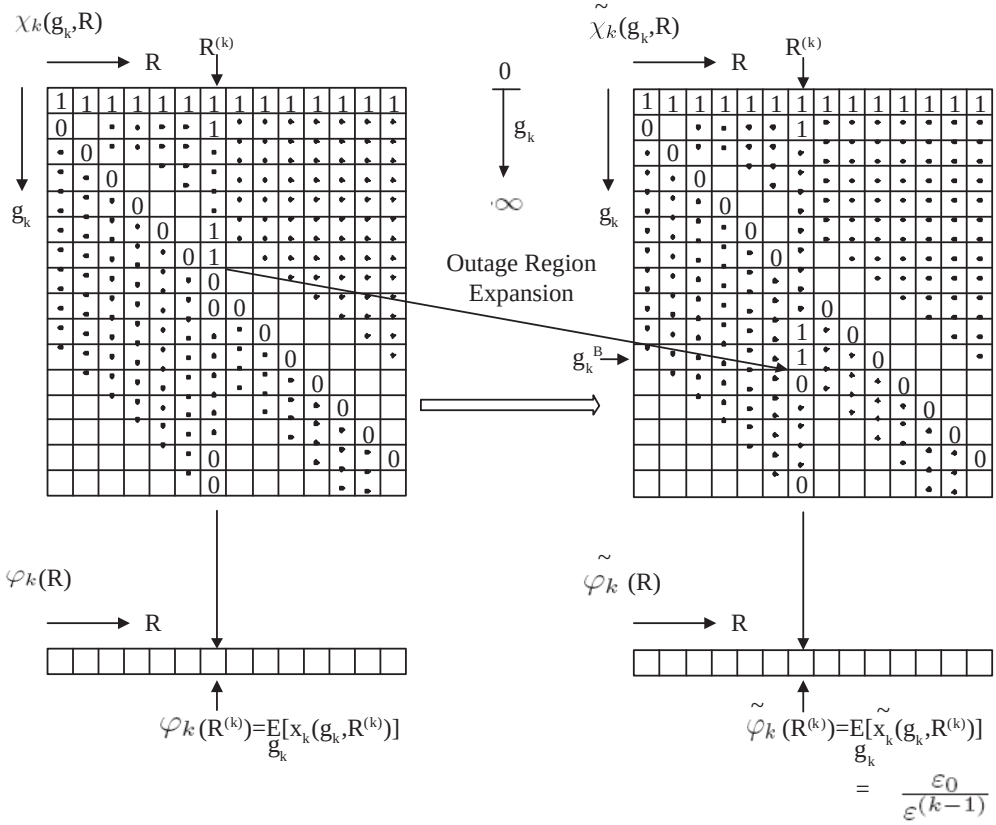

Figure 5.1: An illustration of outage region expansion.

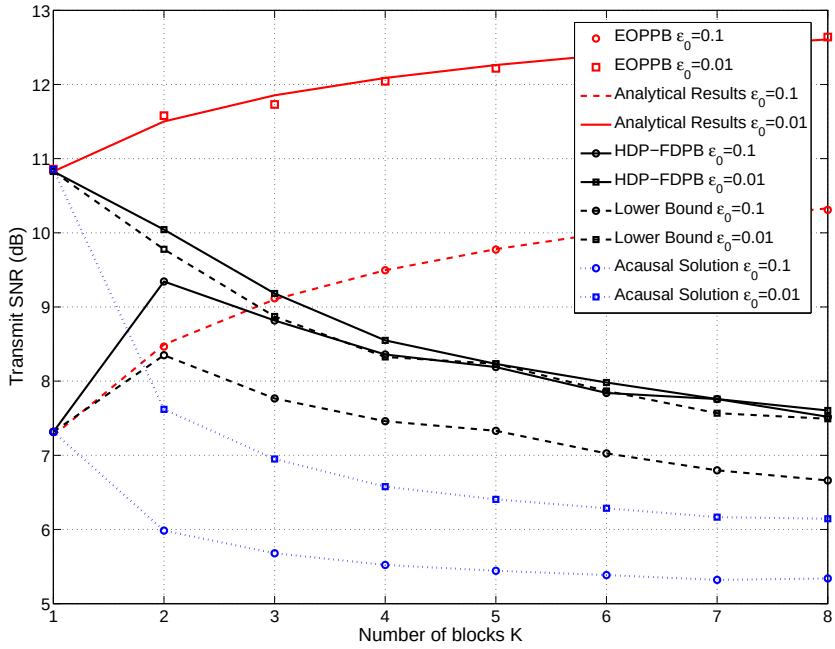

Figure 5.2: Average SNR results versus the number of blocks K with $R_0 = 2$ (bps/Hz).

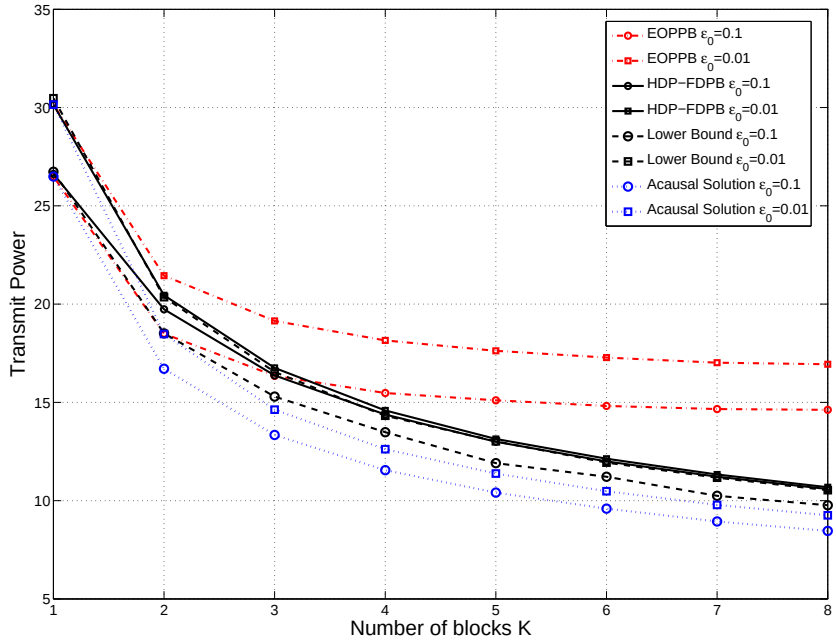

Figure 5.3: Transmit power results versus the number of blocks K with $R_T = R_0 \times K = 8$ (bps/Hz).

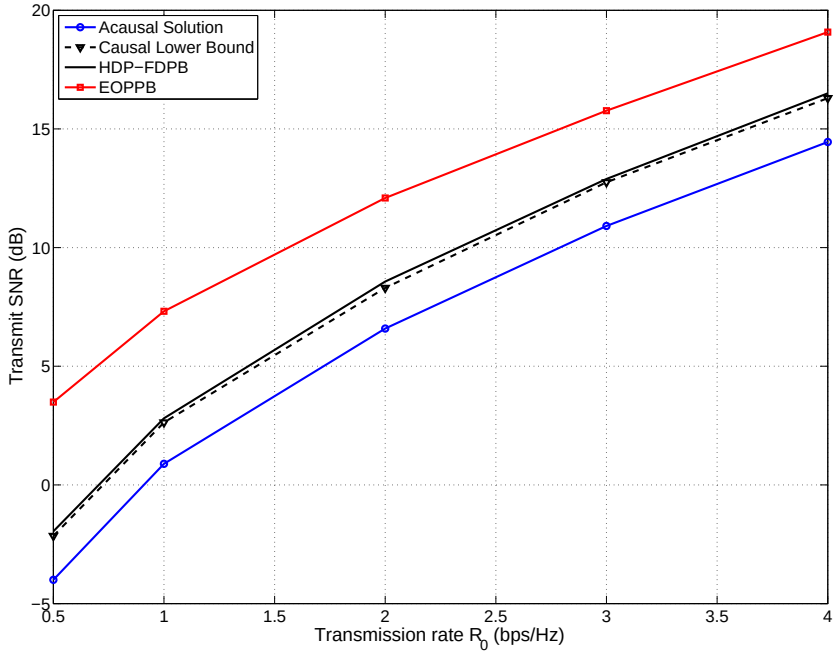

Figure 5.4: SNR results versus the code rate R_0 with $\varepsilon_0 = 0.01$ and $K = 4$.

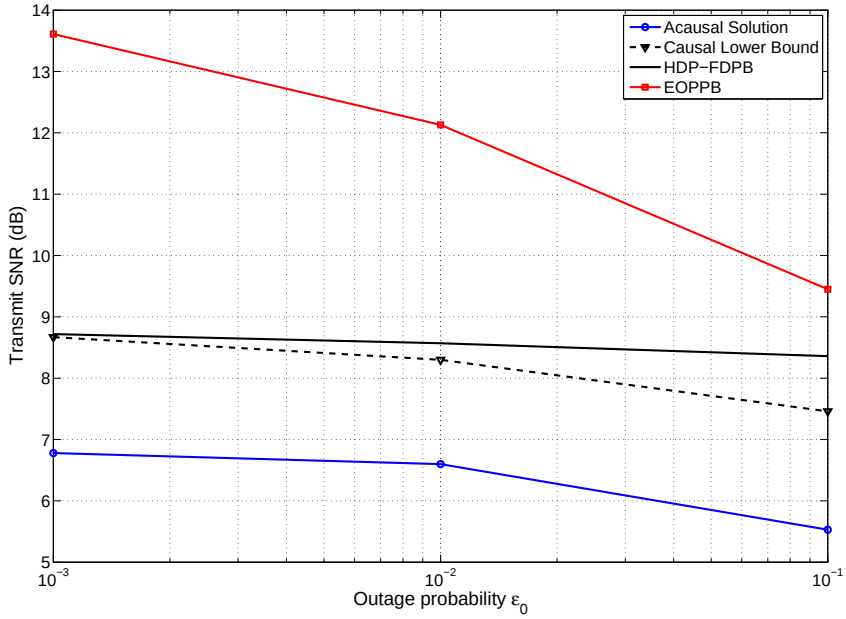

Figure 5.5: SNR results versus the outage probability ε_0 with $R_0 = 2$ (bps/Hz) and $K = 4$.

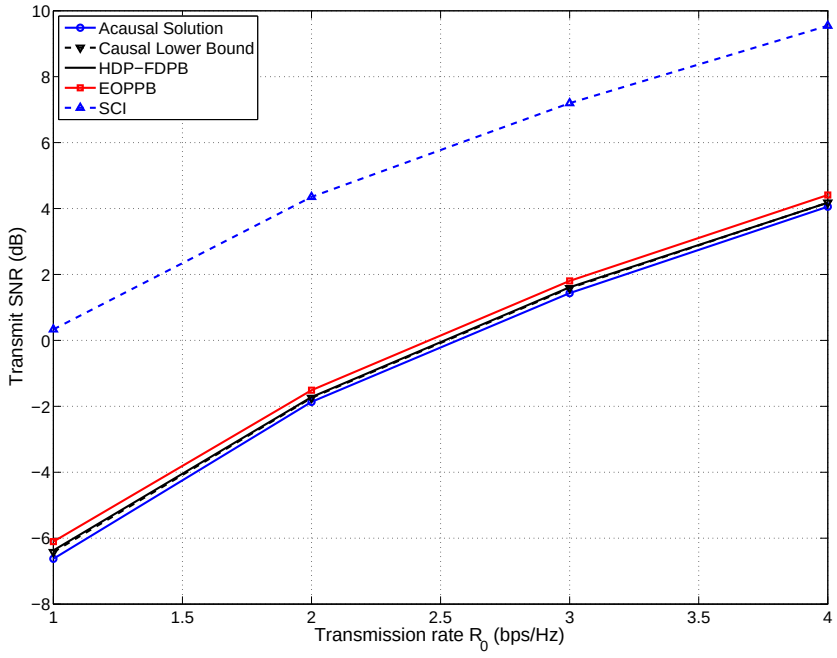

Figure 5.6: SNR results versus the code rate R_0 for a $(3, 2)$ system with $K = 2$ and $\varepsilon_0 = 0.01$.

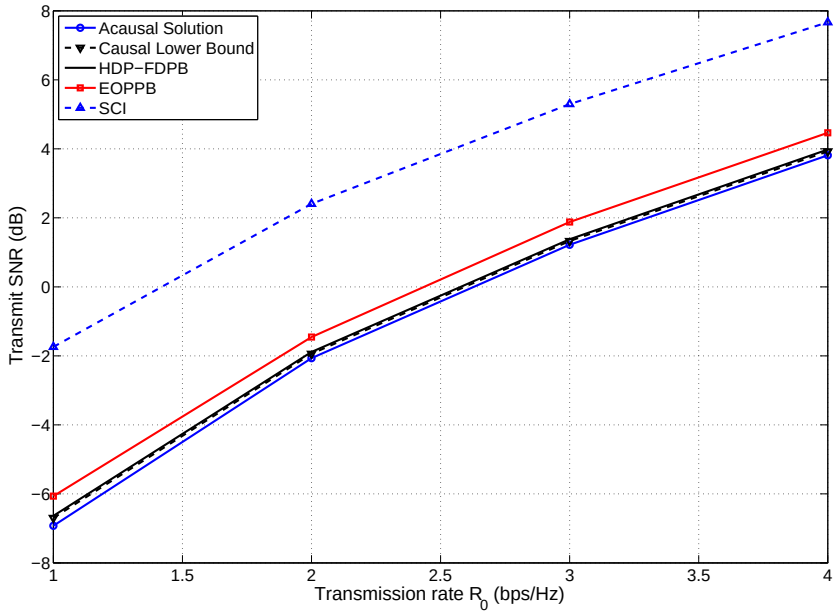

Figure 5.7: SNR results versus the code rate R_0 for a $(3, 2)$ system with $K = 8$ and $\varepsilon_0 = 0.01$.

Chapter 6

Wireless Resource Allocation in Multi-user Systems with Transmitter SCI

In Chapter 6 of the thesis, we aim to minimize the overall transmit power for a time-division multi-user MIMO system with transmitter SCI in BF channels where users are given individual information outage probability constraint. On the basis of the wireless resource allocation problem in the single user system from Chapter 4, this chapter further extends to a multi-user scenario and joint power and time slot allocation are performed. The mathematical tool used in this chapter is convex optimization.

The chapter is organized as follows. Section 6.1 introduces the motivation of the research work in this chapter. In Section 6.2, we present the BF channel model for a time-division multi-user MIMO system and formulate the joint multi-user time-sharing and power allocation problem. Section 6.3 briefly reviews the MPE for searching the minimum power allocation for a single user MIMO BF channel derived from Chapter 4. Section 6.4 proposes a convex problem to obtain

a suboptimal multi-user time-sharing solution. In Section 6.5, an algorithm which finds a joint time-sharing and power allocation solution is presented. Numerical results will be provided in Section 6.6. Finally, we have some concluding remarks in Section 6.7.

6.1 Introduction

In a multi-user system, allocating wireless resource of power, spectrum and/or time slots becomes more complicated as increasing design factors are needed to be taken into consideration. There are limited existing research focusing on energy efficient design to satisfy deterministic QoS in multi-user systems [54, 56]. In [54], the authors solve the problem of minimizing the energy used to transmit multiple packets over wireless additive white Gaussian noise (AWGN) channel within deterministic delay constraint, in which a certain decoding delay bound is needed to be satisfied for each transmission. Energy efficient scheduling for wireless sensor networks is considered in [56], where energy is minimized by varying the transmission time length assigned to different sensor nodes. However, only deterministic decoding delay constraint is considered and the channel gain within transmission time is assumed to be constant which is not applicable for most of the channel conditions.

In contrast to the traditional schemes that only consider deterministic delay constraint, this chapter deals with time-division multi-user MIMO systems where each user is given an individual information outage probability constraint. Our goal is to optimize the power allocation among the users and to schedule the users smartly so that the overall transmit power is minimized while the outage probability constraints of the users are satisfied. Assuming that all users are subjected to a delay tolerance of K blocks, the exact order on how the users are scheduled within the blocks is irrelevant. As a consequence, our aim boils down to finding

the optimal power allocation and the optimal time-sharing (i.e., the number of blocks/time slots assigned) among the users. The problem under investigation is especially crucial if the target rates of the users are predetermined and the cost of transmission is to be minimized with only statistical channel feedback. Note that the research work in this chapter can be thought of as an extension of [100] to MIMO channels.

6.2 System Model and Problem Formulation

We consider a MIMO BF channel with channel matrix represented as $\mathbf{H}_k = \left[h_{i,j}^{(k)}\right] \in \mathbb{C}^{n_r \times n_t}$, where n_t and n_r antennas are, respectively, located at the transmitter and the receiver. The amplitude square of each element, $\left|h_{i,j}^{(k)}\right|^2$, has the p.d.f. of (6.1) as that of c_k,

$$f(c_k) = \begin{cases} e^{-c_k} & c_k \geq 0, \\ 0 & c_k < 0. \end{cases} \tag{6.1}$$

and the elements of $\mathbf{H}_k$ are i.i.d. for different k and antenna pairs. Assuming the received noise power for each user $N_0 = 1$ for normalization, the rate achieved for block k can be written as

$$r_k = \log_2 \det \left| \mathbf{I} + \left(\frac{C_0 P}{n_t K} \right) \mathbf{H}_k \mathbf{H}_k^{\dagger} \right| \quad \text{in bps/Hz} \tag{6.2}$$

where C_0 is the mean power gain, P is the total transmit power over K block channel. In (6.2), we have used the fact that the transmit covariance matrix at time k is $\frac{P\mathbf{I}}{n_t K}$ because the transmitter does not have the instantaneous CSI and thus transmits the same power across the antennas. By transmitting power of $\frac{P}{n_t K}$ at each antenna, the transmit power at each block is kept as $\frac{P}{K}$. For conciseness, in the sequel, we shall assume that $n_t \geq n_r$ and the matrix $\mathbf{H}_k$ is always of full rank. The case of $n_t < n_r$ can be treated in a similar way and thus omitted.

As only SCI is assumed at the transmitter, the exact order of how the users are scheduled for transmission within the K blocks is unimportant and the only thing matters is the amount of channel resources (such as the number of time slots) allocated to the users.

As a result, for a U-user system where w_u time slots are allocated to user u (note that $\sum_{u=1}^{U} w_u \leq K$), we can now assume that user u accesses the channels in time slots (or blocks) k such that

$$k \in \mathcal{D}_u \equiv \left\{ \forall k \in \mathbb{Z} : \sum_{j=1}^{u-1} w_j + 1 \leq k \leq \sum_{j=1}^{u} w_j \right\}. \tag{6.3}$$

Following the model described previously, the IMI attained for user u is given by

$$\frac{1}{w_u} \sum_{k \in \mathcal{D}_u} r_k = \frac{1}{w_u} \sum_{k \in \mathcal{D}_u} \log_2 \det \left| \mathbf{I} + \left(\frac{C_0^{(u)} P_u}{n_t w_u} \right) \mathbf{H}_k^{(u)} \mathbf{H}_k^{(u)\dagger} \right| \quad \text{in bps/Hz} \tag{6.4}$$

where P_u denotes the total transmit power for user u, $\mathbf{H}_k^{(u)}$ is the MIMO channel matrix from the transmitter to user u at slot k, and $C_0^{(u)}$ refers to the mean channel power gain between the transmitter and user u.

Given a target rate $\frac{R_u}{w_u}$ for user u, an outage will occur if $\frac{1}{w_u} \sum_{k \in \mathcal{D}_u} r_k < \frac{R_u}{w_u}$, and the outage tolerance of a user can be characterized by information outage probability constraint

$$\mathsf{Pr} \left\{ \frac{1}{w_u} \sum_{k \in \mathcal{D}_u} r_k < \frac{1}{w_u} R_u \right\} \leq \varepsilon_u \tag{6.5}$$

which can be rewritten as

$$\mathsf{Pr} \left\{ \sum_{k \in \mathcal{D}_u} r_k < R_u \right\} \leq \varepsilon_u, \tag{6.6}$$

where ε_u denotes the maximum allowable outage probability for user u. Note that (6.6) can be viewed as a probabilistic delay constraint which enables us to consider

requirements such as target sum rate R_u, outage tolerance ε_u and time delay in the number of time slots K.

Different from single user system, the problem of interest for multi-user system is to minimize the overall transmit power (i.e., $\sum_{u=1}^{U} P_u$) while ensuring the users' individual outage probability constraints by jointly optimizing the time-sharing (i.e. the number of allocated time slots $\{w_u\}$) and the power allocation (i.e., $\{P_u\}$) for the users. Mathematically, this is written as

$$\mathbb{M} \mapsto \begin{cases} \min\limits_{\{P_u\},\{w_u\}} \sum\limits_{u=1}^{U} P_u \\ \quad \text{s.t. } \mathsf{Pr}\left\{ \sum\limits_{k\in\mathcal{D}_u} r_k < R_u \right\} \le \varepsilon_u \ \forall u, \\ \quad P_u \ge 0 \ \forall u, \\ \quad \sum\limits_{u=1}^{U} w_u \le K, \\ \quad w_u \in \{1, 2, \ldots, K-U+1\} \ \forall u, \end{cases} \tag{6.7}$$

where

P_u the total power allocated to user u; (6.8)

w_u the number of blocks (or the amount of time) allocated to user u; (6.9)

$\mathcal{D}_u$ the set storing the indices of the channels assigned to user u; (6.10)

U the total number of users; (6.11)

K the number of blocks; (6.12)

R_u the target rate for user u; (6.13)

ε_u the outage probability requirement for user u. (6.14)

The challenge of $\mathbb{M}$ is that it is a mixed integer problem which has no known method of achieving the global optimum [84]. The rest of the chapter will be

devoted to solving (6.7).

6.3 MPE

In Chapter 4, we have derived MPE to determine the minimum power required for attaining a given information outage probability constraint if the number of blocks is fixed. In time-division systems, as each block is occupied by one user only, if $\{w_u\}$ are fixed, then the optimization for the users is completely uncoupled and will be equivalent to multiple individual user's power minimizations. Therefore, it suffices to focus on a single user system for a given number of blocks, n, where the outage probability can be written as follows:

$$\mathcal{P}_{\mathsf{out}} \triangleq \mathsf{Pr}\left\{\sum_{k=1}^{n} \log_2 \det\left|\mathbf{I} + \left(\frac{C_0 d^{-\gamma} P}{n_t n}\right)\mathbf{H}_k \mathbf{H}_k^{\dagger}\right| < R\right\}. \tag{6.15}$$

Similar with (4.22) in Chapter 4, we have the MPE based on the above outage constraint as

$$g(P) \triangleq n\rho(P) - \left[\sqrt{2n}\mathsf{erf}^{-1}\left(1-2\varepsilon\right)\right]\sigma(P) - R = 0, \tag{6.16}$$

which can be solved using similar method DIRECT algorithm in 4.5. By denoting $\Gamma_0 \triangleq \frac{C_0 d^{-\gamma} P}{n_t n}$, we have $\rho(P)$ and $\sigma^2(P)$ derived in Section 4.4 as

$$\begin{aligned}
\rho(\Gamma_0) &= \frac{1}{\ln 2} \sum_{\ell=0}^{n_r-1} \sum_{m=0}^{\ell} \frac{\ell!}{(\ell+\delta)!} \left[\frac{1}{m!} \binom{\ell+\delta}{m+\delta} \right]^2 \int_0^\infty \ln(1+\Gamma_0\lambda) \lambda^{\delta+2m} e^{-\lambda} d\lambda \\
&\quad + \frac{2}{\ln 2} \sum_{\ell=1}^{n_r-1} \sum_{i=0}^{\ell-1} \sum_{j=i+1}^{\ell} \frac{\ell!}{(\ell+\delta)!} \frac{(-1)^{i+j}}{i!j!} \binom{\ell+\delta}{i+\delta} \binom{\ell+\delta}{j+\delta} \int_0^\infty \ln(1+\Gamma_0\lambda) \lambda^{\delta+i+j} e^{-\lambda} d\lambda, \\
\sigma^2(\Gamma_0) &= \frac{1}{\ln^2 2} \sum_{\ell=0}^{n_r-1} \sum_{m=0}^{\ell} \frac{\ell!}{(\ell+\delta)!} \left[\frac{1}{m!} \binom{\ell+\delta}{m+\delta} \right]^2 \int_0^\infty \ln^2(1+\Gamma_0\lambda) \lambda^{\delta+2m} e^{-\lambda} d\lambda \\
&\quad + \frac{2}{\ln^2 2} \sum_{\ell=1}^{n_r-1} \sum_{i=0}^{\ell-1} \sum_{j=i+1}^{\ell} \frac{\ell!}{(\ell+\delta)!} \frac{(-1)^{i+j}}{i!j!} \binom{\ell+\delta}{i+\delta} \binom{\ell+\delta}{j+\delta} \int_0^\infty \ln^2(1+\Gamma_0\lambda) \lambda^{\delta+i+j} e^{-\lambda} d\lambda \\
&\quad - \frac{1}{\ln^2 2} \sum_{i=0}^{n_r-1} \sum_{j=0}^{n_r-1} \frac{i!j!}{(i+\delta)!(j+\delta)!} \\
&\quad \times \left[\sum_{m=0}^{i} \sum_{\ell=0}^{j} \frac{(-1)^{m+\ell}}{m!\ell!} \binom{i+\delta}{m+\delta} \binom{j+\delta}{\ell+\delta} \int_0^\infty \lambda^{\delta+m+\ell} e^{-\lambda} \ln(1+\Gamma_0\lambda) d\lambda \right]^2,
\end{aligned} \tag{6.17}$$

where $\delta \triangleq n_t - n_r$. Further, derived from Chapter 4, the integrals of the forms, $\int_0^\infty \lambda^j e^{-\lambda} \ln(1+\Gamma_0\lambda) d\lambda$ and $\int_0^\infty \lambda^j e^{-\lambda} \ln^2(1+\Gamma_0\lambda) d\lambda$, are, respectively, given by (4.16) and (4.19).

To summarize, we now have the MPE to determine the minimum required transmit power for achieving a given information outage probability in an n-block MIMO fading channel. Presumably, if the time-sharing parameters (i.e., $\{w_u\}$) of a time-division multi-user system are known, then the corresponding optimal power allocation for the users can be found from the MPEs. And, the optimal solution of (6.7) could be found using the MPE by an exhaustive search over the space of $\{w_u\}$ (see Section 6.6.1 for details). However, this searching approach will be too complex to do even if the number of users or blocks is moderate. To address this, in the next section, we shall focus on how a sensible solution of $\{w_u\}$ can be found suboptimally.

6.4 Multi-user Time-Sharing from Convex Optimization

In this section, our aim is to optimize the time-sharing parameters $\{w_u\}$ by *joint* consideration of the power consumption and the probability constraints of the users. Ideally, it requires to solve $\mathbb{M}$, i.e., (6.7), which is unfortunately not known. Here, we propose to mimic $\mathbb{M}$ by considering a simpler problem with the probability constraints replaced by some upper bounds, i.e.,

$$\begin{aligned}\mathsf{Pr}\left\{\sum_{k\in\mathcal{D}_u} r_k < R_u\right\} &\le \mathsf{Pr}\left\{\sum_{k\in\mathcal{D}_u} \log_2\left(1+\frac{C_0^{(u)}P_u}{n_t w_u}\cdot\lambda_{\mathsf{max}}^{(u,k)}\right) < R_u\right\} \\ &< \mathsf{Pr}\left\{\chi \triangleq \prod_{k\in\mathcal{D}_u}\lambda_{\mathsf{max}}^{(u,k)} < \left(\frac{w_u n_t}{C_0^{(u)}P_u}\right)^{w_u}\times 2^{R_u}\right\} \triangleq \mathcal{P}_{\mathsf{UB}}^{(u)},\end{aligned} \tag{6.18}$$

where $\lambda_{\mathsf{max}}^{(u,k)}$ denotes the maximum eigenvalue of the channel for user u at time slot k. The first inequality in (6.18) comes from ignoring the rates contributed by the smaller spatial sub-channels while the second inequality removes the unity inside the log expression [which may be regarded as a high SNR approximation]. The p.d.f. of $\lambda_{\mathsf{max}}^{(u,k)}$ is given by [98, 101]

$$\mathcal{F}(\lambda) = \sum_{i=1}^{n_r}\sum_{j=\delta}^{(n_t+n_r)i-2i^2}\left(d_{i,j}\cdot\frac{i^{j+1}}{j!}\right)\lambda^j e^{-i\lambda}, \quad \lambda>0, \tag{6.19}$$

where the coefficients $\{d_{i,j}\}$ are independent of λ. In [98], the values of $d_{i,j}$ for a large number of MIMO settings have been enumerated.

The original outage probability constraint in (6.7) can therefore be ascertained by constraining the upper bound of the outage probability $\left\{\mathcal{P}_{\mathsf{UB}}^{(u)} \le \varepsilon_u\right\}$. The advantage by doing so is substantial. First of all, the optimizing variable P_u can be separated from the random variable, and secondly, the distribution of $\ln\chi$ can

be approximated as Gaussian which permits to evaluate $\mathcal{P}_{\mathsf{UB}}^{(u)}$ as

$$\mathcal{P}_{\mathsf{UB}}^{(u)} = \frac{1}{2} + \frac{1}{2}\mathsf{erf}\left(\frac{\ln\left[\left(\frac{w_u n_t}{C_0^{(u)} P_u}\right)^{w_u} 2^{R_u}\right] - w_u\tilde{\rho}}{\sqrt{2w_u}\tilde{\sigma}}\right) \tag{6.20}$$

where $\tilde{\rho}$ and $\tilde{\sigma}$ are derived in Appendix 10.2 as

$$\begin{aligned}
\tilde{\rho} &= \mathtt{E}[\ln\lambda] = \sum_{i=1}^{n_r} \sum_{j=\delta}^{(n_t+n_r)i-2i^2} d_{i,j}\left(H_j - \gamma_{\mathsf{EM}} - \ln i\right), \\
\tilde{\sigma}^2 &= \mathtt{VAR}[\ln\lambda] \\
&= \sum_{i=1}^{n_r} \sum_{j=\delta}^{(n_t+n_r)i-2i^2} d_{i,j}\left[\gamma_{\mathsf{EM}}^2 + 2(\ln i - H_j)\gamma_{\mathsf{EM}} + \frac{\pi^2}{6} - 2H_j \ln i + (\ln i)^2 + 2\sum_{t=1}^{j-1}\frac{H_t}{t+1}\right] - \tilde{\rho}^2,
\end{aligned} \tag{6.21}$$

where H_ℓ is the Harmonic number defined as $\sum_{m=1}^{\ell}\frac{1}{m}$.

The constraint $\{\mathcal{P}_{\mathsf{UB}}^{(u)} \le \varepsilon_u\}$ can therefore be simplified to

$$P_u \ge \frac{n_t}{C_0^{(u)} e^{\tilde{\rho}}} \cdot \frac{w_u \left(2^{R_u}\right)^{\frac{1}{w_u}}}{\left[e^{-\sqrt{2}\tilde{\sigma}\mathsf{erf}^{-1}(1-2\varepsilon_u)}\right]^{\frac{1}{\sqrt{w_u}}}}. \tag{6.22}$$

Using the upper bound constraints in the multi-user problem (6.7), we then have

$$\tilde{\mathbb{M}} \mapsto \begin{cases} \min\limits_{\{P_u\},\{w_u\}} \sum\limits_{u=1}^{U} P_u \\ \quad \text{s.t. } P_u \ge \dfrac{n_t}{C_0^{(u)} e^{\tilde{\rho}}} \cdot \dfrac{w_u \left(2^{R_u}\right)^{\frac{1}{w_u}}}{\left[e^{-\sqrt{2}\tilde{\sigma}\mathsf{erf}^{-1}(1-2\varepsilon_u)}\right]^{\frac{1}{\sqrt{w_u}}}} \quad \forall u, \\ \quad \sum\limits_{u=1}^{U} w_u \le K, \\ \quad w_u \in \{1, 2, \ldots, K-U+1\} \quad \forall u, \end{cases} \tag{6.23}$$

where the constraints are now written in closed-form.

It is anticipated that the power allocation from the modified problem (6.23) may be quite conservative, i.e., $P_{\mathsf{opt}}|_{\tilde{\mathbb{M}}} \gg P_{\mathsf{opt}}|_{\mathbb{M}}$ because the upper bound may be

loose. However, our conjecture is that the problem structure of $\mathbb{M}$ on $\{w_u\}$ would be accurately imitated by $\tilde{\mathbb{M}}$. Accordingly, we may be able to obtain near-optimal solution for $\{w_u\}$ by solving $\tilde{\mathbb{M}}$ though accurate power consumption cannot be estimated from $\tilde{\mathbb{M}}$. Following the same argument, the exact tightness of the upper bound and how accurate the Gaussian approximation is in evaluating the upper bound probability are not important, as long as $\tilde{\mathbb{M}}$ preserves the structure to balance the users' channel occupancy and power consumption to meet the outage probability requirements.

A remaining difficulty of solving $\tilde{\mathbb{M}}$ is that the optimization is mixed with combinatorial search over the space of $\{w_u\}$ because they are integer-value [84]. To tackle this, we relax $\{w_u\}$ to positive real numbers $\{x_u^2\}$ so that $\tilde{\mathbb{M}}$ can be rewritten as

$$\tilde{\mathbb{M}}_r \mapsto \begin{cases} \min\limits_{\{x_u\}} \dfrac{n_t}{e^{\tilde{\rho}}} \displaystyle\sum_{u=1}^{U} \frac{1}{C_0^{(u)}} \cdot \frac{x_u^2 \, (a_u)^{\frac{1}{x_u^2}}}{(b_u)^{\frac{1}{x_u}}} \\ \text{s.t.} \displaystyle\sum_{u=1}^{U} x_u^2 \le K, \\ \qquad 1 \le x_u \le \sqrt{K-U+1}, \end{cases} \tag{6.24}$$

where $a_u \triangleq 2^{R_u}$ and $b_u \triangleq e^{-\sqrt{2}\tilde{\sigma}\mathsf{erf}^{-1}(1-2\varepsilon_u)}$. Apparently, both constraints in (6.24) are convex, and if the cost is also convex, the problem can be solved using known convex programming routines [84].

Now, let us turn our attention on a function of the form

$$f(x) = x^2 \cdot \frac{a^{\frac{1}{x^2}}}{b^{\frac{1}{x}}} \equiv x^2 h(x), \text{ for } a, b, x > 0, \tag{6.25}$$

where $h(x) \triangleq \frac{a^{\frac{1}{x^2}}}{b^{\frac{1}{x}}}$. Our interest is to examine if $f(x)$ is convex, or equivalently whether $f''(x) > 0$ which is the second order condition of convex function from

Section 3.1.2. To show this, we first obtain

$$\begin{cases} h'(x) = h(x)\left(\dfrac{-2\ln a}{x^3} + \dfrac{\ln b}{x^2}\right), \\ h''(x) = h(x)\left(\dfrac{-2\ln a}{x^3} + \dfrac{\ln b}{x^2}\right)^2 + h(x)\left(\dfrac{6\ln a}{x^4} - \dfrac{2\ln b}{x^3}\right). \end{cases} \tag{6.26}$$

Applying these results, $f''(x)$ can be found as

$$\frac{f''(x)}{h(x)} = \left(\frac{-2\ln a}{x^2} + \frac{\ln b}{x}\right)^2 + \left(\frac{-2\ln a}{x^2} + \frac{2\ln b}{x}\right) + 2. \tag{6.27}$$

Letting $\alpha = \frac{2\ln a}{x^2}$ and $\beta = \frac{-\ln b}{x}$, we have

$$\frac{f''(x)}{h(x)} = (\alpha+\beta)^2 - \alpha - 2\beta + 2 = \left(\alpha - \frac{1}{2}\right)^2 + (\beta-1)^2 + \frac{3}{4} + 2\alpha\beta > 0, \tag{6.28}$$

since $\alpha, \beta > 0$, which can be seen from the definition of (a, b) that $\alpha > 0$ and $\beta > 0$ for $\varepsilon_u < 0.5$.[1] Together with the fact that $h(x) > 0$ for all $x > 0$, we can conclude that $f''(x) > 0$ and therefore, $f(x)$ is convex. As the cost function in (6.24) is a summation of the functions of the form $f(x)$, it is convex and hence the problem (6.24), or $\tilde{\mathbb{M}}_r$.

With $\tilde{\mathbb{M}}_r$ being convex, we can find the global optimal $\{x_u\}_{\mathsf{opt}}$ at polynomial-time complexity. In particular, the complexity grows at most as $\mathcal{O}(U^3)$, which is scalable with the number of users [84]. The remaining task, however, is to derive the integer-valued $\{w_u\}$ from $\{x_u\}$. Simply setting $w_u = x_u^2$ would result in non-integer solutions while rounding them off could lead to violation of the outage rate probability constraints. In the following section, a greedy-approach will be presented to obtain a feasible solution of $\{w_u\}$ from $\{x_u\}$.

[1] It should be emphasized that the convexity of f is subjected to the condition that $\varepsilon_u < 0.5$. However, in practice, it would not make sense to have a system operating with outage probability greater than 50%.

6.5 The Proposed Algorithm

Thus far, we have presented two main approaches: one that determines the optimal transmit power $\{P_u\}$ based on MPE (see Section 6.3); another one that finds the suboptimal (relaxed) time-sharing parameters $\{w_u\}$ by constraining the upper bound probability (see Section 6.4). In this section, we shall devise an algorithm that combines the two approaches to jointly optimize the power allocation and time-sharing of the users. Our idea is to first map the optimal solution $\{x_u\}_{\mathsf{opt}}$ from $\tilde{\mathbb{M}}_r$ to a proper $\{w_u\}$ in $\mathbb{M}$ by rounding the results to the nearest positive integers, and then to step-by-step allocate one more block to the user who can minimize the overall required power using MPE. The proposed algorithm is described as follows:

1. Solve $\{x_u\}$ in $\tilde{\mathbb{M}}_r$ (6.24) using convex optimization routines such as interior-point method [84] and optimization toolbox provided in Matlab.

2. Initialize $w_u = \lfloor x_u^2 \rfloor \ \forall u$ where $\lfloor y \rfloor$ returns the greatest integer that is smaller than y. Notice that at this point, $\{w_u\}$ and $\{P_u\}$ from $\tilde{\mathbb{M}}_r$ may not give a feasible solution to $\mathbb{M}$, and some outage probability constraints may not be satisfied.

3. For each user u, compute the minimal required power to ensure the outage probability constraint by solving MPE

$$\begin{aligned} P_u &= \arg\left\{g_u(P|w_u) = 0\right\} \\ &= \arg\min_{P\geq 0} \left|g_u(P|w_u)\right|, \end{aligned} \tag{6.29}$$

   where the function $g_u(P|w_u)$ is defined similarly as in (6.16). The notation $(P|w_u)$ is used to emphasize the fact that w_u is given as a fixed constant.

4. Then, initialize $m = K - \sum_{u=1}^{U} w_u$.

5. Compute the power reduction metrics

$$\triangle P_u = P_u - \arg\min_{P\geq 0} |g_u(P|w_u + 1)| \quad \forall u. \tag{6.30}$$

6. Find $u^* = \arg\max_u \triangle P_u$ and update

$$\begin{cases} w_{u^*} := w_{u^*} + 1, \\ P_{u^*} := P_{u^*} - \triangle P_{u^*}, \\ m := m - 1. \end{cases} \tag{6.31}$$

   If $m \geq 1$, go back to Step 5. Otherwise, go to Step 7.

7. Optimization is completed and the solutions for both $\{w_u\}$ and $\{P_u\}$ have been found.

A first look at the algorithm reveals that the required complexity of the proposed algorithm is

$$\begin{aligned} \mathcal{C}_{\mathsf{proposed}} &= \mathcal{O}(U^3) + mU\mathcal{C}_{\mathsf{DIRECT}} \\ &\lesssim \mathcal{O}(U^3) + U^2\mathcal{C}_{\mathsf{DIRECT}} \end{aligned} \tag{6.32}$$

where $\mathcal{C}_{\mathsf{DIRECT}}$ denotes the required complexity for solving MPE by using DIRECT algorithm.

6.6 Simulation Results

6.6.1 Simulation Setup and Benchmarks

Computer simulations are conducted to evaluate the performance of the proposed algorithm for the power-minimization problem with outage probability constraints. Only SCI has been assumed and the total transmit SNR, defined as $\left(\sum_{u=1}^{U} P_u\right)$, is considered as the performance metric. To compute the required SNR for a given

set of simulation parameters such as the number of users and blocks, the users' target rates and outage probabilities, the algorithm presented in Section 6.3, which iteratively solves the MPE, is used. Note that the MPE itself has already taken into account the randomness of the channel for outage evaluation.

Results for the proposed algorithm will be compared with the following benchmarks:

- *Global Optimum*—With MPE presented in Section 6.3, it is possible to find the global optimal solution of time and power allocation for the users by solving $\mathbb{M}$ over the space of $\{w_u\}$ at the expense of much greater complexity, i.e.,

$$\mathbb{M} \mapsto \begin{cases} \min\limits_{\{w_u\}} & \sum\limits_{u=1}^{U} \arg\left\{g_u(P_u|w_u) = 0\right\} \\ \text{s.t.} & \sum\limits_{u=1}^{U} w_u \leq K, \\ & w_u \in \{1, 2, \ldots, K-U+1\}\ \forall u. \end{cases} \tag{6.33}$$

  The required complexity is given by

$$\begin{aligned} \mathcal{C}_{\mathsf{optimum}} &= \binom{K}{U} U \mathcal{C}_{\mathsf{DIRECT}} \\ &\approx \binom{K}{U} \left[\frac{\mathcal{C}_{\mathsf{proposed}} - \mathcal{O}(U^3)}{U}\right]. \end{aligned} \tag{6.34}$$

  For large K, we have

$$\frac{\mathcal{C}_{\mathsf{optimum}}}{\mathcal{C}_{\mathsf{proposed}}} \approx \frac{1}{U}\binom{K}{U} \tag{6.35}$$

  and the complexity saving by the proposed scheme is enormous. For instance, if $U = 4$ and $K = 30$, the ratio is approximately 6851.

- *Equal-Time with Optimized Power*—An interesting benchmark is the system where each user is allocated more or less an equal number of blocks (i.e., $w_u \approx \lfloor \frac{K}{U} \rfloor\ \forall u$ and with $\sum_{u=1}^{U} w_u = K$) while the power allocation for each user is optimized by solving MPE. Obviously, if the system has homogeneous

users (e.g., users with the same channel statistics and outage requirements), then equal-time allocation should be optimal. This system can show how important time-sharing optimization is if the system has highly heterogeneous users.

- *Equal-Time with Suboptimal Power (6.22)*—A suboptimal power allocation to achieve a given outage probability can be found by (6.22) based on the upper bound probability without relying on the MPE. This system enables us to see how important the MPE is.

6.6.2 Results

Results in Fig. 6.1 are provided for the transmit power against the outage probability requirements for a 3-user system with 20 blocks (i.e., $K = 20$). The users are considered to have target rates $(R_1, R_2, R_3) = (8, 12, 16)$ bps/Hz, channel power gains $(C_0^{(1)}, C_0^{(2)}, C_0^{(3)}) = (0.8, 1, 1.2)$, multiple receive antennas $(n_r^{(1)}, n_r^{(2)}, n_r^{(3)}) = (2, 3, 2)$ and the same outage probability requirements (ε) ranging from 10^{-5} to 10^{-1}. The number of transmit antennas at the base station is set to be $n_t = 4$. Results in this figure show that the total transmit power of the proposed scheme decreases if the required outage probability increases. For example, there is about 2dB power reduction when ε increases from 10^{-5} to 10^{-1}. Results also illustrate that the proposed method performs nearly the same as the global optimum. However, compared with the equal-time method with optimum power solution, there is only about 0.2dB reduction in power by the proposed method. This is because the optimal strategy tends to allocate similar number of blocks to the users, which can be observed from Configuration 1 of Table 6.1. In addition, as can be seen, the transmit power of the equal-time method with suboptimal power is much greater than that with optimum power, which shows that the MPE is very important in optimizing the power allocation. In particular, more than 3dB in power is required when compared with the equal-time method with optimal power solution.

The power results against the target rate for a 3-user system with total number of blocks $K = 15$, $(\varepsilon_1, \varepsilon_2, \varepsilon_3) = (10^{-4}, 10^{-3}, 10^{-2})$, and $(C_0^{(1)}, C_0^{(2)}, C_0^{(3)}) = (0.5, 1, 1.5)$ are plotted in Fig. 6.2. Also, 3 transmit antennas and 2 receive antennas per user are considered, and all the users have the same target rate R. Results indicate that the total transmit power increases dramatically with R (e.g., 10dB increase from 8bps/Hz to 32bps/Hz for the proposed method). As can be observed, the increase in power is almost linear with R. In addition, the proposed method consistently performs nearly as the global optimum although the gap between the proposed method and the equal-time method with optimal power solution is not very obvious.

In Fig. 6.3, we have the results for the transmit power against the total number of blocks K for a 3-user system with $(\varepsilon_1, \varepsilon_2, \varepsilon_3) = (10^{-2}, 10^{-3}, 10^{-4})$, $(R_1, R_2, R_3) = (16, 20, 24)$ bps/Hz, and $(C_0^{(1)}, C_0^{(2)}, C_0^{(3)}) = (1.5, 1, 0.5)$. The number of transmit antennas is 4 while users' number of receive antennas are $(n_r^{(1)}, n_r^{(2)}, n_r^{(3)}) = (4, 3, 2)$. Note that in this case, we have set the conditions for different users, such as users' requirements and channel conditions, to be quite different from each other to see how the proposed scheme performs. As we can see, the total transmit power decreases as K increases. In particular, the power for the proposed method decreases by 8dB when K increases from 6 to 18. Again, results show that the performance of the proposed scheme is nearly optimal while this time, the gap between the proposed method and the equal-time methods becomes more obvious (about 5dB for $K = 6$ and 2dB for $K = 18$). This is because the optimal strategy tends to allocate more blocks for high-demand poor-channel-condition users (the numbers of allocated blocks for the users for different methods with $K = 12$ are shown in Configuration 3 of Table 6.1).

Fig. 6.4 plots the power results against the number of receive antennas for a 3-user system with $K = 20$, $(\varepsilon_1, \varepsilon_2, \varepsilon_3) = (10^{-1}, 10^{-3}, 10^{-4})$, $(R_1, R_2, R_3) = (8, 16, 24)$ bps/Hz, $(C_0^{(1)}, C_0^{(2)}, C_0^{(3)}) = (1.5, 1, 0.5)$ and $n_t = 4$. As expected, the required

transmit power decreases with the number of receive antennas. This can be explained by the fact that the transmission rate mainly depends on the rank of the MIMO system, which is limited by the number of receive antennas (n_r). The actual number of block allocation for various methods is provided in Configuration 4 of Table 6.1.

6.7 Summary

This chapter has addressed the optimization problem of power allocation and scheduling for a time-division multi-user MIMO system in Rayleigh BF channels when the transmitter has only the SCI of the users and the users are given individual information outage probability constraint. We have applied MPE on the basis of single user power allocation in Chapter 4 to determine the minimum power for attaining a given outage rate probability constraint if the number of blocks for a user is fixed. On the other hand, we have proposed a convex programming approach to find the suboptimal number of blocks allocated to the users. The two main techniques have then been combined to obtain a joint solution for both power and time allocation for the users. Results have demonstrated that the proposed method achieves near-optimal performance. The problem when CSI is available at the transmitter is considered in the next chapter.

Table 6.1: Various configurations tested from Figs. 6.1–6.4. The superscript $\star$ highlights the solution that is not the same as the optimum.

Configuration	u	R_u	ε_u	$C_0^{(u)}$	$n_r^{(u)}$	$(w_u)_{\text{opt}}$	$(w_u)_{\text{proposed}}$	$(w_u)_{\text{equal}}$
1 ($n_t = 4$, $K = 20$)	1	8	10^{-3}	0.8	2	6	6	6
	2	12	10^{-3}	1	3	5	$6^\star$	$7^\star$
	3	16	10^{-3}	1.2	2	9	$8^\star$	$7^\star$
2 ($n_t = 3$, $K = 15$)	1	16	10^{-4}	0.5	2	6	6	$5^\star$
	2	16	10^{-3}	1	2	5	5	5
	3	16	10^{-2}	1.5	2	4	4	$5^\star$
3 ($n_t = 4$, $K = 12$)	1	16	10^{-2}	1.5	4	2	2	$4^\star$
	2	20	10^{-3}	1	3	3	3	$4^\star$
	3	24	10^{-4}	0.5	2	7	7	$4^\star$
4 ($n_t = 4$, $K = 20$)	1	8	10^{-1}	1.5	3	2	$3^\star$	$6^\star$
	2	16	10^{-3}	1	3	6	$7^\star$	$7^\star$
	3	24	10^{-4}	0.5	3	12	$10^\star$	$7^\star$

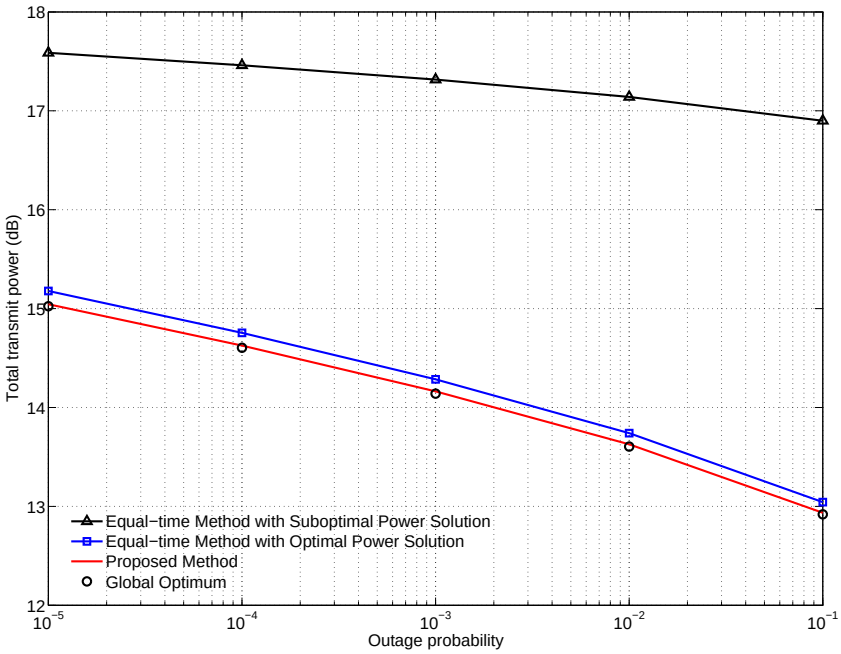

Figure 6.1: Results of the transmit power versus the outage probability when $U = 3$, $K = 20$, $(R_1, R_2, R_3) = (8, 12, 16)$ bps/Hz, $(C_0^{(1)}, C_0^{(2)}, C_0^{(3)}) = (0.8, 1, 1.2)$, $n_t = 4$ and $(n_r^{(1)}, n_r^{(2)}, n_r^{(3)}) = (2, 3, 2)$.

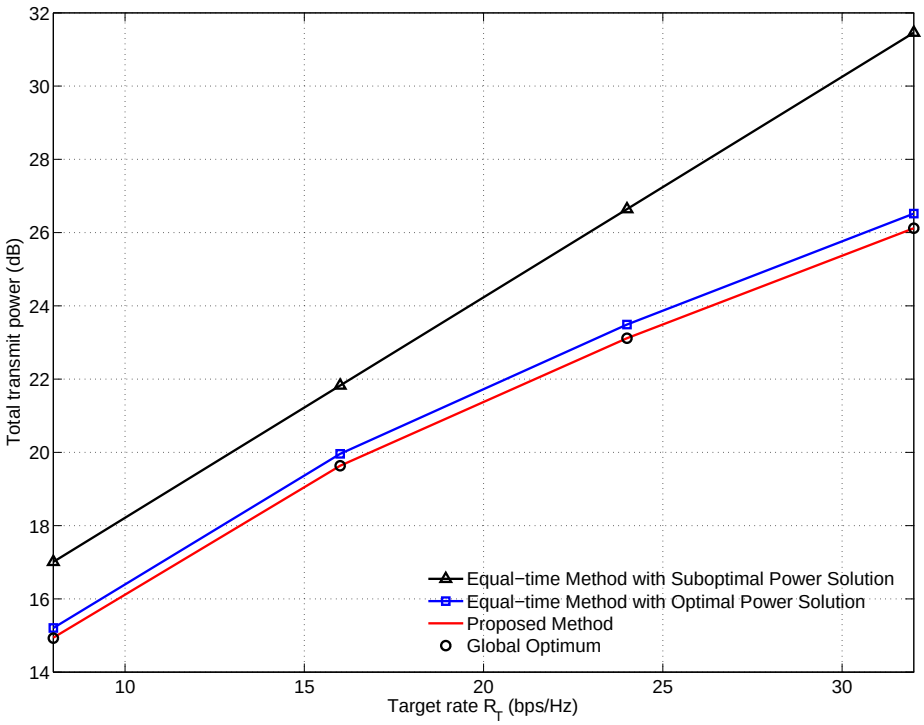

Figure 6.2: Results of the transmit power versus the target rate when $U = 3$, $K = 15$, $(\varepsilon_1, \varepsilon_2, \varepsilon_3) = (10^{-4}, 10^{-3}, 10^{-2})$, $(C_0^{(1)}, C_0^{(2)}, C_0^{(3)}) = (0.5, 1, 1.5)$, $n_t = 3$ and $(n_r^{(1)}, n_r^{(2)}, n_r^{(3)}) = (2, 2, 2)$.

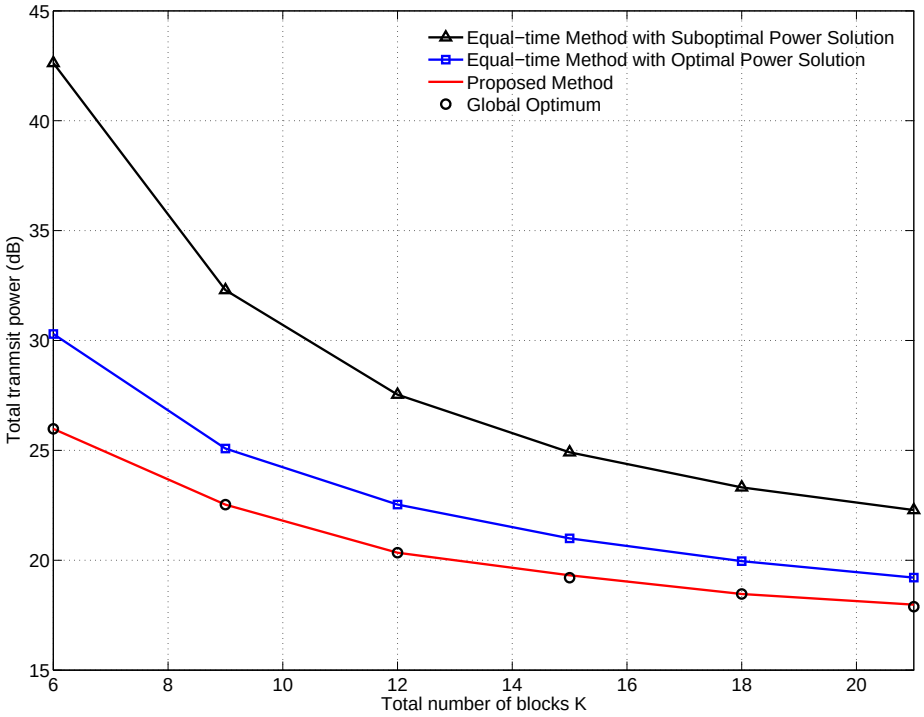

Figure 6.3: Results of the transmit power versus the number of blocks when $U = 3$, $(\varepsilon_1, \varepsilon_2, \varepsilon_3) = (10^{-2}, 10^{-3}, 10^{-4})$, $(R_1, R_2, R_3) = (16, 20, 24)$ bps/Hz, $(C_0^{(1)}, C_0^{(2)}, C_0^{(3)}) = (1.5, 1, 0.5)$, $n_t = 4$ and $(n_r^{(1)}, n_r^{(2)}, n_r^{(3)}) = (4, 3, 2)$.

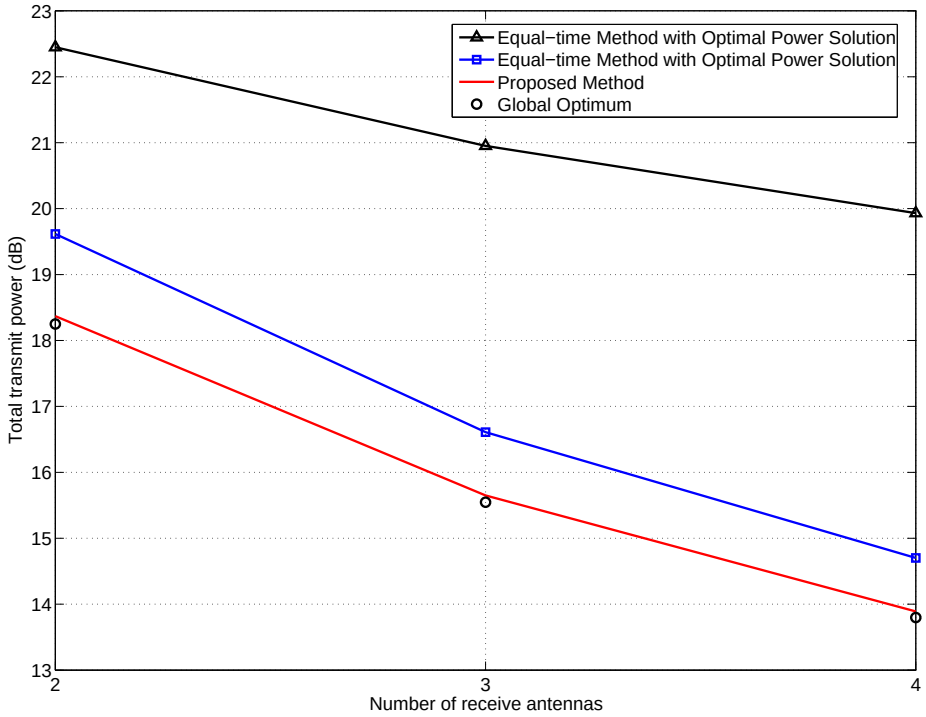

Figure 6.4: Results of the transmit power versus the receive antennas when $U = 3$, $K = 20$, $(\varepsilon_1, \varepsilon_2, \varepsilon_3) = (10^{-1}, 10^{-3}, 10^{-4})$, $(R_1, R_2, R_3) = (8, 16, 24)$ bps/Hz $(C_0^{(1)}, C_0^{(2)}, C_0^{(3)}) = (1.5, 1, 0.5)$ and $n_t = 4$.

Chapter 7

Wireless Resource Allocation in Multi-user Systems with Transmitter CSI

In Chapter 7 of the thesis, we aim to minimize the overall transmit power for a time-division multi-user system with transmitter CSI in BF channels where users are given individual information outage probability constraint. Based on the wireless resource allocation problem in single user system from Chapter 5, this chapter further extends to multi-user scenario and joint power and time slot allocation are performed. Only SISO system is considered in this chapter, as for MIMO system the spatial diversity reduces the need for dynamically allocating power over blocks (time diversity) which is observed from Chapter 5. The mathematical tools used in this chapter are convex optimization and nonlinear optimization.

This chapter is organized as follows. Section 7.1 gives an introduction of the research work in this chapter. In Section 7.2, we present system model and problem formulation. A simplified suboptimal single user power allocation strategy is proposed in Section 7.3, while the multi-user time sharing problem is solved in Section

7.4. The overall strategy is developed in Section 7.5. Simulation results are shown in Section 7.6 and Section 7.7 summaries this chapter.

7.1 Introduction

In this chapter, we aim to develop an energy-efficient scheduling algorithm for multi-user TDMA system in BF channel, subject to individual user's QoS requirements in terms of information outage probability, by jointly adapting time-sharing and power allocation for the users in accordance with the causal CSI at the transmitter. We first derive a simple closed-form power allocation solution for a single user BF channel to attain a given outage probability constraint with causal CSI. Then, we construct an optimization problem to permit a joint consideration of power consumptions and QoS requirements of the users, which we solve to obtain a near-optimal solution for users' time-sharing. The two solutions are combined to provide a suboptimal joint solution for adaptive time-sharing and power allocation of the system. Results reveal that the proposed scheme is near-optimal and a significant energy saving of more than 15 (dB) is possible between systems with and without adaptation on time-sharing and/or power allocation.

7.2 System Model

Let us first assume a single user system where the channel is assumed to be in BF which fades i.i.d. from one block to another. The channel is in Rayleigh fading and the channel power gain g_k for block k has the following p.d.f.

$$f(g_k) = \begin{cases} e^{-g_k} & \text{if } g_k \geq 0, \\ 0 & \text{if } g_k < 0. \end{cases} \tag{7.1}$$

For a given block k, the transmission rate achievable is given by

$$r_k = \log_2(1 + p_k g_k) \quad (\text{bps/Hz}), \tag{7.2}$$

where the noise power is assumed unity and p_k is the transmit power assigned for the block. Outage is said to occur if the transmitted code rate, R_0, is less than the IMI of the channel, i.e.,

$$\mathcal{I}_K \triangleq \frac{1}{K}\sum_{k=1}^{K} r_k < R_0. \tag{7.3}$$

Under this single user system model, our aim is to minimize the average transmit power over the K blocks to attain a given outage probability QoS, ε_0. Mathematically, that is,

$$\min_{\boldsymbol{p}} \mathsf{E}_{\boldsymbol{g}}\left[\frac{1}{K}\sum_{k=1}^{K} p_k\right] \quad \text{s.t.} \quad \mathsf{Pr}\left\{\frac{1}{K}\mathcal{I}_K(\boldsymbol{p},\boldsymbol{g}) < R_0\right\} \le \varepsilon_0, \tag{7.4}$$

where $\boldsymbol{p} \triangleq (p_1, p_2, \cdots p_K)$, and $\boldsymbol{g} \triangleq (g_1, g_2, \cdots g_K)$. Apparently, the minimum power will be achieved when the constraint becomes active, i.e., $\mathsf{Pr}\left\{\frac{1}{K}\mathcal{I}_K(\boldsymbol{p},\boldsymbol{g}) < R_0\right\} = \varepsilon_0$ because the random variable $\mathcal{I}$ is an increasing function of the transmit power.

In a TDMA system, the channel is shared by multiple users in time. In particular, we assume that there are U users, each transmitting at a sum rate R_u over w_u blocks with an outage requirement ε_u. Also, it is assumed that user u accesses the channel in continuous time slots (or blocks), $k \in \mathcal{D}_u$, such that

$$\mathcal{D}_u \equiv \left\{\forall k \in \mathbb{Z} : \sum_{j=1}^{u-1} w_j + 1 \le k \le \sum_{j=1}^{u} w_j\right\}. \tag{7.5}$$

In other words, users occupy and share the wireless channel in a sequential manner, and that user u is given w_u blocks of the channel, which is to be optimized. We further assume that the decoding delay for all the users is K, so that $\sum_{u=1}^{U} w_u \le K$.

Following (7.4), the optimization in a multi-user TDMA system requires to solve

$$\begin{cases} \min\limits_{\{p_k \geq 0\},\{w_u\}} \mathsf{E}\left[\dfrac{1}{K}\displaystyle\sum_{u=1}^{U}\sum_{k\in\mathcal{D}_u} p_k\right] \\ \quad \text{s.t. } \mathsf{Pr}\left\{\displaystyle\sum_{k\in\mathcal{D}_u} r_n < R_u\right\} \leq \varepsilon_u \ \forall u, \\ \quad \displaystyle\sum_{u=1}^{U} w_u \leq K, \\ \quad w_u \in \{1, 2, \ldots, K-U+1\} \ \forall u. \end{cases} \tag{7.6}$$

7.3 Single User Power Allocation

The problem of the single user power allocation (7.4) has been studied in Section 5.3 from Chapter 5 and it was shown that the optimal power allocation can be obtained by a DP approach to fully exploit the causality of the CSI. However, this approach is computationally involved.

To avoid the complexity of DP, in this section, we propose OLLP, which unlike the backward optimization of DP, is a forward suboptimal power allocation algorithm. In OLLP, the mutual information in block k can be obtained by

$$r_k(g_k, r^{(k)}) = \arg \min_{\substack{0\leq r_k \leq r^{(k)} \\ 1\leq k\leq K-1}} \left[p_k(r_k, g_k) + S_{k+1}\big(r^{(k)} - r_k(g_k, r^{(k)})\big)\right], \tag{7.7}$$

where $r^{(k)} \triangleq KR_0 - \sum_{j=1}^{k-1} r_j\left(\boldsymbol{g}^{(j)}\right)$ is the remaining data to be fulfilled from block k to K, with $\boldsymbol{g}^{(j)} \triangleq (g_1, g_2, \ldots, g_j)$ and $r_K = r^{(K)}$. The variable S_{k+1}, which we derive later, can be regarded as the expected power for the future blocks. In other words, OLLP intends to balance the rates for the current block, r_k, and the future blocks, $r^{(k+1)}$, with consideration of the causal CSI (i.e., the previous CSI or fulfilled data so far, the current CSI, and the channel statistics (7.1)). Nonetheless, solving (7.7) requires an analytical expression for S_{k+1}, which will unfortunately involve knowing the exact IMI distribution over the future channel blocks. Note

that the future channel blocks are i.i.d., S_{k+1} can be approximated by considering equal expected power for each future channel block [This approximation is particularly accurate for high-rate and short-time transmission (i.e., large R_0 and small K) because in that case, the transmitter tends to allocate the same amount of information for each block]. As a result, it can be shown that p_k and S_{k+1} have the following relationships

$$\begin{cases} p_k = \left.\dfrac{2^{r_k}-1}{g_k}\right|_{g_k \geq g^B}, \\ S_{k+1} \approx \dfrac{1-(1-\widetilde{\varepsilon}_0)^{K-k}}{\widetilde{\varepsilon}_0} \times \\ \qquad \left(2^{\frac{r^{(k)}-r_k}{K-k}}-1\right) E_1\left(-\ln(1-\widetilde{\varepsilon}_0)\right), \end{cases} \tag{7.8}$$

where $\widetilde{\varepsilon}_0 \triangleq 1-(1-\varepsilon_0)^{\frac{1}{K}}$ can be viewed as the per-block outage probability and $g^B = -\ln(1-\widetilde{\varepsilon}_0)$ is the channel gain threshold to determine whether a given channel is in an outage region. The approximation of S_{k+1} is an alternative form of (5.39).

Using (7.8) into (7.7), we can then work out the IMI needed for any block k if $g_k \geq g^B$ or the channel is not in outage by solving

$$r_k(g_k, r^{(k)}) = \arg \min_{\substack{0 \leq r_k \leq r^{(k)} \\ 1 \leq k \leq K-1}} \left\{ \frac{2^{r_k}-1}{g_k} + \frac{1-(1-\widetilde{\varepsilon}_0)^{K-k}}{\widetilde{\varepsilon}_0} \left(2^{\frac{r^{(k)}-r_k}{K-k}}-1\right) E_1\left(-\ln(1-\widetilde{\varepsilon}_0)\right) \right\}. \tag{7.9}$$

Note that if $g_k < g^B$ or the channel is in outage, no power or IMI is allocated.

Now, we define

$$\begin{cases} a \triangleq \dfrac{1}{g_k}, \\ b \triangleq \dfrac{1-(1-\widetilde{\varepsilon}_0)^{K-k}}{\widetilde{\varepsilon}_0} E_1\left(-\ln(1-\widetilde{\varepsilon}_0)\right), \\ c \triangleq \dfrac{1}{K-k}, \\ d \triangleq -a-b. \end{cases} \tag{7.10}$$

Then, (7.9) can be rewritten as

$$\min_{0\leq r_k\leq r^{(k)}} a\cdot 2^{r_k}+b\cdot 2^{c[r^{(k)}-r_k]}+d. \tag{7.11}$$

It can be observed that (7.11) is convex, and can permit a closed-form optimal solution, given by

$$\varsigma_k=\frac{1}{c+1}\left(\log_2\frac{bc}{a}+cr^{(k)}\right). \tag{7.12}$$

Considering the constraint region of ς_k, we can further express the IMI allocation for block k for the case $\{g_k\geq g^B\forall k\leq K\}$ as

$$r_k=\begin{cases}0, & \text{if } \varsigma_k<0 \text{ and } k<K,\\ \varsigma_k, & \text{if } 0\leq\varsigma_k\leq r^{(k)} \text{ and } k<K,\\ r^{(k)}, & \text{if } \varsigma_k>r^{(k)} \text{ or } k=K.\end{cases} \tag{7.13}$$

Otherwise, as said before if the channel is in outage, $r_k=0$ and hence, the power allocation $p_k=0$.

Utilizing the IMI allocation for block k in (7.13), the power allocation for the block can be easily found as

$$p_k=\frac{2^{r_k}-1}{g_k}. \tag{7.14}$$

7.4 Multi-user Time-Sharing

Here, we attempt to deal with the multi-user problem (7.6). In particular, the aim is to optimize the time-sharing parameters $\{w_u\}$ by joint consideration of the users' power consumption and QoS constraints in terms of outage probability. Ideally, in order to find the optimal $\{w_u\}$, it needs to obtain the optimal power allocation $\{p_k\}$, which depends upon the instantaneous causal CSI that will require DP. To overcome this difficulty, we use the formula for S_{k+1} in the previous section to

estimate the power consumption for each user. As such, we can approximate the total power for user u, denoted as P_u, by

$$P_u = \frac{\varepsilon_u \left(2^{\frac{R_u}{w_u}} - 1\right)}{1 - (1-\varepsilon_u)^{\frac{1}{w_u}}} E_1\left(-\frac{1}{w_u}\ln(1-\varepsilon_u)\right). \tag{7.15}$$

As a consequence, (7.6) can be simplified as

$$\begin{cases} \min\limits_{\{x_u\}} \sum\limits_{u=1}^{U} \dfrac{\varepsilon_u \left(2^{\frac{R_u}{x_u}} - 1\right)}{1 - (1-\varepsilon_u)^{\frac{1}{x_u}}} E_1\left(-\dfrac{1}{x_u}\ln(1-\varepsilon_u)\right), \\ \text{s.t.} \sum\limits_{u=1}^{U} x_u \le K, \\ \quad 0 < x_u < K - U + 1 \ \forall u, \end{cases} \tag{7.16}$$

where $\{x_u\}$ are real-valued variables and the relaxed version of $\{w_u\}$. To solve (7.16), we propose to apply nonlinear optimization methods, i.e., SQP introduced in Section 3.2, which can achieve quadratic convergence globally, to search for the optimal $\{x_u\}$. Note that a local minimum is at least achieved by applying this algorithm. After $\{x_u\}$ are found, the following algorithm is used to obtain $\{w_u\}$:

1. Initialize $w_u = \lfloor x_u \rfloor \ \forall u$, where $\lfloor y \rfloor$ returns the greatest integer that is smaller than y.

2. For each user u, compute the minimal required power, $P_u(w_u, R_u, \varepsilon_u)$, to ensure the given information outage probability constraint by (7.15).

3. Then, initialize $m = K - \sum_{u=1}^{U} w_u$.

4. Compute the power reduction metrics

$$\triangle P_u = P_u - \arg\min_{P_u \ge 0} P_u(w_u + 1, R_u, \varepsilon_u) \ \ \forall u. \tag{7.17}$$

5. Find $u^* = \arg\max_u \triangle P_u$ and update

$$\begin{cases} w_{u^*} := w_{u^*} + 1, \\ m := m - 1. \end{cases} \tag{7.18}$$

If $m \geq 1$, go back to Step 4. Otherwise, the optimization is completed and $\{w_u\}$ have been found.

7.5 The Proposed Method

The proposed single user power allocation solution and the multi-user time-sharing scheme can be combined to minimize the transmit power of a multi-user TDMA system with users' individual outage probability constraints. The proposed algorithm is given as follows:

1. Given the users' QoS $\{R_u, \varepsilon_u\}$, find the time-sharing parameters $\{w_u\}$ using the method in Section 7.4.

2. Randomly select a user u which has not been served and use the OLLP power allocation policy in Section 7.3 to assign the power for the blocks based on the causal CSI and update w_u to be the actual blocks used for transmission.

3. Revise the total number of blocks to $K - w_u$ and remove this user from the user list. Go back to Step 1. If the user list is empty, then the optimization has been completed.

7.6 Simulation Results

Computer simulations are conducted to evaluate the performance of the proposed method. The metric, defined as the expected total transmit power, i.e.,

$E[\sum_{u=1}^{U}\sum_{k=1}^{w_u} p_k]$, will be used. As normalized noise is assumed, we can use (dB) as the unit for measurement. Figs. 7.1 and 7.2 give the results for single user systems, while Figs. 7.3 and 7.4 show the results for multi-user scenarios.

7.6.1 Benchmarks

In the following, we present several benchmarks we use to compare the results in multi-user systems.

- *The Global Optimal Solution and a Lower Bound*—The optimal solution for the joint multi-user time-sharing and power allocation with the aid of causal CSI, can be, in theory, obtained by an exhaustive search over the space of any possible $\{w_u\}$ while the optimal power allocation is found by the DP method in [81].

  The required complexity to obtain the optimum is given by $\mathcal{O}(\mathcal{G}^{2K+1}\mathcal{R}\mathcal{P}^2) + \mathcal{O}(\mathcal{M}KU) + \mathcal{O}(K^U)$, where $\mathcal{G}$, $\mathcal{R}$ and $\mathcal{P}$ correspond to, respectively, the sizes of the discretization spaces for g_k, R and P, while $\mathcal{M}$ represents the number of numerical simulations needed to estimate the power. Note that the required very high complexity of the optimal solution makes it infeasible to use, and even for the sake of performance comparison, it is too complex for simulations. For this reason, in this chapter, a lower bound, which is derived from the global optimal solution (see the following for details), is used instead.

  If we assume that for each user the channel is identified to be in outage at the beginning of the first block so that power is never wasted, we can obtain the lower bound of the global optimal solution by an exhaustive search over the space of any possible $\{w_u\}$ while applying the single user power allocation lower bound by DP. This provides an important benchmark for evaluation of the proposed algorithm.

- *Equal-Time Sharing with OLLP Power Allocation*—This benchmark refers to the system where each user is given the same number of blocks (i.e., $w_u \approx \lfloor \frac{K}{U} \rfloor \ \forall u$ and with $\sum_{u=1}^{U} w_u \leq K$) while the power allocation for each user is optimized by the proposed OLLP in Section 7.3. Obviously, if the system has homogeneous users (i.e., users of the same QoS requirements), then an equal-time sharing is optimal. This system is useful to demonstrate the importance of time-sharing optimization if users are highly heterogeneous.

- *Optimized Time-Sharing with Equal IMI Allocation*—To reveal the effect of the proposed OLLP power allocation solution in Section 7.3, a suboptimal power allocation which simply ensures that the same IMI is achieved for any given block, is used in conjunction with the multi-user time-sharing solution in Section 7.4.

In Fig. 7.1, average total transmit power results are provided for single user systems against the number of blocks K with sum rate $R = KR_0 = 8$ (bps/Hz) and $\varepsilon_0 = 0.1$ while similar results are plotted in Fig. 7.2 but with $\varepsilon_0 = 0.01$. As can be seen, the required power decreases with the number of blocks K. This can be explained by the fact that coding over a longer period of time increases time diversity for better fading and noise removal.

In addition, results indicate that the proposed OLLP power allocation solution yields near-optimal power allocation results when K is small or the delay is very stringent. This illustrates that the approximation of equal expected IMI allocation over future blocks used in OLLP is particularly accurate when the required IMI per block is large. Nevertheless, if K is large, then the difference between the proposed OLLP and the lower bound starts to appear. In particular, the difference can be as large as 2 (dB) if $\varepsilon_0 = 0.1$ and 5 (dB) if $\varepsilon_0 = 0.01$.

Results in Figs. 7.3 and 7.4 provide the expected total transmit power results for 3-user TDMA systems under the setting $(R_1, R_2, R_3) = (8, 12, 16)$ (bps/Hz)

with $(\varepsilon_1, \varepsilon_2, \varepsilon_3) = (0.1, 0.01, 0.001)$, and then the setting $(R_1, R_2, R_3) = (8, 8, 8)$ (bps/Hz) with $(\varepsilon_1, \varepsilon_2, \varepsilon_3) = (0.01, 0.01, 0.01)$, respectively. Note that the first setting corresponds to the case where users have distinct QoS constraints while the second setting assumes that users are homogeneous. In both figures, the results for the lower bound are only given up to $K = 8$, due to the unaffordable complexity of DP.

As we can see from the results in these figures, similar with the single user case, the expected total transmit power decreases with K. Also, the proposed multi-user time-sharing with OLLP yields near-optimal performance as it matches well with the lower bound. More remarkably, it is noted that OLLP is particularly important in saving power for multi-user systems since it can easily give 20 (dB) reduction in power as compared to the power allocation method based on equal-IMI allocation, even though multi-user time-sharing is optimized. Multi-user time-sharing is also important if users are heterogeneous. This can be seen from the results in Fig. 7.3 that for small number of blocks K, the difference between with and without time-sharing optimization can be as large as 15 (dB).

7.7 Summary

In this chapter, we have devised an energy efficient joint time-sharing and power allocation algorithm for multi-user TDMA systems in SISO BF channels, with individual user's QoS in terms of outage probability requirement, by exploiting the causal CSI at the transmitter. First, a simple closed-form OLLP power allocation solution has been presented to meet a given outage probability constraint for a single user BF channel. We have also constructed an optimization problem which enables a joint consideration of power consumption and QoS of the users to find a suboptimal multi-user time-sharing solution. This time-sharing solution in conjunction with the OLLP solution forms the joint solution for the multi-user

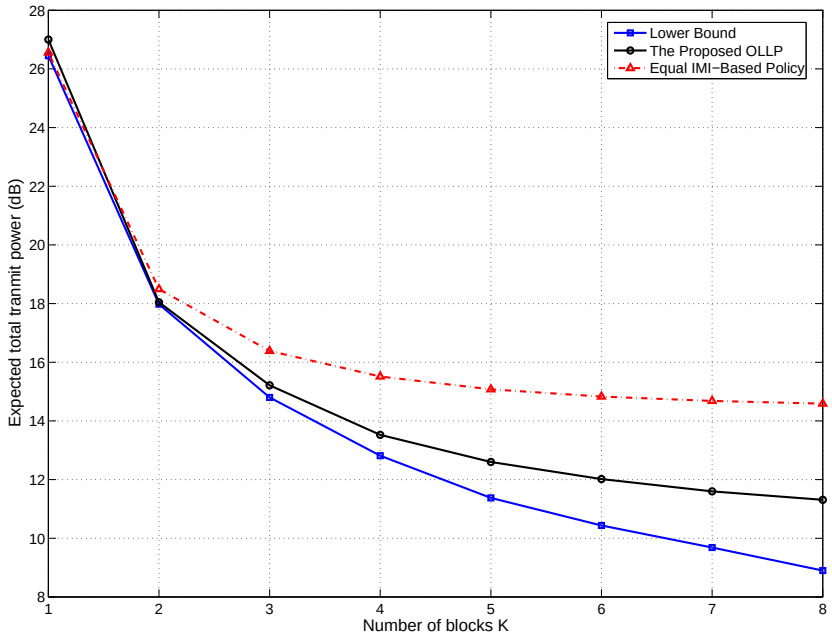

Figure 7.1: The expected total transmit power versus the number of blocks K with $R = 8$ (bps/Hz) and $\varepsilon_0 = 0.1$ for single user systems.

TDMA system, which has been shown by simulation results to be near-optimal and yields a significant energy saving as compared to those without time-sharing optimization or OLLP power allocation.

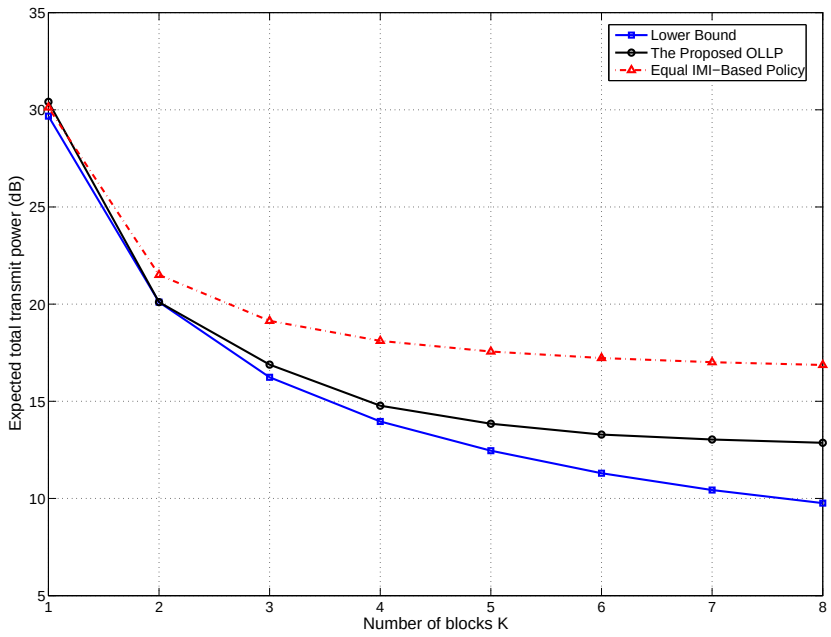

Figure 7.2: The expected total transmit power versus the number of blocks K with $R = 8$ (bps/Hz) and $\varepsilon_0 = 0.01$ for single user systems.

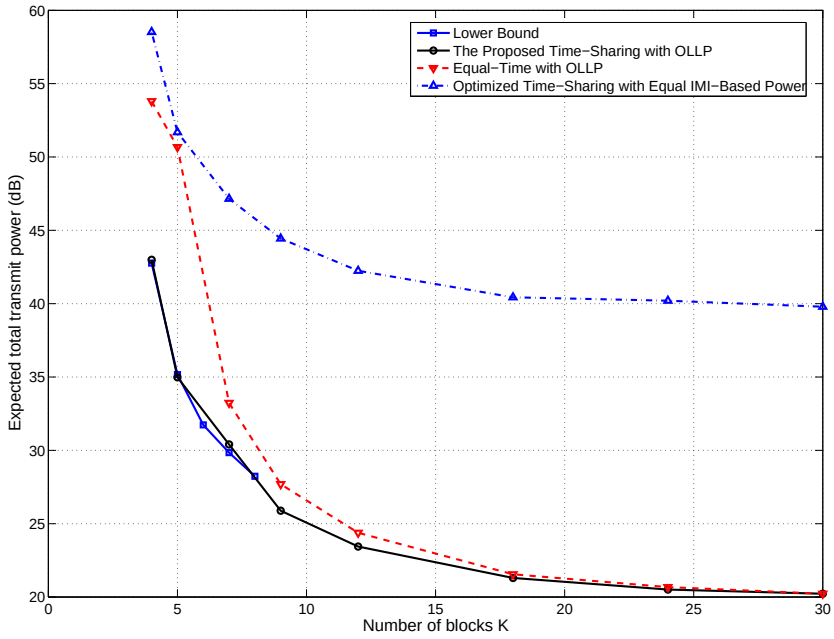

Figure 7.3: The expected total transmit power versus the number of blocks K with $U = 3$, $(R_1, R_2, R_3) = (8, 12, 16)$ (bps/Hz) and $(\varepsilon_1, \varepsilon_2, \varepsilon_3) = (0.1, 0.01, 0.001)$ for multi-user systems.

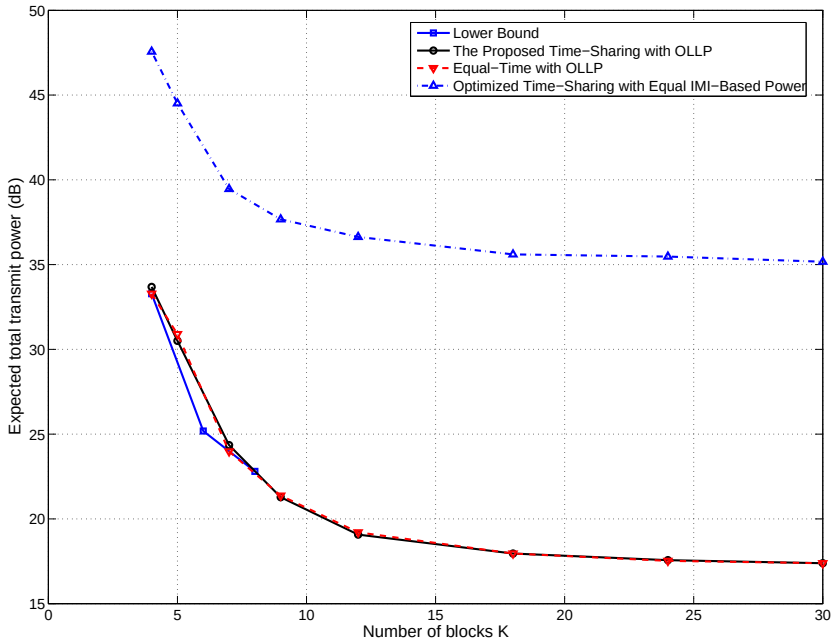

Figure 7.4: The expected total transmit power versus the number of blocks K with $U = 3$, $(R_1, R_2, R_3) = (8, 8, 8)$ (bps/Hz) and $(\varepsilon_1, \varepsilon_2, \varepsilon_3) = (0.01, 0.01, 0.01)$ for multi-user systems.

Chapter 8

Wireless Resource Allocation in Multi-user Systems with EC Constraints

In this chapter, our aim is to devise energy efficient strategies for single and multi-user MIMO systems with individual user's EC constraint. The use of EC provides a metric to quantify the maximum achievable rate under a certain queueing delay constraint. The objective is to minimize the transmission energy of the MIMO systems while meeting the EC constraints, with the aid of the channel statistics. Instead of considering decoding delay constraint from physical layer in Chapters 4-8, this chapter considers queueing delay constraint from data link layer.

The chapter is organized as follows. Section 8.1 overviews the motivation and aim of the research in this chapter. System model is presented in Section 8.2. A closed-form expression for the EC of a single user link MIMO system is derived in Section 8.3. In Section 8.4, we address multi-user MIMO systems, in which dynamic time sharing and multi-user power control are performed when TDMA is used. The results are extended to dynamic spectrum allocation in FDMA system

in Section 8.5. Simulation results are presented in Section 8.6 and Section 8.7 summaries this chapter.

8.1 Introduction

Due to the instability nature of wireless channels, supporting a given QoS is extremely challenging and guarantee of a deterministic delay-bound is not practical. In particular, delay bound violation probability has been used for measuring the performance of wireless networks, which allows a certain violation probability for a given delay bound. EC, which measures the maximum achievable source rate under a delay bound, has been presented as a QoS metric for wireless channels. The details of EC has been introduced in Section 2.3.

There are limited studies considering the provision of EC in wireless systems. In [102], an EC model is derived for frequency-selective fading channel. In [79, 103, 104, 105, 106], energy-efficient scheduling to satisfy each user's EC constraint has been considered by exploiting the CSI at the transmitter side. However, the CSI at the transmitter may not be possible in many applications and MIMO systems have not been considered in such cases.

In this chapter, we consider the power allocation and scheduling optimization for multi-user MIMO antenna systems, in which users are given individual EC constraint and operate in a TDMA or FDMA fashion, with the aid of SCI at the transmitter. Our contributions are that firstly, we derive a closed-form expression for the EC of a single user MIMO system, which is then applied to solve the power minimization problem. In addition, we show that the optimal solutions for the multi-user MIMO TDMA or FDMA systems can be obtained by DP. Thirdly, low-complexity heuristic approach for achieving near-optimum solutions is presented.

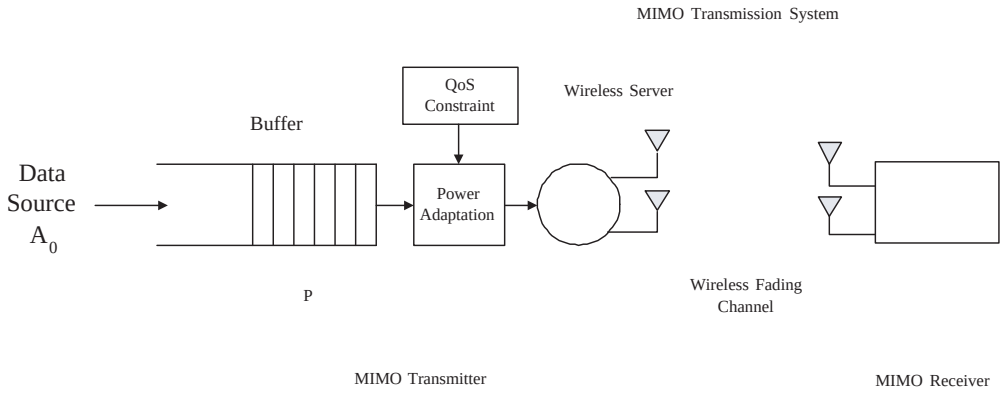

Figure 8.1: The system model of a single user MIMO system.

8.2 System Model

8.2.1 Single User MIMO Systems and EC

We first consider the single user MIMO system as shown in Fig. 8.1. The channel is assumed to be an independent discrete-time stationary and ergodic stochastic service process, which fades i.i.d. from one frame to another, but the fade is static within a frame length. The channel can be written as a matrix $\mathbf{H}(t) = \left[h_{i,j}^{(t)}\right] \in \mathbb{C}^{n_r \times n_t}$ for time t, where n_t and n_r denote the number of antennas at the transmitter and the receiver respectively. In particular, the amplitude square of each entry $\left|h_{i,j}^{(t)}\right|^2$ in $\mathbf{H}(t)$ has the p.d.f. of (8.1) as that of c,

$$f(c) = \begin{cases} e^{-c} & \text{for } c \geq 0, \\ 0 & \text{for } c < 0. \end{cases} \tag{8.1}$$

Since CSI is unavailable at the transmitter side, equal-power allocation over the antennas is adopted. Now, let B denote the bandwidth and T_f denote the frame length. In addition, we assume the noise power $N_0 = 1$ for normalization. Then,

we have the achievable rate for the frame at time t as [33]

$$R(t) = BT_f \log_2 \det \left| \mathbf{I} + \mathbf{H}(t)\mathbf{H}(t)^\dagger \left(\frac{P}{n_t} \right) \right| \text{ (bits/frame)} \tag{8.2}$$

where P is the total transmit power. As the processing is similar for different time, in the sequel, the index t will be omitted. In addition, we assume that transmitter has the knowledge of achievable rate R (i.e., a feedback link exists from receiver to transmitter for rate feedback), so that rate-adaptive transmissions and strong channel coding are used to transmit packets without error and the transmission rate is equal to R.

Data are assumed to arrive at the transmitter at a constant rate of $A(t) = A_0$, for all t. Before being transmitted to the user, they are stored in a buffer of infinite-length. Due to the time-varying nature of wireless channels, the service rate for the queueing data is time-varying. In what follows, the data stored in the buffer will experience delays. In particular, the probability of experiencing a delay exceeding a given bound $D_{\max}$ can be approximated by [77, 105] (Section 2.3)

$$\Pr\left(\text{Delay} > D_{\max}\right) \approx e^{-\theta A_0 D_{\max}}. \tag{8.3}$$

QoS assurance in a wireless system can therefore be recast as to achieve a given QoS exponent θ, which can be regarded as the exponential decay rate of the delay violation probability [77]. In addition, for an uncorrelated channel service process, EC can be expressed as [77, 79] (Section 2.3)

$$E_c(\theta, P) = -\frac{1}{\theta} \log_2 \mathrm{E}_{\mathbf{H}} \left[e^{-\theta R(P)} \right]. \tag{8.4}$$

Under this single user system model, our aim is to minimize the required transmit power to achieve a given QoS exponent (delay constraint) θ, when an EC of A_0 is

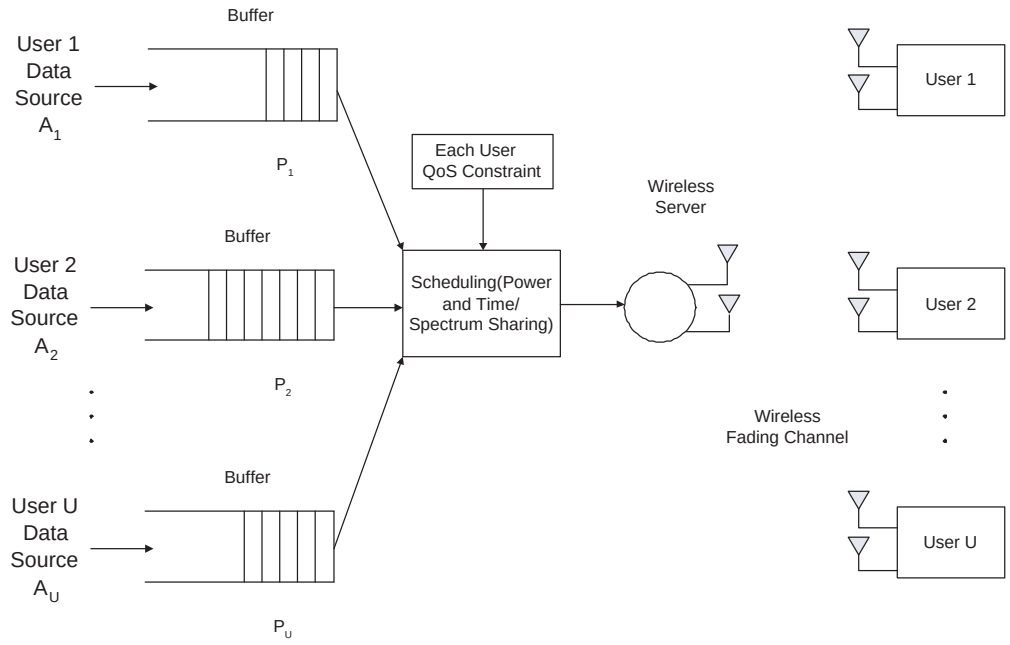

Figure 8.2: The system model for a multi-user MIMO system.

to be attained, by exploiting the SCI at the transmitter. Mathematically, that is,

$$\mathbb{P}_{\mathsf{SU}} : \min_{P>0} P \quad \text{s.t.} \quad E_c(\theta, P) = A_0. \tag{8.5}$$

8.2.2 Multi-user MIMO Systems

Extending the above model to a system with multiple users as shown in Fig. 8.2, each user is given a QoS requirement specified by (A_u, θ_u) for $u = 1, 2, \ldots, U$ where U is the total number of users. In this chapter, we consider that users are served in either a TDMA or FDMA fashion so that at any given time (or frequency), only one user is accommodated and users are orthogonal to one another. For convenience, we assume that each user has the same number of antennas n_r and the users' channels are i.i.d. and denoted as $\{\mathbf{H}_u\}_{u=1,2,\ldots,U}$.

Our aim for the multi-user system is to minimize the total transmit power $\sum_{u=1}^{U} P_u$ while satisfying each user's QoS constraint (A_u, θ_u) by jointly optimizing the power allocation $\{P_u\}$ and the number of time-slots (or subcarriers) $\{\alpha_u\}$ (or $\{\beta_u\}$) with

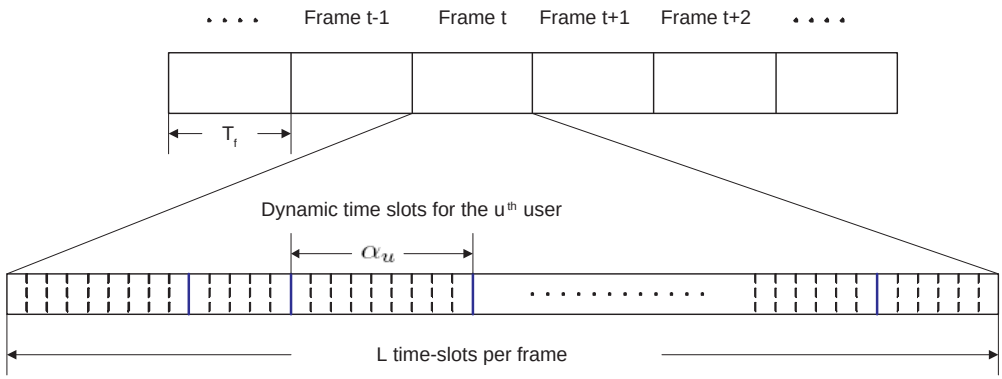

Figure 8.3: The adaptive frame structure for the multi-user TDMA system.

the aid of SCI (see Fig. 8.3 for the adaptive frame structure for TDMA). In the case of TDMA, we thus have

$$\begin{aligned} \mathbb{P}_{\mathsf{TDMA}}: \min_{\{P_u\},\{\alpha_u\}} \quad & \sum_{u=1}^{U} \alpha_u P_u \\ \text{s.t.} \quad & \sum_{u=1}^{U} \alpha_u \le L \\ & -\frac{1}{\theta_u} \log \mathbf{E}_{\mathbf{H}_u}\left[e^{-\theta_u R_u}\right] = A_u \forall u \\ & R_u = \alpha_u B T_0 \log_2 \det\left|\mathbf{I} + \mathbf{H}_u \mathbf{H}_u^{\dagger}\left(\frac{P_u}{n_t}\right)\right|, \end{aligned} \tag{8.6}$$

where L, T_0 and B denote, respectively, the total number of time-slots per frame, the period for each time-slot and the total transmission bandwidth. It is assumed that the channel within each frame remains static, while it will fade independently and identically from one frame to another.

For FDMA with M subcarriers, the problem becomes

$$\begin{aligned}\mathbb{P}_{\mathsf{FDMA}}: \min_{\{P_u\},\{\beta_u\}} \quad & \sum_{u=1}^{U} \beta_u P_u \\ \text{s.t.} \quad & \sum_{u=1}^{U} \beta_u \le M \\ & -\frac{1}{\theta_u}\log \mathtt{E}_{\mathbf{H}_u}\left[e^{-\theta_u R_u}\right] = A_u \forall u \\ & R_u = \sum_f B_0 T_f \log_2 \det\left|\mathbf{I} + \mathbf{H}_u(f)\mathbf{H}_u(f)^\dagger\left(\frac{P_u}{n_t}\right)\right|,\end{aligned} \tag{8.7}$$

where T_f and B_0 denote, respectively, the frame length and the bandwidth for each subcarrier. Similarly, the channels at different subcarriers and for different users are i.i.d..

8.3 The Optimal Power Allocation for single user MIMO Systems

In this section, we solve $\mathbb{P}_{\mathsf{SU}}$ by first deriving a closed-form expression for the EC of a single user MIMO channel. This is done by first substituting (8.2) into (8.4) to yield

$$E_c(\theta, P) = -\frac{1}{\theta}\log \mathtt{E}_{\mathbf{H}}\left[e^{-\theta B T_f \log_2 \det\left|\mathbf{I}+\mathbf{H}\mathbf{H}^\dagger\left(\frac{P}{n_t}\right)\right|}\right]. \tag{8.8}$$

By using the random matrix theory results in [107, 108], we have the EC expression presented in the following theorem.

Theorem 1: (A closed-form expression for EC) Denoting $k = \min(n_t, n_r)$ and $d = \max(n_t, n_r) - k$, the EC of an uncorrelated MIMO central Wishart channel is given by

$$E_c(\theta, P) = -\frac{1}{\theta}\log \frac{\det \mathbf{G}(\theta, P)}{\prod_{i=1}^{k}\Gamma(d+i)}, \tag{8.9}$$

in which $\Gamma(z) = \int_0^\infty x^{z-1}e^{-x}dx$ and $\mathbf{G}$ is a $k \times k$ Hankel matrix whose (i,j)th entry is

$$\begin{aligned}[\mathbf{G}(\theta, P)]_{i,j} &= \int_0^\infty e^{-x}x^{i+j+d}\left(1+\frac{P}{n_t}x\right)^{-s(\theta)}dx \\ &= \left(\frac{p}{n_t}\right)^{-(1+i+j+d)}\Gamma(1+i+j+d)\times \\ &\quad U\left(1+i+j+d, 2+i+j+d-s(\theta); \frac{n_t}{P}\right),\end{aligned} \tag{8.10}$$

where $s(\theta) = \frac{\theta BT_f}{\ln 2}$, and $U(\alpha, \gamma; z)$ is the confluent hypergeometric functions defined as [95, (9.210)]

$$U(\alpha,\gamma;z) = 1 + \frac{\alpha}{\gamma}\frac{z}{1!} + \frac{\alpha(\alpha+1)}{\gamma(\gamma+1)}\frac{z^2}{2!} + \frac{\alpha(\alpha+1)(\alpha+2)}{\gamma(\gamma+1)(\gamma+2)}\frac{z^3}{3!} + \cdots. \tag{8.11}$$

Using Theorem 1, the optimal power allocation strategy for a single user MIMO system to meet a given EC constraint can be found by numerically solving

$$-\frac{1}{\theta}\log\frac{\det\mathbf{G}(\theta,P)}{\prod_{i=1}^k\Gamma(d+i)} = A_0. \tag{8.12}$$

8.4 Multi-user MIMO-TDMA with EC Constraints

8.4.1 The Optimal DP Solution

According to Theorem 1, the EC for user u in the TDMA system is therefore given by (8.9) with θ, P and $s(\theta)$ replaced by, respectively, θ_u, P_u and $s(\theta_u, \alpha_u) = \frac{\alpha_u\theta_u BT_0}{\ln 2}$. As such, the minimum power P_u with a given number of allocated time-slots $1 \le \alpha_u \le L+1-U$ can be found by solving

$$E_c(\theta_u, P_u, \alpha_u) = A_u. \tag{8.13}$$

As a result, by going through every possible combination of users' allocated time-slots $\{\alpha_u\}$, it is possible to find the joint-optimal $\{P_u, \alpha_u\}$ for total power minimization. Remarkably, the required optimization can be achieved by formulating the system as a U-stage dynamic system, with the system state x_u representing the number of remaining time-slots for users u to U, the control parameter l_u representing the number of allocated time-slots for user u and the cost function $c_u(l_u) = l_u P_u(l_u)$ denoting the power consumption for user u. The cost accumulated from stage u to U can be found as

$$J_u(x_u) = \min_{1 \le l_u \le x_u} c_u(l_u) + J_{u+1}(x_u - l_u) \text{ for } 1 \le x_u \le L, \tag{8.14}$$

with $J_{U+1} = 0$. In what follows, the total system cost is given by $J_1(x_1)$, and the optimal control policy $\boldsymbol{l}^* = (l_1^*, l_2^*, \ldots, l_U^*)$ is the one that minimizes the cost and can be found from (8.14) by proceeding backward for $u = U, U-1, \ldots, 1$ [83]. Then, the optimal time-slot allocation can be found from

$$\begin{cases} \alpha_u(x_u) = l_u^*(x_u), \\ x_{u+1} = L - \sum_{i=1}^{u} \alpha_i, \\ x_1 = L. \end{cases} \tag{8.15}$$

8.4.2 A Suboptimal Convex Optimization Approach

Because of the high-complexity of DP, we now develop a heuristic solution by convex optimization. Note that the EC expression involves a complex function in $\mathbf{G}$, further analysis is extremely difficult and we therefore resort to relaxation to bound the EC. In particular, we rewrite the QoS constraint as

$$\mathrm{E}_{\mathbf{H}_u}\left[e^{-\theta_u B \alpha_u T_0 \log_2 \det\left|\mathbf{I}+\mathbf{H}_u \mathbf{H}_u^{\dagger}\left(\frac{P_u}{n_t}\right)\right|}\right] \ge e^{-\theta_u A_u}. \tag{8.16}$$

Writing $\delta_u = \frac{\theta_u B T_0}{\ln 2}$, we have

$$\begin{aligned}
&\mathtt{E}_{\mathbf{H}_u}\left[e^{-\alpha_u\delta_u \ln\det\left|\mathbf{I}+\mathbf{H}_u\mathbf{H}_u^\dagger\left(\frac{P_u}{n_t}\right)\right|}\right] \\
=&\mathtt{E}_{\mathbf{H}_u}\left[\det\left|\mathbf{I}+\mathbf{H}_u\mathbf{H}_u^\dagger\left(\frac{P_u}{n_t}\right)\right|^{-\alpha_u\delta_u}\right] \\
\geq&\left(\mathtt{E}_{\mathbf{H}_u}\left[\det\left|\mathbf{I}+\mathbf{H}_u\mathbf{H}_u^\dagger\left(\frac{P_u}{n_t}\right)\right|\right]\right)^{-\alpha_u\delta_u} \\
\geq&\left(\mathtt{E}_{\mathbf{H}_u}\left[\det\left|\mathbf{H}_u\mathbf{H}_u^\dagger\left(\frac{P_u}{n_t}\right)\right|\right]\right)^{-\alpha_u\delta_u} \\
=&\left(\frac{P_u}{n_t}\right)^{-\alpha_u\delta_u}\left(\mathtt{E}_{\mathbf{H}_u}\left[\det\left|\mathbf{H}_u\mathbf{H}_u^\dagger\right|\right]\right)^{-\alpha_u\delta_u}.
\end{aligned} \tag{8.17}$$

The first inequality comes from Jensen's Inequality introduced in Section 3.1. The original QoS constraint (8.16) can be guaranteed by ensuring that the above lower bound equals to the right-hand-side of (8.16), which gives the power allocation for user u as

$$P_u = n_t e^{\frac{\theta_u A_u}{\alpha_u\delta_u}}\left(\mathtt{E}_{\mathbf{H}_u}\left[\det\left|\mathbf{H}_u\mathbf{H}_u^\dagger\right|\right]\right)^{-1}. \tag{8.18}$$

In addition, it is known in [107] that

$$\mathtt{E}_{\mathbf{H}_u}\left[\det\left|\mathbf{H}_u\mathbf{H}_u^\dagger\right|\right] = \prod_{i=0}^{k-1}\frac{\Gamma(d+k-i+1)}{\Gamma(d+k-i)}. \tag{8.19}$$

As a consequence, $\mathbb{P}_{\mathsf{TDMA}}$ can be reformulated as

$$\begin{aligned}
\tilde{\mathbb{P}}_{\mathsf{TDMA}}: \min_{\{\alpha_u\}}\ & \sum_{u=1}^{U} n_t\alpha_u e^{\frac{\theta_u A_u}{\alpha_u\delta_u}}\left(\mathtt{E}\left[\det\left|\mathbf{H}_u\mathbf{H}_u^\dagger\right|\right]\right)^{-1} \\
\text{s.t.}\ & \begin{cases} \sum_{u=1}^{U}\alpha_u = L \\ \alpha_u = \{1,2,\ldots,L+1-U\}\ \forall u. \end{cases}
\end{aligned} \tag{8.20}$$

This combinatorial problem, $\tilde{\mathbb{P}}_{\mathsf{TDMA}}$, can then be simplified by relaxing $\{\alpha_u\}$ as positive real numbers, such that we have

$$\begin{aligned} \min_{\{0\leq\alpha_u\leq L+1-U\}} \quad & \sum_{u=1}^{U} n_t \alpha_u e^{\frac{\theta_u A_u}{\alpha_u \delta_u}} \left(\mathtt{E}\left[\det\left|\mathbf{H}_u\mathbf{H}_u^{\dagger}\right|\right]\right)^{-1} \\ \text{s.t.} \quad & 0 \leq \sum_{u=1}^{U} \alpha_u \leq L. \end{aligned} \tag{8.21}$$

It can be easily seen that (8.21) is a convex problem and therefore can be optimally and efficiently solved [84]. Nevertheless, the solutions obtained from (8.21) are not the feasible solutions for (8.20). To address this, we present the following algorithm:

1. Initialize $\alpha_u = \lceil \alpha_u \rceil \ \forall u$, where $\lceil y \rceil$ returns the smallest integer that is larger than y;

2. For each user u, compute the minimal required power $P_u(\alpha_u, A_u, \theta_u)$ by solving (8.13) to ensure that a given delay constraint is met;

3. Then, initialize $v = L - \sum_{u=1}^{U} \alpha_u$;

4. Compute the power reduction metrics

$$\triangle P_u = \arg\min_{P_u \geq 0} P_u(\alpha_u - 1, A_u, \theta_u) - P_u \ \forall u; \tag{8.22}$$

5. Find $u^* = \arg\min_u \triangle P_u$ and update

$$\begin{cases} \alpha_{u^*} := \alpha_{u^*} - 1, \\ v := v - 1. \end{cases} \tag{8.23}$$

If $v \geq 1$, go back to Step 4. Otherwise, the optimization is completed and $\{\alpha_u\}$ and $\{P_u\}$ are obtained.

8.5 Multi-user MIMO-FDMA with EC Constraints

8.5.1 The Optimal DP Solution

In the multi-user MIMO-FDMA system, if β_u is fixed, then the EC can be obtained as

$$\begin{aligned} E_c(\theta_u, P_u, \beta_u) &= -\frac{1}{\theta_u} \log \mathtt{E}\left[e^{-\theta_u T_f B_0 \sum_f \log_2 \det\left|\mathbf{I}+\mathbf{H}_u(f)\mathbf{H}_u(f)^\dagger\left(\frac{P_u}{n_t}\right)\right|} \right] \\ &= -\frac{1}{\theta_u} \log \mathtt{E}\left[e^{-\theta_u T_f B_0 \log_2 \det\left|\mathbf{I}+\mathbf{H}_u(f)\mathbf{H}_u(f)^\dagger\left(\frac{P_u}{n_t}\right)\right|} \right]^{\beta_u}. \end{aligned} \tag{8.24}$$

Using Theorem 1, we can then obtain

$$E_c(\theta_u, P_u, \beta_u) = -\frac{1}{\theta_u} \log \left(\frac{\det \mathbf{G}(\theta_u, P_u)}{\prod_{i=1}^{k} \Gamma(d+i)} \right)^{\beta_u}, \tag{8.25}$$

where the elements of $\mathbf{G}(\theta_u, P_u)$ is given in (8.10) with $s_u = \frac{\theta_u B_0 T_f}{\ln 2}$. Therefore the minimum power P_u can be solved by

$$E_c(\theta_u, P_u, \beta_u) = A_u. \tag{8.26}$$

Similarly as in the case of TDMA, the joint-optimal solution for $\{P_u, \beta_u\}$ can be found by a U-stage DP at the expense of complexity (see Section 8.4.1). The details are omitted here.

8.5.2 A Suboptimal Convex Optimization Approach

With a view to reducing the complexity for solving $\mathbb{P}_{\mathsf{FDMA}}$, we investigate the use of relaxation and convex optimization as in Section 8.4.2. To do this, we first

rewrite the QoS constraint for user u as

$$\mathtt{E}_{\mathbf{H}_u}\left[e^{-\theta_u T_f B_0 \log_2 \det\left|\mathbf{I}+\mathbf{H}_u(f)\mathbf{H}_u(f)^\dagger\left(\frac{P_u}{n_t}\right)\right|}\right]^{\beta_u} = e^{-\theta_u A_u}. \tag{8.27}$$

Also, we can obtain the following lower bound

$$\begin{aligned}
&\mathtt{E}_{\mathbf{H}_u}\left[e^{-\theta_u T_f B_0 \log_2 \det\left|\mathbf{I}+\mathbf{H}_u(f)\mathbf{H}_u(f)^\dagger\left(\frac{P_u}{n_t}\right)\right|}\right]^{\beta_u} \\
=&\mathtt{E}_{\mathbf{H}_u}\left[\det\left|\mathbf{I}+\mathbf{H}_u(f)\mathbf{H}_u(f)^\dagger\left(\frac{P_u}{n_t}\right)\right|^{-s_u}\right]^{\beta_u} \\
\geq&\left(\mathtt{E}_{\mathbf{H}_u}\left[\det\left|\mathbf{I}+\mathbf{H}_u(f)\mathbf{H}_u(f)^\dagger\left(\frac{P_u}{n_t}\right)\right|^{-s_u}\right]\right)^{\beta_u} \\
\geq&\left[\left(\mathtt{E}_{\mathbf{H}_u}\left[\det\left|\mathbf{H}_u(f)\mathbf{H}_u(f)^\dagger\left(\frac{P_u}{n_t}\right)\right|\right]\right)^{-s_u}\right]^{\beta_u} \\
=&\left[\left(\frac{P_u}{n_t}\right)^{-s_u}\left(\mathtt{E}_{\mathbf{H}_u}\left[\det\left|\mathbf{H}_u(f)\mathbf{H}_u(f)^\dagger\right|\right]\right)^{-s_u}\right]^{\beta_u}.
\end{aligned} \tag{8.28}$$

The EC constraint can therefore be ascertained by having

$$P_u = n_t e^{\frac{\theta_u A_u}{s_u \beta_u}}\left(\mathtt{E}_{\mathbf{H}_u}\left[\det\left|\mathbf{H}_u(f)\mathbf{H}_u(f)^\dagger\right|\right]\right)^{-1}. \tag{8.29}$$

The optimization problem $\mathbb{P}_{\mathsf{FDMA}}$ can hence be reformulated as

$$\begin{aligned}
\tilde{\mathbb{P}}_{\mathsf{FDMA}}: \min_{\{\beta_u\}} \quad & \sum_{u=1}^{U} n_t \beta_u e^{\frac{\theta_u A_u}{\beta_u s_u}}\left(\mathtt{E}\left[\det\left|\mathbf{H}_u(f)\mathbf{H}_u(f)^\dagger\right|\right]\right)^{-1} \\
\text{s.t.} \quad & \begin{cases} \sum_{u=1}^{U} \beta_u = M, \\ \beta_u = \{1, 2, \ldots, M+1-U\}\ \forall u, \end{cases}
\end{aligned} \tag{8.30}$$

the solution $\{\beta_u\}$ can be found similarly using the method in Section 8.4.2. Due to the relaxation, the solution is in general suboptimal. The performance of this heuristic method will be evaluated by simulation results in Section 8.6.

8.6 Simulation Results

In this section, we provide simulation results for both single and multi-user MIMO systems. Fig. 8.4 shows the normalized EC results for a single user (3, 2) system under various power constraints, P. In the simulations, we have assumed $B = 10$ (kHz) and $T_f = 20$ (ms). Results in this figure illustrate perfect agreement for the analytical results presented in Theorem 1 with the Monte-Carlo simulation results. In addition, we can see that EC decreases with the increasing of QoS exponent θ, which indicates the tradeoff between delay performance and the supported source rate. It is also observed that for a given θ, greater transmission power leads to a higher EC.

The simulation results for multi-user scenarios are shown in Figs. 8.5 and 8.6. In these simulations, it is considered that $U = 3$, $n_t = 3$ and $n_r = 2$. In Fig. 8.5, the results are provided for MIMO-TDMA systems while the results for MIMO-FDMA systems are plotted in Fig. 8.6. We have furthermore assumed that $B = 10$ (kHz), $T_0 = 1$ (ms) and $T_f = 20$ (ms). In addition, users are required to have $\{A_1, A_2, A_3\} = \{0.08, 0.32, 1\}$ (kbits/frame) (or $\{4, 16, 50\}$ (kbps) normalized rate) with the QoS exponents $\{\theta_1, \theta_2, \theta_3\} = \{10^{-1}, 10^{-2}, 10^{-3}\}$. As can be seen, the proposed suboptimal solution performs as well as the optimal DP solution with inappreciable difference. Moreover, it is possible to have power saving as large as 10 (dB) between systems with and without dynamic time-slot allocation.

The MIMO-FDMA results in Fig. 8.6 have considered $B_0 = 1$ (kHz), $B = 20$ (kHz) and $T_f = 10$ (ms). The users' requirements are specified by $\{A_1, A_2, A_3\} = \{0.08, 0.32, 1.32\}$ (kbits/frame) (or $\{8, 32, 132\}$ (kbps) normalized rate) with the QoS exponents $\{\theta_1, \theta_2, \theta_3\} = \{10^{-1}, 10^{-2}, 10^{-3}\}$. Again, it is seen that the proposed suboptimal solution and the optimal DP solution perform similarly, and there is a significant gain observed by systems with dynamic subcarrier allocation.

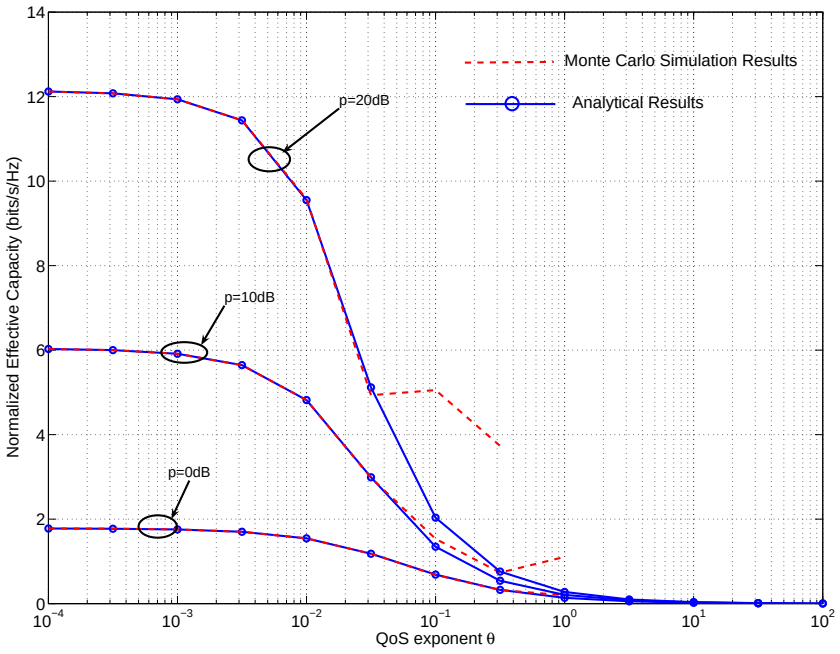

Figure 8.4: The normalized EC for (3,2) systems.

8.7 Summary

In this chapter, we investigated the power minimization problems for multi-user MIMO-TDMA and MIMO-FDMA systems with individual user's EC constraint. By jointly optimizing the power allocation and the number of time-slot (or sub-carrier) allocation for all of the users with the aid of SCI at the transmitter, we have shown that a significant power saving is achieved. In particular, we have demonstrated that the optimal solutions can be obtained by DP and near-optimal heuristics at much reduced complexity are possible.

Chapters 4-8 of the thesis concentrate on theoretical investigation of wireless resource allocation for delay constrained communications. In the following chapter, we will modify and apply the theoretic framework into wireless standard IEEE 802.16 WiMAX.

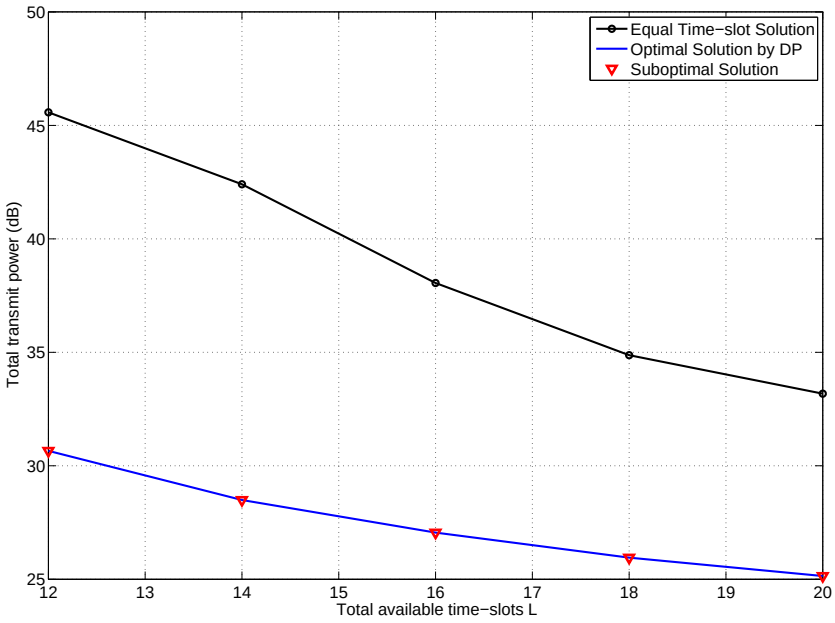

Figure 8.5: The required transmit power for 3-user MIMO-TDMA systems.

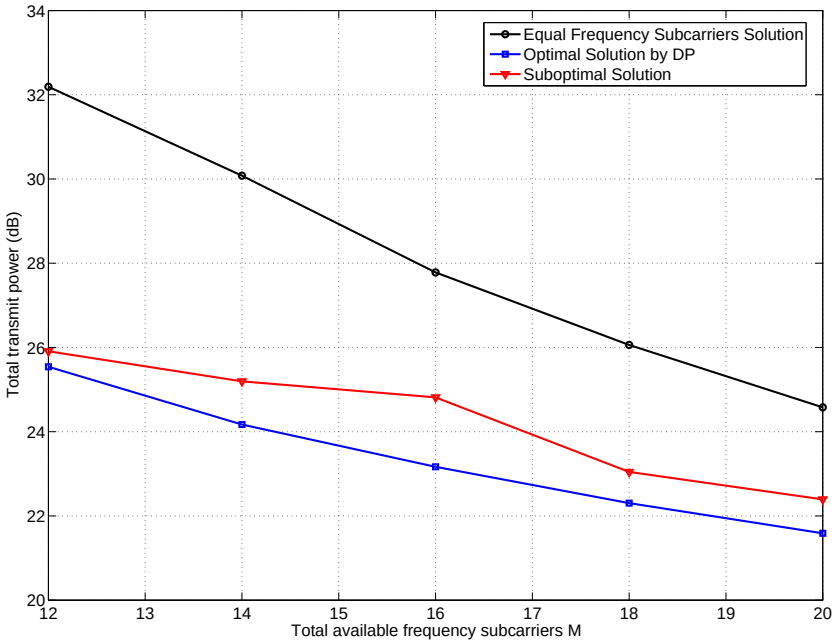

Figure 8.6: The required transmit power for 3-user MIMO-FDMA systems.

Chapter 9

Scheduling in WiMAX

In this chapter, we will apply the developed framework in Chapters 4-8 to IEEE 802.16e mobile WiMAX. Based on a BT WiMAX simulation platform, we modify the theoretical framework in the previous chapters, develop uplink scheduling algorithms and further validate the results under different system settings. The results in this chapter prove the significance of the research work in the thesis for ways to design energy efficient wireless resource allocation strategies.

This chapter is organized as follows. Section 9.1 introduces the motivation for the chapter and the key specifications of the physical layer IEEE 802.16e based WiMAX. The resource allocation framework is modified for WiMAX uplink scheduling for the single user case and multi-user case respectively in Sections 9.2 and 9.3. At the end, Section 9.4 summarises this chapter.

9.1 Introduction

The IEEE 802.16 standard aims at providing wireless access over long distance in a variety of ways: point-to-point links and full mobile cellular type access as shown in Fig. 9.1 [109]. Unlike WLAN where no QoS can be guaranteed, WiMAX

considers several types of QoS mechanisms for services such as data, voice and video. Therefore, it is crucial to develop resource allocation mechanisms for QoS guarantees such as delay, delay jitter and throughput requirements. As scheduling mechanisms in WiMAX are not standardized by the IEEE and implementation differentiations are allowed, scheduling design is of special interest to all WiMAX equipment vendors and network providers.

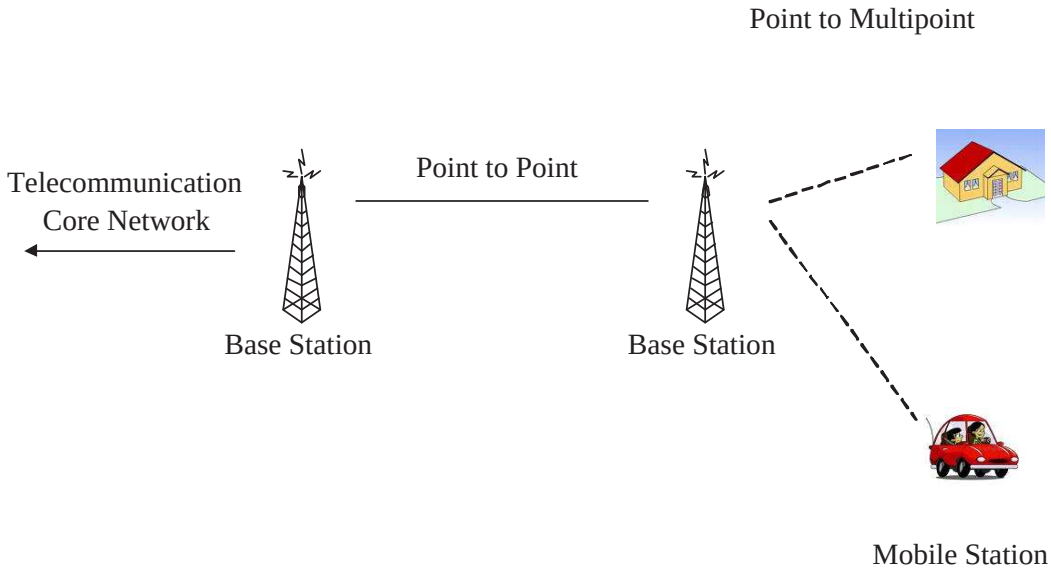

Figure 9.1: An illustration of WiMAX.

In this chapter, we modify the theoretical framework in the previous chapters and develop energy efficient scheduling for WiMAX uplink, which refers to the link from mobile terminals to base station. The reason we consider uplink instead of downlink lies in three aspects. First, transmit energy accounts for a major proportion of energy consumption for the uplink. Mobile devices are usually more limited in energy consumption than the base station. Due to the above reason, although in the downlink, the base station tends to transmit at the maximum available power, for the uplink, the power allocation becomes more important as only a limited amount of energy is available. Second, optimal uplink scheduling is much more difficult to achieve. This is because the trade-off between transmit power and transmission bandwidth is different on the uplink compared to the downlink.

For example, two mobile stations can simultaneously transmit at full power to the base station using half of the channel bandwidth, whereas on the downlink, partitioning the traffic in frequency does not yield any more transmit power. Third, the cell size of WiMAX is often limited by the uplink performance and in particular inter-cellular interference. By designing energy efficient scheduling for the uplink, the performance of cell edge users can be significantly improved, which leads to better cell coverage. A reduction in the number of base stations required for achieving a given coverage brings cost reduction for network providers. Note that although we have taken 802.16e WiMAX as a platform to apply the uplink scheduling algorithms, they can be equally well applied to other technologies, such as LTE.

9.1.1 A Brief Review of the WiMAX Physical Layer

In order to understand the uplink scheduling design, in the following sections, we review some of the key features of the WiMAX physical layer, which are the use of orthogonal frequency division multiplexing (OFDM) and advanced multiple antenna techniques, such as MIMO [51]. OFDM is a modulation method to use a large number of closely-spaced orthogonal subcarriers to carry data. Consequently, the input data stream is divided into several parallel sub-streams of reduced symbol rate and each sub-stream is modulated and transmitted on a separate orthogonal subcarrier. By applying this, the symbol duration is increased, which improves the robustness of OFDM to delay spread.

For multiple antenna techniques, maximum ratio combining (MRC), space time coding (STC) and spatial multiplexing (SM) are supported in WiMAX. In MRC, multiple copies of the received signal are combined at the receiver weighted with respect to their SNR. In STC, spatial diversity is provided by using Alamouti coding to reduce the error probability [51, 3]. In SM, multiple streams are transmitted over multiple antennas. For example, with 2×2 MIMO system, the SM

increases the data rate two-fold by transmitting two independent data stream over the antennas [1]. The WiMAX downlink support all the above multiple antenna techniques, while, WiMAX uplink only support MRC.

The WiMAX physical layer also supports different modulation schemes, of which the basic principle is to encode an information bit stream into a carrier signal which is then transmitted over a communications channel. The modulation types are amplitude and phase modulation, in which the information bit stream is encoded in the amplitude and/or phase of the transmitted signal [2]. In Table 9.1, we have summarized the modulation schemes used in WiMAX system. Here, Phase Shift Keying (PSK) means all the information is encoded in the phase of the transmit signal. For Quadrature Amplitude Modulation (QAM), the information bits are encoded in both the amplitude and phase of the transmitted signal. The notation m within mPSK/QAM represent the amount of information carried by the modulated signal. Therefore, the number of bits required to carry m amount of information is $\log_2 m$. In addition, QPSK is short for Quadrature Phase Shift Keying where $m = 4$ in this case. n-CTC stands for Convolutional Turbo Coding where n is the CTC code rate. tX is repetition code indicating the original information is reproduced t times. For a specified modulation type, the number of bits per subcarrier B_s is calculated as

$$B_s = n \times \frac{1}{t} \times \log_2 m. \tag{9.1}$$

For each of the OFDM transmission frame, a number of subchannels, which are groups of data and pilot subcarriers, are dedicated to a certain user, where subcarriers for pilots are used for estimation and synchronization purposes. Each frame contains a number of OFDM symbols and the number of symbols can be adjusted between downlink and uplink. The physical layer parameters for WiMAX are

[1] Note that the actual data rate increase is between 1.5 and 2 due to the time overheads taken into account.

Table 9.1: Modulation and coding scheme in WiMAX.

Modulation	Code Rate	Downlink	Bit Rate (bits/subcarrier)	Uplink	Bit Rate (bits/subcarrier)
QPSK	1/2-CTC,6X	QPSK 1/12	1/6	QPSK 1/12	1/6
	1/2-CTC,4X	QPSK 1/8	1/4	QPSK 1/8	1/4
	1/2-CTC,2X	QPSK 1/4	1/2	QPSK 1/4	1/2
	1/2-CTC,1X	QPSK 1/2	1	QPSK 1/2	1
	3/4-CTC	QPSK 3/4	1.5	QPSK 3/4	1.5
16QAM	1/2-CTC	16-QAM 1/2	2	16-QAM 1/2	2
	3/4-CTC	16-QAM 3/4	3	16-QAM 3/4	3
64QAM	1/2-CTC	64-QAM 1/2	3	-	-
	2/3-CTC	64-QAM 2/3	4	-	-
	3/4-CTC	64-QAM 3/4	4.5	-	-
	5/6-CTC	64-QAM 5/6	5	-	-

listed in Table 9.2 for a 10MHz partial usage of subcarriers (PUSC), where part of the subcarriers are used for data transmission. The parameters for both uplink and downlink are listed for completeness. In the later sections, we will only refer to the parameters for uplink data. An example to calculate the data rate R within a frame for uplink transmission under a fixed modulation type is

$$R = \frac{1}{F_l} \times S_f \times C_a \times N \times (1 - O_h) \times B_s \qquad \text{in (bps)}, \tag{9.2}$$

where N is the number of subchannels in use. However, the actual number of bits carried by the modulated signal is dependent on the SNR according to Shannon's theory. We use $f_T(\gamma)$ to represent the actual throughput (bits/subcarrier) as the function of receive SNR γ and the data rate is rewritten as

$$R = \frac{1}{F_l} \times S_f \times C_a \times N \times (1 - O_h) \times f_T(\gamma) \qquad \text{in (bps)}. \tag{9.3}$$

Table 9.2: System settings for WiMAX with 10MHz channel and PUSC.

Parameter	Downlink	Uplink
Subcarriers for data	720	560
Subcarriers for pilot	120	280
Symbols/frame (S_f)	29	18
Number of subchannels	30	35
Subcarriers/subchannel (C_a)	24	16
Proportion of overhead (O_h)	0.3	0.25
Frame length (ms) (F_l)	5	5
Frame symbol rate (symbols/second)	200	200
Bandwidth (MHz)	10	10

9.2 Uplink Scheduling for Single User System

9.2.1 Uplink Throughput

Based on a BT simulation system, we have WiMAX uplink throughput plotted in Fig. 9.2. As only MRC is supported by uplink and the highest order of modulation and coding scheme is 16-QAM 3/4, the maximum throughput of the uplink is 3 (bits/subcarrier) shown in Table 9.1. In order to further design the scheduling algorithm, we can estimate the throughput results closely using a polynomial approximation as

$$f_T(\gamma) = \max(0, \min(6 \times 10^{-5} \times \gamma^3 + 0.0042 \times \gamma^2 + 0.0749 \times \gamma + 0.3792, 3)). \quad (9.4)$$

For a single user system, the relationship between transmit power P and the number of subchannels K used is investigated in this section. We assume receive noise power $N_0 = 1$, therefore, the transmit power P equals to the SNR γ. By substituting (9.4) into (9.3), we can numerically solve (9.3) to find the amount of transmit power P required for a given data rate R, where we rewrite the equation as

$$\frac{1}{F_l} \times S_f \times C_a \times N \times (1 - O_h) \times f_T(\frac{P}{N}) = R. \quad (9.5)$$

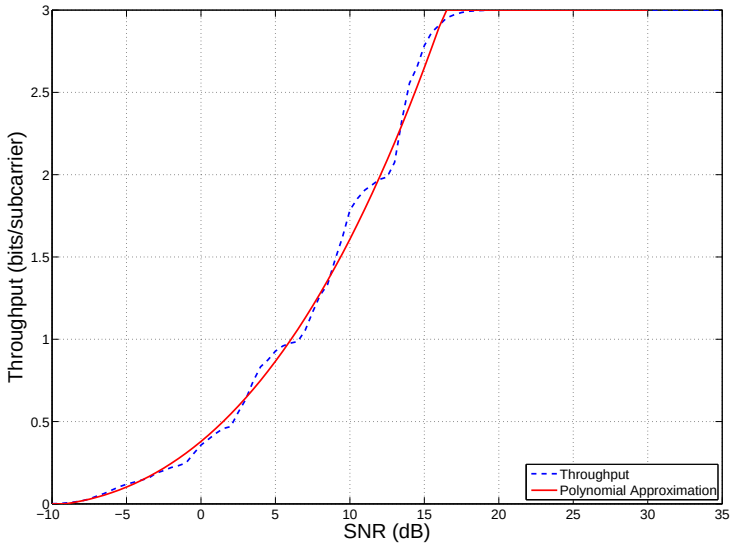

Figure 9.2: WiMAX uplink throughput from BT simulation system.

In Fig. 9.3, we plot the required transmit power versus the number of subchannels for different required transmission data rates. It is shown that the optimal number of subchannels used to achieve the minimum transmit power varies for different data rates. In addition, it indicates that using a greater number of subchannels may not be able to achieve better performance. The reason is because with the limited transmit power for the uplink, increasing the number of subchannels results in decreasing of SNR, which may lead to a lower data rate.

9.2.2 Problem Formulation and Solutions

In a single user system, the energy efficient uplink scheduling problem aims to minimize the transmit power by adjusting the number of subchannels N while satisfying the user's data rate requirement from (9.3). Mathematically, the problem is constructed as

$$\begin{aligned} \min_{N} \quad & P \\ s.t. \quad & \frac{1}{F_l} \times S_f \times C_a \times N \times (1 - O_h) \times f_T(\frac{P}{N}) = R. \end{aligned} \tag{9.6}$$

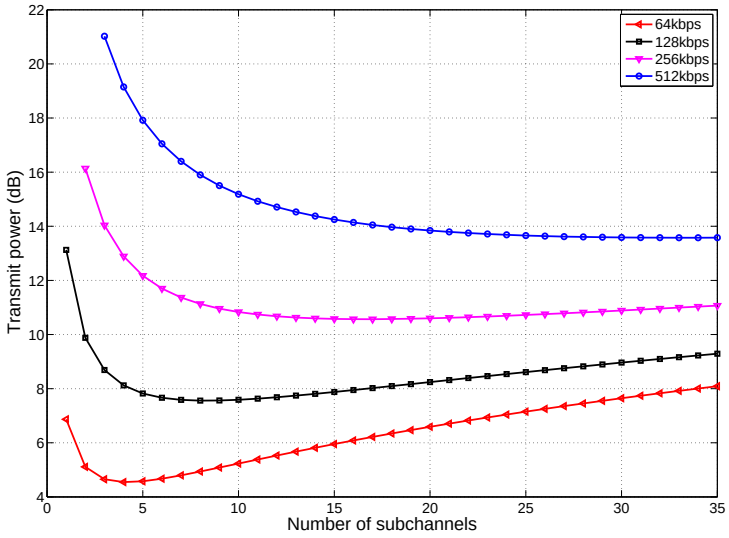

Figure 9.3: Required transmit power versus the number of subchannels.

The problem can be solved by going through all the possible N to find the minimum power P^*, based on which, the optimal number of subchannels N^* and throughput $f_T(\frac{P^*}{N^*})$ can be derived. The modulation scheme can be consequently obtained by referring to Table 9.1.

Fig. 9.4 shows the optimal number of subchannels allocated as the function of required data rate. It shows that the optimal number of subchannels increases with transmission data rate. Furthermore, we list the optimal modulation scheme used for different transmission data rates in Table 9.3. Interestingly, it is observed that QPSK 1/2 is the modulation scheme used for achieving minimum power for all different transmission data rates.

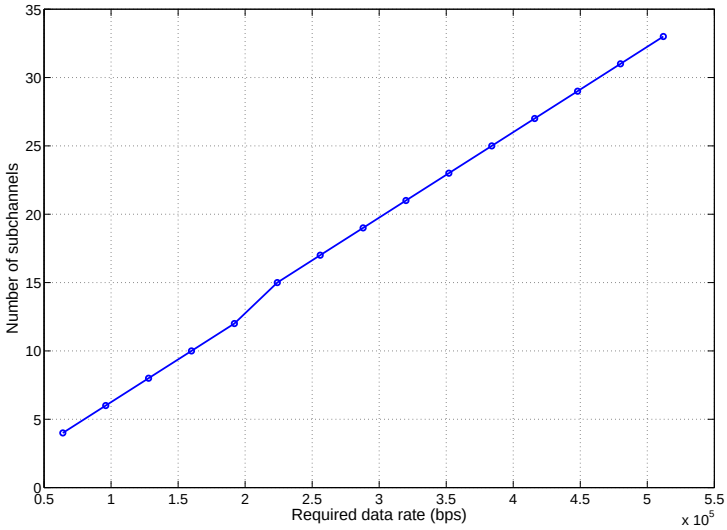

Figure 9.4: Optimal number of subchannels allocated versus the required data rate.

Table 9.3: Modulation scheme for different data rate requirements.

Data rate (kbps)	64	128	192	256
Throughput (bits/subcarrier)	0.28	0.28	0.28	0.26
Number of subchannels	4	8	12	17
Modulation scheme	QPSK 1/2	QPSK 1/2	QPSK 1/2	QPSK 1/2
Data rate (kbps)	320	384	448	512
Throughput (bits/subcarrier)	0.26	0.27	0.27	0.27
Number of subchannels	21	25	29	33
Modulation scheme	QPSK 1/2	QPSK 1/2	QPSK 1/2	QPSK 1/2

9.3 Uplink Scheduling for Multi-user System

9.3.1 Problem Formulation and Solutions

For a multi-user system, the problem of interest is to minimize the total transmit power while satisfying each user's data rate requirement by jointly adjusting power and number of subchannels w_u allocated for each user, which is formulated as follows.

$$\begin{aligned} \min_{\{w_u\},\{P_u\}} \quad & \sum_{1}^{U} P_u \\ s.t. \quad & \sum_{1}^{U} w_u \leq M \\ & R_u = \frac{1}{F_l} \times S_f \times C_a \times w_u \times (1 - O_h) \times f_T \left(\frac{C_o^{(u)} P_u}{w_u} \right), \end{aligned} \tag{9.7}$$

in which P_u is the power allocated to user u, and w_u is the number of subchannels allocated to user u. In addition, the available number of subchannels is limited by M. $C_o^{(u)}$ is the channel coefficient or mean power gain of user u. The total number of users is U.

The above problem can be solved optimally by using DP. The power consumption P_u with a given number of allocated subchannels $1 \leq w_u \leq M + 1 - U$ can be found by solving

$$\frac{1}{F_l} \times S_f \times C_a \times w_u \times (1 - O_h) \times f_T \left(\frac{C_o^{(u)} P_u}{w_u} \right) = R_u. \tag{9.8}$$

As a result, by going through every possible combination of users' allocated subchannels $\{w_u\}$, it is possible to find the joint-optimal $\{P_u^*\}, \{w_u^*\}$ for total power minimization. Remarkably, the required optimization can be achieved by formulating the system as a U-stage dynamic system, with the system state x_u representing

the number of remaining subchannels for users u to U, the control parameter μ_u representing the number of allocated subchannels for user u and the cost function $c_u(\mu_u) = P_u(\mu_u)$ denoting the power consumption for user u. The cost accumulated from stage u to U can be found as

$$J_u(x_u) = \min_{1 \le \mu_u \le x_u} c_u(\mu_u) + J_{u+1}(x_u - \mu_u) \text{ for } 1 \le x_u \le M, \tag{9.9}$$

with $J_{U+1} = 0$. In what follows, the total system cost is given by $J_1(x_1)$, and the optimal control policy $\boldsymbol{\mu}^* = (\mu_1^*, \mu_2^*, \ldots, \mu_U^*)$ is the one that minimizes the cost and can be found from (9.9) by proceeding backward for $u = U, U-1, \ldots, 1$ [83]. Then, the optimal time-slot allocation can be found from

$$\begin{cases} w_u(x_u) = \mu_u^*(x_u), \\ x_{u+1} = M - \sum_{i=1}^{u} w_i, \\ x_1 = M. \end{cases} \tag{9.10}$$

9.3.2 Simulation Results

Simulation results are presented and compared in this section. Table 9.4 lists the system settings used in the simulation. Figs. 9.5, 9.6, 9.7 and 9.8 plot the total transmit power verse the number of subchannels for different settings. And the modulation schemes in accordance with the settings are shown in Table 9.5, 9.6, 9.7 and 9.8.

From the simulation results, it can be observed that the optimal solution by DP always has better performance than the equal subchannel solution, indicating different numbers of subchannels are usually allocated for different users. For lightly loaded systems as shown in Figs. 9.5 and 9.6, the total required transmit power is not always decreasing with the increasing of number of subchannels. The optimal required number of subchannels is usually less than the maximum available

Table 9.4: System settings for the simulation.

Settings	Configurations
Setting 1	$U = 5,\ R = [128, 64, 256, 128, 64]$kbps
	$C_o = [0.1, 0.5, 1, 10, 20]$
Setting 2	$U = 7,\ R = [128, 64, 256, 128, 64, 256, 64]$kbps
	$C_o = [0.01, 0.1, 0.5, 1, 10, 20, 30]$
Setting 3	$U = 10,\ R = [128, 128, 128, 128, 128, 128, 128, 128, 128, 128]$kbps
	$C_o = [0.01, 0.05, 0.1, 0.5, 1, 5, 10, 15, 20, 30]$
Setting 4	$U = 8,\ R = [256, 512, 256, 512, 256, 512, 256, 256]$kbps
	$C_o = [0.01, 0.1, 0.5, 1, 10, 20, 30, 50]$

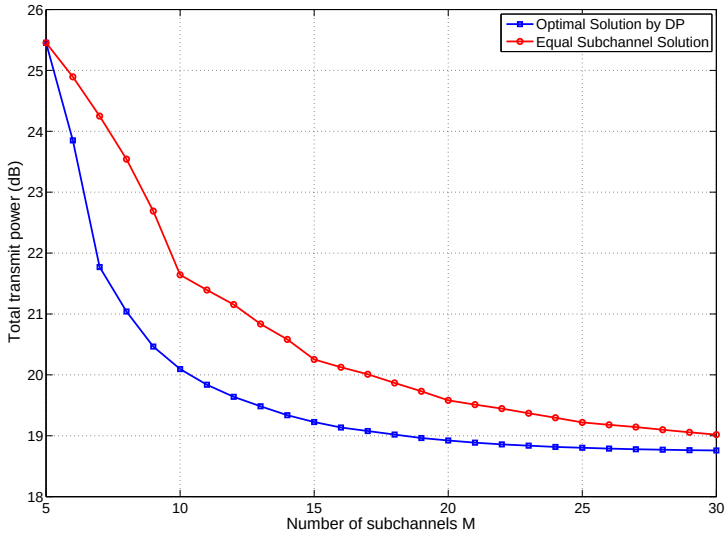

Figure 9.5: Total transmit power versus the number of subchannels for Setting 1.

subchannels, which is 35 for a 10MHz channel. In addition, the users tend to use a lower modulation and coding scheme, such as QPSK1/4, etc.

For heavily loaded systems as shown in Figs. 9.7 and 9.8, the total transmit power usually decreases with the increasing number of subchannels, indicating more subchannels are required in this case. For bad channel conditions with a high data rate requirement, the users are more likely to be allocated with more subchannels and a lower modulation scheme, while for better channel conditions, fewer subchannels are allocated with higher modulation scheme used.

Table 9.5: Modulation scheme for different data rate requirements for Setting 1.

	Subchannels (35)	Modulation Scheme	Throughput (bits/subcarrier)
U1	8	QPSK 1/4	0.28
U2	4	QPSK 1/4	0.28
U3	16	QPSK 1/4	0.28
U4	5	QPSK 1/4	0.44
U5	2	QPSK 1/2	0.56

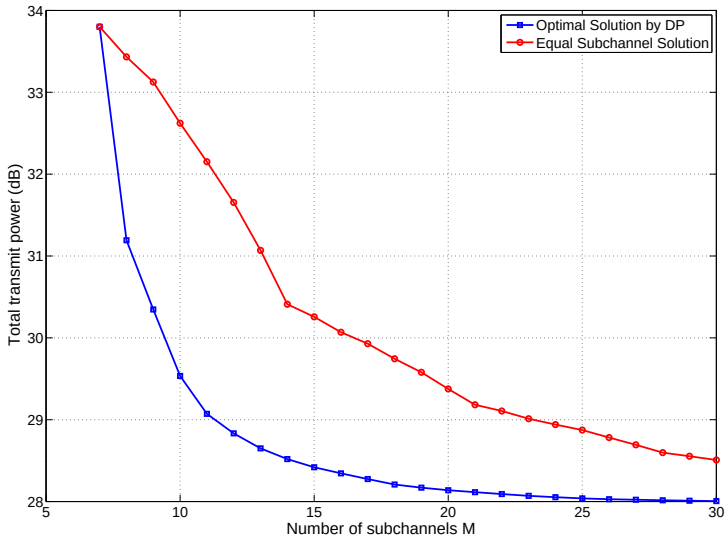

Figure 9.6: Total transmit power versus the number of subchannels for Setting 2.

Table 9.6: Modulation scheme for different data rate requirements for Setting 2.

	Subchannels (35)	Modulation Scheme	Throughput (bits/subcarrier)
U1	8	QPSK 1/4	0.28
U2	4	QPSK 1/4	0.28
U3	12	QPSK 1/4	0.37
U4	5	QPSK 1/4	0.44
U5	1	QPSK 3/4	1.11
U6	4	QPSK 3/4	1.11
U7	1	QPSK 3/4	1.11

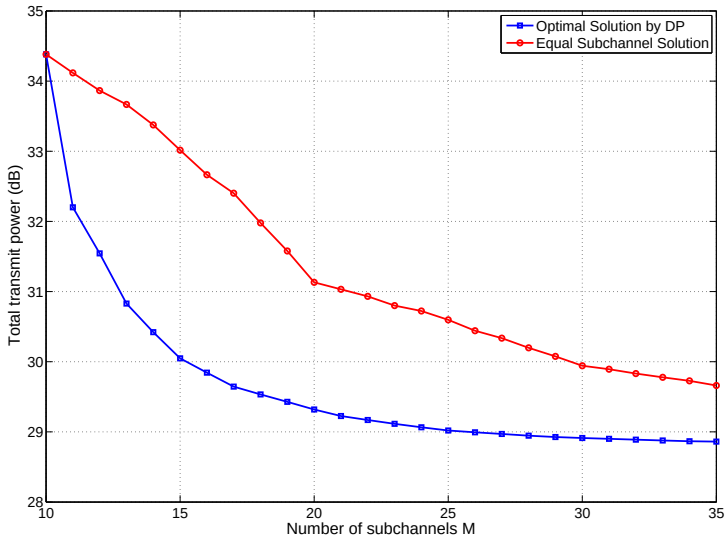

Figure 9.7: Total transmit power versus the number of subchannels for Setting 3.

Table 9.7: Modulation scheme for different data rate requirements for Setting 3.

	Subchannels (35)	Modulation Scheme	Throughput (bits/subcarrier)
U1	8	QPSK 1/4	0.28
U2	7	QPSK 1/4	0.28
U3	6	QPSK 1/4	0.37
U4	4	QPSK 1/2	0.44
U5	3	QPSK 1/2	1.11
U6	2	QPSK 3/4	1.11
U7	2	QPSK 3/4	1.11
U8	1	16QAM 3/4	2.22
U9	1	16QAM 3/4	2.22
U10	1	16QAM 3/4	2.22

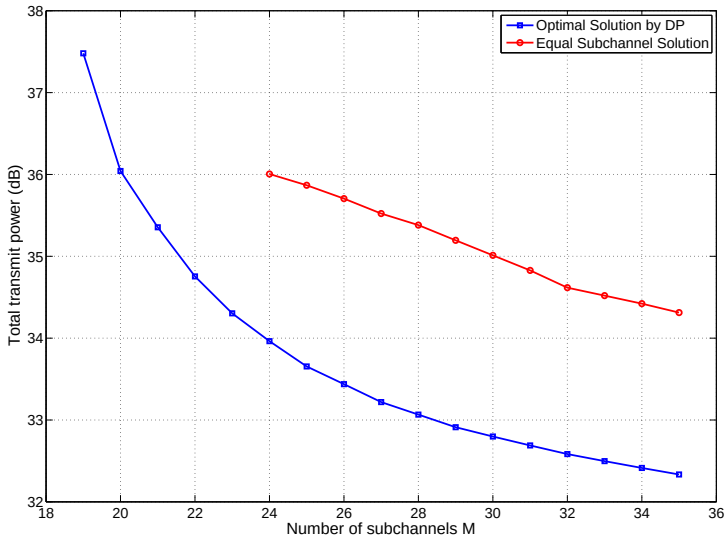

Figure 9.8: Total transmit power versus the number of subchannels for Setting 4.

Table 9.8: Modulation scheme for different data rate requirements for Setting 4.

	Subchannels (35)	Modulation Scheme	Throughput (bits/subcarrier)
U1	10	QPSK 1/4	0.44
U2	9	QPSK 1/2	0.99
U3	3	QPSK 3/4	1.48
U4	4	16QAM 3/4	2.22
U5	2	16QAM 3/4	2.22
U6	3	16QAM 3/4	2.96
U7	2	16QAM 3/4	2.22
U8	2	16QAM 3/4	2.22

9.4 Summary

In this chapter, we have modified the framework developed in Chapters 4-8 and applied it to the IEEE 802.16e WiMAX uplink. The physical layer specifications of WiMAX have been reviewed, where the concept of OFDM, multiple antenna techniques and different modulation schemes have been presented. Based on a BT WiMAX simulation platform, we have reconstructed the uplink scheduling problem for the single user case to minimize the transmit power by adjusting the number of subchannels while satisfying the user's data rate requirement. The problem has been solved numerically and the simulation results show that the transmitter tends to use QPSK 1/2 as modulation scheme and increasing the number of subchannels may require more transmission power. This is different from the theoretical results in which it is always better to allocate more subchannels. This is because the throughput-SNR function is not exact logarithm function in the practical situation as opposed to that in Shannon's theory.

For the multi-user case, the problem has been constructed to minimize the total transmit power while satisfying each user's data rate requirement by optimizing the number of subchannels allocated for each user. We have developed an algorithm to solve the optimization problem by using DP, which has linear complexity with the number of users. Furthermore, we have validated the results under different system settings considering both lightly and heavily loaded traffic. The results have shown a remarkable improvement in power consumption using the proposed algorithm compared with the results without optimizing the number of subchannels. The results in this chapter have proved the significance of the proposed energy efficient design for wireless system resource allocation to satisfy the requirements both in energy reduction and QoS satisfaction.

Chapter 10

Conclusions and Future Work

In Chapter 10 of the thesis, we summarize the research work in the thesis and propose the future research directions.

10.1 Summary of Thesis

In the thesis, we have presented a concrete framework for solving wireless resource allocation problems for delay constrained communications. The uniqueness is to consider delay constraints as statistical QoS requirement in terms of information outage probability under certain decoding delay constraint from the viewpoint of physical layer and queueing delay bound violation probability constraint (or EC constraint) from the viewpoint of data link layer when designing power and time/spectrum allocation mechanisms. Moreover, the consideration of MIMO channels in physical layer is another contribution of the thesis.

Following Chapter 1, which gave the motivation, measurable aims, contributions,

list of publications and outline of the thesis, an overview of SISO and MIMO channel capacity and EC has been presented in Chapter 2. In addition, the mathematical methodologies used in the thesis, including convex optimization theory, nonlinear optimization theory and DP and optimal control theory, have been briefly described in Chapter 3.

Chapter 4 dealt with the power minimization problem in single user transmission systems with SCI at the transmitter to achieve a given information outage probability under certain decoding delay constraint. Gaussian approximation with mean and variance derived as closed-form has been used to represent the p.d.f. of MIMO channel. Transmit power is minimized using a nonlinear optimization method DIRECT algorithm. Numerical results for different MIMO systems (SISO is considered as a special case) with different QoS requirements have been compared and analyzed and the accuracy of Gaussian approximation has been verified.

Chapter 5 has addressed the power minimization problem in a single user transmission system with causal CSI to achieve an information outage probability under certain decoding delay constraint. The problem has been solved for both SISO and MIMO systems optimally using high complexity DP by adapting the power and rate allocation. An efficient near optimal solution at much less complexity has been proposed which is asymptotically optimal for small information outage probability values. A lower bound and a simpler method termed EOPPB have also been developed for comparisons.

Based on single user resource allocation results in Chapters 4 and 5, Chapters 6 and 7 have extended the results to multi-user systems. In Chapter 6, we have solved the power minimization and scheduling problem for a time-division multi-user MIMO system with transmitter SCI under individual user's information outage probability constraint. We have used MPE in Chapter 4 to determine the minimum power for attaining a given outage probability constraint under certain decoding delay constraint by using Gaussian approximation for channel capacity. On the

other hand, we proposed a convex optimization approach to find the suboptimal number of blocks allocated to the users. The two main techniques have been combined to obtain a joint solution for both power and time allocation for the users. Simulation results have demonstrated that the proposed method achieves near-optimal performance.

Chapter 7 has devised an energy efficient joint time-sharing and power allocation algorithm in multi-user TDMA system with individual user's QoS constraint in terms of information outage probability requirement, by exploiting the causal CSI at the transmitter. A simple closed-form OLLP power allocation solution has been developed to meet a given outage probability constraint for a single user channel. The optimization problem has been constructed enabling a joint consideration of power consumption and users' QoS constraints to find a suboptimal multi-user time-sharing solution. This time-sharing solution in conjunction with the OLLP solution has formed the joint solution for the multi-user TDMA system, which has been shown by simulation results to be near-optimal and yield a significant energy saving as compared to those without time-sharing optimization or OLLP power allocation.

While Chapters 4 to 7 have considered delay in terms of information outage probability constraint from the viewpoint of physical layer, Chapter 8 formulated delay constraint in terms of queueing delay bound violation probability from data link layer. In Chapter 8, we have investigated the power minimization problem for multi-user TDMA/FDMA MIMO systems with individual user's QoS constraint in terms of EC function with the aid of SCI at the transmitter side. We have obtained a closed-form expression for the EC of a single user MIMO system, which is then applied to solve the power minimization problem in the multi-user system. By jointly optimizing the power allocation and the number of time-slot (or subcarrier) allocation for all of the users with the aid of SCI at the transmitter, we have shown that a significant power saving is achieved. In particular, we have

demonstrated that the optimal solutions can be obtained by DP and near-optimal heuristics at much reduced complexity are possible.

Finally, Chapter 9 has revised the designed algorithm and applied the results to IEEE 802.16e WiMAX for uplink scheduling. Energy efficient uplink scheduling algorithm has been designed to satisfy users' QoS requirement in terms of data rate. The simulations have been carried out under different system settings, including lightly loaded traffic and heavily loaded traffic. The significance of the proposed energy efficient design for wireless system resource allocation has been verified in terms of power reduction and users' QoS satisfaction.

10.2 Future Work

As a result of the work undertaken so far presented in the thesis, several directions for future research have been proposed here to extend the obtained results.

- In the thesis, we have designed wireless resource allocation strategies considering i.i.d. Rayleigh fading channel. This has provided a fundamental framework which can be further applied to different physical channel models. Several interesting directions the research can be further investigated are listed as follows.

  - Firstly, the framework can be further extended to MIMO spatial correlated channel model, where channel correlation matrix is added in at the transmitter and receiver sides, or time correlated channels, where the channel fading among blocks are correlated. However, for spatial or time correlated channels, with SCI at the transmitter, the Gaussian approximation used in Chapter 4 may not be accurate to measure the distribution of the "sum-rate" [110].

– Secondly, delayed feedback is another possible research direction based on the framework in the thesis. When the feedback is delayed, instantaneous CSI model is not accurate. The reason is because in block fading channel model, we assume the block fading is i.i.d. Since there is no time correlation among different blocks, the delayed feedback cannot be used to estimate the channel state as the channels may change. However, delayed feedback has smaller or no impact for SCI feedback. In this case, the SCI model can still be applied for delayed feedback. This has pointed out one of the possible future research directions to consider delayed feedback situation.

– Thirdly, the channel fading model of Rayleigh fading can be generalized to other fading channel models, such as Rician fading, where there exists a dominant communication path between the transmitter and receiver. With SCI at the transmitter, the distribution of the "sum-rate" can still be accurately approximated by Gaussian distribution in Rician fading channels [94]. However, for other types of channel fading, such as Nakagami channel fading, as far as our knowledge, no research work has investigated the mutual information distribution. According to central limit theory, Gaussian approximation can still be accurate for large number of blocks K. However, when K is small, the distribution may not be accurately approximated by Gaussian approximation.

– Fourthly, when the channel is very noisy, the feedback information from the receiver to the transmitter is not accurate. Since in the thesis, only accurate feedback with no error is considered, the model used in the thesis is not able to capture such situations and the desired QoS performance may not be achieved (resulting in QoS degradation) when the feedback is inaccurate. This has pointed out one of the possible future research directions to investigate the model and resource allocation strategies taking the feedback error into consideration.

- The thesis developed the joint power allocation and time-sharing scheduling method for multi-user resource allocation. Based on the built framework, the scheduling method can be further combined with other scheduling methods. For example, the framework can be combined with scheduling using multi-user diversity when CSI is available at the transmitter side, where transmitter can opportunistically choose the user for data transmission based on the CSI feedback [22, 23, 24, 26, 27, 25]. Another example is to combine the scheduling with MIMO spatial division multiple access technique, where within a time slot a best set of users is allocated resource instead of a single user [48].

- The thesis has dealt with the wireless resource allocation problem within a single cell. Due to the heavy demand for expensive and precious spectrum and aggressive reuse of the spectral resources to increase network capacity, it is inevitable to cause an increased level of interference throughout the network. This brings about an interesting research direction to investigate resource allocation for improving the overall efficiency of the system taking into consideration the interference among different cells through the use of multiple base station cooperation. Based on this, the research can be further investigated in the following directions. Firstly, the research work can be applied for developing multi-cell base station cooperation and resource allocation strategies. The energy efficient scheduling can be applied for multicell resource allocation of a single mobile operator, where the power, time slots and spectrum among different cells can be optimized. Secondly, the research can be applied to developing inter operator base station cooperation and resource allocation strategies. The energy efficient scheduling algorithms can be applied for multiple operators' base station cooperation and resource sharing, which can further improve users' QoS and reduce the transmit energy.

Appendices

A.1 Evaluation of $\int_0^\infty \ln(1+a\lambda)\lambda^j e^{-\lambda} d\lambda$

To evaluate the expectation, we first reformulate it as

$$\begin{aligned} &\int_0^\infty \lambda^j e^{-\lambda} \ln(1+a\lambda) d\lambda \\ =&\frac{e^{\frac{1}{a}}}{a^{j+1}} \int_1^\infty e^{-\frac{1}{a}t} (t-1)^j \ln t dt \\ =&\frac{e^{\frac{1}{a}}}{a^{j+1}} \sum_{n=0}^{j} (-1)^n \binom{j}{n} \int_1^\infty e^{-\frac{1}{a}t} t^{j-n} \ln t dt \end{aligned} \tag{1}$$

where $t = 1 + a\lambda$ and the integrand of the last line can be derived as follows:

$$
\begin{aligned}
&\int_1^\infty e^{-\frac{1}{a}t} t^{j-n} \ln t dt \\
=&a \int_1^\infty e^{-\frac{1}{a}t} \left[t^{j-n-1} + t^{j-n-1}(j-n) \ln t \right] dt \\
=&a \left(a \left\{ e^{-\frac{1}{a}} + \int_1^\infty e^{-\frac{1}{a}t} \left[(j-n-1)t^{j-n-2} + (j-n)t^{j-n-2} + (j-n-1)(j-n)t^{j-n-2} \ln t \right] \right\} \right) \\
&\vdots \\
=&(a)^2 e^{-\frac{1}{a}} + (a)^3 \left[(j-n-1) + (j-n) \right] \\
&+ (a)^4 \left[(j-n)(j-n-2) + (j-n-1)(j-n-2) + (j-n-1)(j-n) \right] + \cdots \\
&+ \frac{(j-n)!}{\left(\frac{1}{a}\right)^{j-n}} \left(\int_1^\infty e^{-\frac{1}{a}t} \ln t dt + \sum_{p=1}^{j-n} \frac{1}{j-n+1-p} \int_0^\infty e^{-\frac{1}{a}t} dt \right) \\
=&e^{-\frac{1}{a}} \sum_{p=2}^{j-n} (a)^p \frac{(j-n)!}{(j-n-p+1)!} \sum_{q=1}^{p-1} \frac{1}{j-n-q+1} \\
&+ \frac{(j-n)!}{\left(\frac{1}{a}\right)^{j-n}} \left(\int_1^\infty e^{-\frac{1}{a}} \ln t dt + a e^{-\frac{1}{a}} \sum_{p=1}^{j-n} \frac{1}{j-n+1-p} \right).
\end{aligned}
\tag{2}
$$

Substituting this result into (1), we have

$$
\begin{aligned}
\int_0^\infty \lambda^j e^{-\lambda} \ln(1+a\lambda) d\lambda = \frac{e^{\frac{1}{a}}}{a^j} \sum_{n=0}^{j} (-1)^n \binom{j}{n} \frac{(j-n)!}{\left(\frac{1}{n}\right)^{j-n}} \int_1^\infty e^{-\frac{1}{a}t} \ln t dt \\
+ \frac{1}{a^j} \sum_{n=0}^{j-1} \sum_{p=1}^{j-n} (-1)^n \binom{j}{n} \frac{(j-n)!}{\left(\frac{1}{n}\right)^{j-n}} \cdot \frac{1}{j-n+1-p} \\
+ \frac{1}{a^{j+1}} \sum_{n=0}^{j-2} \sum_{p=2}^{j-n} \sum_{q=1}^{p-1} (-1)^n \binom{j}{n} \frac{(j-n)!}{(j-n-p+1)!} (a)^p \frac{1}{j-n-q+1}.
\end{aligned}
\tag{3}
$$

From (4.331-2) [95], we can obtain

$$
\int_1^\infty e^{-\frac{1}{a}t} \ln t dt = a E_1 \left(\frac{1}{a} \right) \tag{4}
$$

where $E_1(\cdot)$ denotes the exponential integral.

A.2 Evaluation of σ_ρ^2

Similarly, we start by considering $\int_0^\infty [\ln(1+a\lambda)]^2 \lambda^j e^{-\lambda} d\lambda$ which equals

$$\frac{e^{\frac{i}{a}}}{a^{j+1}} \int_1^\infty e^{-\frac{1}{a}t}(t-1)^j (\ln t)^2 dt = \frac{e^{\frac{1}{a}}}{a^{j+1}} \sum_{n=0}^{j} (-1)^n \binom{j}{n} \int_1^\infty e^{-\frac{1}{a}t} (\ln t)^2 t^{j-n} dt. \quad (5)$$

Then, the challenge is to derive the remaining integrand which has been done as follows:

$$\begin{aligned}
&\int_1^\infty e^{-\frac{1}{a}t} (\ln t)^2 t^{j-n} dt \\
=&a \int_1^\infty e^{-\frac{1}{a}t} \left[2t^{j-n-1} \ln t + t^{j-n-1}(j-n)(\ln t)^2\right] dt \\
=&(a)^2 \int_1^\infty e^{-\frac{1}{a}t} \left[2t^{j-n-2} + 2t^{j-n-2}(j-n-1)\ln t + 2t^{j-n-2}(j-n)\ln t\right] dt \\
&+ (a)^2 \int_1^\infty e^{-\frac{1}{a}t} t^{j-n-2}(j-n-1)(j-n)(\ln t)^2 dt \\
&\vdots \\
=&2e^{-\frac{1}{a}} \sum_{t=3}^{j-n} (a)^t \frac{(j-n)!}{(j-n-t+1)!} \sum_{p=1}^{t-2} \sum_{q=p+1}^{t-1} \frac{1}{(j-n+1-p)(j-n+1-q)} \\
&+ \frac{(j-n)!}{\left(\frac{1}{a}\right)^{j-n}} \left(\int_1^\infty e^{-\frac{1}{a}t} (\ln t)^2 dt + 2 \sum_{p=1}^{j-n} \frac{1}{j-n+1-p} \int_1^\infty e^{-\frac{1}{a}t} \ln t dt \right) \\
&+ \frac{(j-n)!}{\left(\frac{1}{a}\right)^{j-n}} \int_1^\infty 2e^{-\frac{1}{a}t} \sum_{p=1}^{j-n-1} \sum_{q=p+1}^{j-n} \frac{1}{(j-n+1-p)(j-n+1-q)} dt.
\end{aligned} \quad (6)$$

Using (6) into (5), we get

$$\int_0^\infty [\ln(1+a\lambda)]^2 \lambda^j e^{-\lambda} d\lambda = \frac{e^{\frac{1}{a}}}{a^{j+1}} \sum_{n=0}^{j} (-1)^n \binom{j}{n} \frac{(j-n)!}{\left(\frac{1}{a}\right)^{j-n}} \int_1^\infty e^{-\frac{1}{a}t} (\ln t)^2 dt$$
$$+ \frac{2e^{\frac{1}{a}}}{a^{j+1}} \int_1^\infty e^{-\frac{1}{a}t} \ln t dt \sum_{n=0}^{j-1} \sum_{p=1}^{j-n} (-1)^n \binom{j}{n} \frac{(j-n)!}{\left(\frac{1}{a}\right)^{j-n}} \cdot \frac{1}{j-n+1-p}$$
$$+ \frac{2}{a^j} \sum_{n=0}^{j-2} \sum_{p=1}^{j-n-1} \sum_{q=p+1}^{j-n} (-1)^n \binom{j}{n} \frac{(j-n)!}{\left(\frac{1}{a}\right)^{j-n}} \cdot \frac{1}{(j-n+1-p)(j-n+1-q)}$$
$$+ \frac{2}{a^{j+1}} \sum_{n=0}^{j-3} \sum_{t=3}^{j-n} \sum_{p=1}^{t-2} \sum_{q=p+1}^{t-1} (-1)^n \binom{j}{n} (a)^t \frac{(j-n)!}{(j-n-t+1)!} \cdot \frac{1}{(j-n+1-p)(j-n+1-q)}. \tag{7}$$

in which the result of $\int_1^\infty e^{-\frac{1}{a}t} \ln t dt$ is shown in 4 and from (4.335) [95], we have

$$\int_1^\infty e^{-\frac{1}{a}t} (\ln t)^2 dt = a \left[\left(\ln \frac{i}{a} - \gamma_{\mathsf{EM}} \right)^2 + \frac{\pi^2}{6} \right] - 2_3F_3 \left([1,1,1]; [2,2,2]; -\frac{1}{a} \right), \tag{8}$$

where γ_{EM} is Euler constant and ${}_pF_q$ denotes the generalized hypergeometric function.

A.3 Derivation of $\tilde{\rho} = \mathtt{E}[\ln \lambda]$ and $\tilde{\sigma}^2 = \mathtt{VAR}[\ln \lambda]$

Given a random variable λ with p.d.f. in (6.19), $\tilde{\rho}$ can be found by

$$\begin{aligned}
\tilde{\rho} &= \mathtt{E}[\ln \lambda] \\
&= \sum_{m=1}^{n_r} \sum_{\ell=\delta}^{(n_t+n_r)m-2m^2} \left(\frac{d_{m,\ell} \times m^{\ell+1}}{\ell!} \right) \times \int_0^\infty \lambda^\ell e^{-m\lambda} \ln \lambda d\lambda \\
&= \sum_{m=1}^{n_r} \sum_{\ell=\delta}^{(n_t+n_r)m-2m^2} \left(\frac{d_{m,\ell} \times m^{\ell+1}}{\ell!} \right) \times \left(\frac{1}{m} \right)^{\ell+1} \int_0^\infty t^\ell e^{-t} \ln \left(\frac{t}{m} \right) dt \\
&= \sum_{m=1}^{n_r} \sum_{\ell=\delta}^{(n_t+n_r)m-2m^2} \left(\frac{d_{m,\ell} \times m^{\ell+1}}{\ell!} \right) \times \left(\frac{1}{m} \right)^{\ell+1} \left[\int_0^\infty t^\ell e^{-t} \ln t dt - (\ln m) \int_0^\infty t^\ell e^{-t} dt \right] \\
&\equiv \sum_{m=1}^{n_r} \sum_{\ell=\delta}^{(n_t+n_r)m-2m^2} \left(\frac{d_{m,\ell} \times m^{\ell+1}}{\ell!} \right) \times \left(\frac{1}{m} \right)^{\ell+1} \left[S(\ell) - (\ln m)\ell! \right],
\end{aligned} \tag{9}$$

where $S(\ell) \triangleq \int_0^\infty t^\ell e^{-t} \ln t dt$, and we have used the fact that

$$\int_0^\infty t^\ell e^{-t} dt = \Gamma(\ell + 1) = \ell!. \tag{10}$$

In (4.352) from [95], we have $S(\ell) = \ell!(H_\ell - \gamma_{\mathsf{EM}})$ in which H_ℓ denotes the Harmonic number. As a result, we have

$$\tilde{\rho} = \sum_{m=1}^{n_r} \sum_{\ell=\delta}^{(n_t+n_r)m-2m^2} d_{m,\ell} \left(H_\ell - \gamma_{\mathsf{EM}} - \ln m \right). \tag{11}$$

The variance $\tilde{\sigma}^2$ can be derived as follows. Firstly, $\mathtt{E}[(\ln\lambda)^2]$ can be evaluated as

$$\begin{aligned}
\mathtt{E}[(\ln\lambda)^2] &= \sum_{m=1}^{n_r}\sum_{\ell=\delta}^{(n_t+n_r)m-2m^2}\left(\frac{d_{m,\ell}\times m^{\ell+1}}{\ell!}\right)\times\int_0^\infty \lambda^\ell e^{-m\lambda}(\ln\lambda)^2 d\lambda \\
&= \sum_{m=1}^{n_r}\sum_{\ell=\delta}^{(n_t+n_r)m-2m^2}\left(\frac{d_{m,\ell}\times m^{\ell+1}}{\ell!}\right)\times\int_0^\infty \left(\frac{t}{m}\right)^\ell e^{-t}(\ln t-\ln m)^2\frac{dt}{m} \\
&= \sum_{m=1}^{n_r}\sum_{\ell=\delta}^{(n_t+n_r)m-2m^2}\left(\frac{d_{m,\ell}}{\ell!}\right)\times\int_0^\infty t^\ell e^{-t}\left[(\ln t)^2-2(\ln t)(\ln m)+(\ln m)^2\right]dt \\
&= \sum_{m=1}^{n_r}\sum_{\ell=\delta}^{(n_t+n_r)m-2m^2}\left(\frac{d_{m,\ell}}{\ell!}\right)\times\left[W(\ell)-2(\ln m)S(\ell)+(\ln m)^2\ell!\right]
\end{aligned} \tag{12}$$

where $W(\ell)\triangleq\int_0^\infty t^\ell e^{-t}(\ln t)^2dt$. In Appendix 10.2, we have shown that

$$W(\ell)=\ell!\left(\gamma_{\mathsf{EM}}^2-2\gamma_{\mathsf{EM}}+\frac{\pi^2}{6}\right)+2\ell!\sum_{j=1}^{\ell-1}\frac{H_j-\gamma_{\mathsf{EM}}}{j+1}. \tag{13}$$

Using this result together with that for $S(\ell)$, we can express $\mathtt{E}[(\ln\lambda)^2]$ in closed-form as

$$\begin{aligned}
&\mathtt{E}[(\ln\lambda)^2] \\
&=\sum_{m=1}^{n_r}\sum_{\ell=\delta}^{(n_t+n_r)m-2m^2} d_{m,\ell}\left[\gamma_{\mathsf{EM}}^2+2(\ln m-H_\ell)\gamma_{\mathsf{EM}}+\frac{\pi^2}{6}-2(\ln m)H_\ell+(\ln m)^2+2\sum_{j=1}^{\ell-1}\frac{H_j}{j+1}\right].
\end{aligned} \tag{14}$$

As such, the variance is given by

$$\begin{aligned}
\tilde{\sigma}^2 &= \mathtt{E}[(\ln\lambda)^2]-(\mathtt{E}[\ln\lambda])^2 \\
&=\sum_{m=1}^{n_r}\sum_{\ell=\delta}^{(n_t+n_r)m-2m^2} d_{m,\ell}\left[\gamma_{\mathsf{EM}}^2+2(\ln m-H_\ell)\gamma_{\mathsf{EM}}+\frac{\pi^2}{6}-2(\ln m)H_\ell+(\ln m)^2+2\sum_{j=1}^{\ell-1}\frac{H_j}{j+1}\right] \\
&\quad-\left[\sum_{m=1}^{n_r}\sum_{\ell=\delta}^{(n_t+n_r)m-2m^2} d_{m,\ell}\left(H_\ell-\gamma_{\mathsf{EM}}-\ln m\right)\right]^2.
\end{aligned} \tag{15}$$

A.3.1 Evaluation of $W(\ell) = \int_0^\infty t^\ell e^{-t}(\ln t)^2 dt$

Similar technique can be applied to derive $W(\ell)$, and the first is to have the recursive relation:

$$\begin{aligned} W(\ell) &= \int_0^\infty t^\ell e^{-t}(\ln t)^2 dt \\ &= -\int_0^\infty t^\ell (\ln t)^2 de^{-t} \\ &= 2\int_0^\infty t^{\ell-1} e^{-t} \ln t dt + \ell \int_0^\infty t^{\ell-1} e^{-t} (\ln t)^2 dt \\ &= 2S(\ell-1) + \ell W(\ell-1). \end{aligned} \tag{16}$$

Applying this further, we can get

$$\begin{aligned} W(\ell) &= 2S(\ell-1) + \ell W(\ell-1) \\ &= 2S(\ell-1) + 2\ell S(\ell-2) + \ell(\ell-1)W(\ell-2) \\ &= 2S(\ell-1) + 2\ell S(\ell-2) + 2\ell(\ell-1)S(\ell-3) + \ell(\ell-1)(\ell-2)W(\ell-3) \\ &\vdots \\ &= 2S(\ell-1) + 2\ell S(\ell-2) + 2\ell(\ell-1)S(\ell-3) + \cdots \\ &+ 2\ell(\ell-1)(\ell-2)\cdots(2)S(0) + \ell(\ell-1)(\ell-2)\cdots(2)(1)W(0) \\ &= \ell! W(0) + 2\left(\frac{\ell!}{1!}\right) S(0) + 2\left(\frac{\ell!}{2!}\right) S(1) + 2\left(\frac{\ell!}{3!}\right) S(2) + \cdots \\ &+ 2\left(\frac{\ell!}{(\ell-1)!}\right) S(\ell-2) + 2\left(\frac{\ell!}{\ell!}\right) S(\ell-1) \\ &= \ell! W(0) + 2\ell! \sum_{j=0}^{\ell-1} \frac{S(j)}{(j+1)!}. \end{aligned} \tag{17}$$

Again, note that

$$W(0) = \int_0^\infty e^{-t}(\ln t)^2 dt = \gamma_{\mathsf{EM}}^2 + \frac{\pi^2}{6} \tag{18}$$

and we can find $W(\ell)$ as

$$W(\ell) = \ell! \left(\gamma_{\mathsf{EM}}^2 - 2\gamma_{\mathsf{EM}} + \frac{\pi^2}{6}\right) + 2\ell! \sum_{j=1}^{\ell-1} \frac{H_j - \gamma_{\mathsf{EM}}}{j+1}. \tag{19}$$

References

[1] V. McDonald, "The cellular concept," *Bell System Technical Journal*, 1979.

[2] A. Goldsmith, *Wireless communications*. Cambridge University Press, 2005.

[3] D. Tse and P. Viswanath, *Fundamentals of wireless communication*. Cambridge University Press, 2005.

[4] M. Zeng, A. Annamalai, and V. Bhargava, "Recent advances in cellular wireless communications," *IEEE Communications Magazine*, vol. 37, no. 9, pp. 128–138, 1999.

[5] F. Gfeller and U. Bapst, "Wireless in-house data communication via diffuse infrared radiation," *Proceedings of IEEE*, vol. 67, no. 11, pp. 1474–1486, 1979.

[6] Qualcomm, *Evolution of wireless applications and services*, 2007.

[7] J. Schiller, *Mobile communications*. Addison-wesley, second ed., 2003.

[8] A. Han and K. Liu, *Resource allocation for wireless networks*. Cambridge, first ed., 2008.

[9] Z. Han, *An optimization theoretical framework for resource allocation over wireless networks*. PhD thesis, University of Maryland, USA, 2003.

[10] S. Grandhi, R. Yates, and D. Goodman, "Resource allocation for cellular radio systems," *IEEE Transactions on Vehicular Technology*, vol. 46, no. 3, pp. 571–587, 1997.

[11] J. Zander, "Performance of optimal transmit power control in cellular radio systems," *IEEE Transactions on Vehicular Technology*, vol. 41, no. 1, pp. 57–62, 1992.

[12] R. Yates, "A framework for uplink power control in cellular radio systems," *IEEE Journal on Selected Areas in Communications*, vol. 13, no. 7, pp. 1341–1347, 1995.

[13] S. Grandhi, R. Vijayan, D. Goodman, D. Goodman, and J. Zander, "Centralized power control in cellular radio systems," *IEEE Transactions on Vehicular Technology*, vol. 42, no. 4, pp. 466–468, 1993.

[14] S. Grandhi, R. Vijayan, and D. Goodman, "Distributed power control in cellular radio systems," *IEEE Transactions on Vehicular Technology*, vol. 42, no. 4, pp. 226–228, 1994.

[15] S. Das, S. Ganu, N. Rivera, and R. Roy, "Performance Analysis of Downlink Power Control in CDMA Systems," in *IEEE Sarnoff Student Symposium*, (USA), 2004.

[16] B. Jabbari, "Teletraffic aspects of evolving and next-generation wireless communication networks," *IEEE Personal Communications Magazine*, vol. 3, no. 6, pp. 4–9, 1996.

[17] A. Goldsmith and P. Varaiya, "Capacity of fading channels with channel side information," *IEEE Transactions on Information Theory*, vol. 43, no. 6, pp. 1986–1992, 1997.

[18] M. Alouini and A. Goldsmith, "Capacity of Rayleigh fading channels under different adaptive transmission and diversity-combining techniques," *IEEE Transactions on Vehicular Technology*, vol. 48, no. 4, pp. 1165–1181, 1999.

[19] H. Fattah and C. Leung, "An overview of scheduling algorithms in wireless multimedia networks," *IEEE Transactions on Wireless Communication*, vol. 9, pp. 76–83, Oct. 2002.

[20] W. Ajib and D. Haccoun, "An overview of scheduling algorithms in MIMO-based fourth-generation wireless systems," *IEEE Transactions on Network*, vol. 19, pp. 43–48, Sept. 2005.

[21] M. Andrews, K. Kumaran, K. Ramanan, A. Stolyar, P. Whiting, and R. Vijayakumar, "Providing quality of service over a shared wireless link," *IEEE Communications Magazine*, vol. 39, no. 2, pp. 150–154, 2001.

[22] R. Knopp and P. Humblet, "Information capacity and power control in single-cell multiuser communications," in *IEEE International Conference on Communications*, vol. 1, pp. 18–22, 1995.

[23] D. Tse, "Optimal power allocation over parallel Gaussian broadcast channels," in *IEEE International Symposium on Information Theory*, p. 27, 1997.

[24] D. Gesbert and M. Alouini, "How much feedback is multi-user diversity really worth?," in *IEEE International Conference on Communications*, vol. 1, pp. 234–238, 2004.

[25] P. Viswanath, D. Tse, and R. Laroia, "Opportunistic beamforming using dumb antennas," *IEEE Transactions on Information Theory*, vol. 48, no. 6, pp. 1277–1294, 2002.

[26] V. Hassel, *Design issues and performance analysis for opportunistic scheduling algorithms in wireless networks*. PhD thesis, Department of Electronics

and Telecommunications, Norwegian University of Science and Technology, Norway, Jan. 2007.

[27] V. Hassel, G. Oien, and D. Gesbert, "Throughput guarantees for wireless networks with opportunistic scheduling: a comparative study," *IEEE Transactions on Wireless Communications*, vol. 6, pp. 4215–4220, Dec. 2007.

[28] L. Yang and M. Alouini, "Performance analysis of multiuser selection diversity," in *IEEE International Conference on Communications*, (France), 2004.

[29] S. Kulkarni and C. Rosenberg, "Opportunistic scheduling policies for wireless systems with short term fairness constraints," in *IEEE Global Communications Conference*, vol. 1, pp. 533–537, Dec. 2003.

[30] V. Hassel, M. Hanssen, and G. Oien, "Spectral efficiency and fairness for opportunistic round robin scheduling," in *IEEE International Conference on Communications*, vol. 2, pp. 784–789, 2006.

[31] X. Qin and R. Berry, "Exploiting multiuser diversity for medium access control in wireless networks," in *IEEE Conference on Computer Communications*, vol. 2, (USA), pp. 1084–1094, 2003.

[32] G. Foschini and M. Gans, "On limits of wireless communication in a fading environment when using multiple antennas," *Wireless Personal Communications*, vol. 6, no. 3, pp. 311–335, 1998.

[33] E. Telatar, "Capacity of multi-antenna Gaussian channels," *European Transactions on Telecommunications*, vol. 10, pp. 585–595, Nov. 1999.

[34] Q. Spencer, *Transmission strategies for wireless multi-user, multiple-input, multiple-output communication.* PhD thesis, Brigham Young University, 2004.

[35] A. Paulraj, D. Gore, R. Nabar, and H. Bolcskei, "An overview of MIMO communications-a key to gigabit wireless," *IEEE Proceedings*, vol. 92, pp. 198–218, Feb. 2004.

[36] IEEE 802.16 working group, *Air interface for fixed broadband wireless access systems*.

[37] IEEE 802.16 working group, *Air interface for fixed and mobile broadband wireless access systems*.

[38] 3rd Generation Partnership Project, *Overview of 3GPP release 8 v0.0.6*.

[39] S. Jayaweera and H. Poor, "Capacity of multiple-antenna systems with both receiver and transmitter state information," *IEEE Transactions on Information Theory*, vol. 49, pp. 2697–2709, Oct. 2003.

[40] L. Hanlen and A. Grant, "Optimal transmit covariance for MIMO channels with statistical transmitter side information," in *IEEE International Symposium on Information Theory*, Sept. 2005.

[41] A. Grant, "Capacity of ergodic MIMO channels with complete transmitter channel knowledge," in *Communication Theory Workshop*, Feb. 2005.

[42] N. Sharma and L. Ozarow, "A study of opportunism for multiple-antenna systems," *IEEE Transactions on Information Theory*, vol. 51, pp. 1804–1814, May 2005.

[43] I. Kim, S. Hong, S. Chassemzadeh, and V. Tarokh, "Opportunistic beamforming based on multiple weighting vectors," *IEEE Transactions on Wireless Communications*, vol. 4, no. 6, pp. 2683–2687, 2005.

[44] M. Kim, Z. Yi, D. Kim, and W. Chung, "Improved opportunistic beamforming in Ricean channels," *IEEE Transactions on Communications*, vol. 54, no. 12, pp. 2199–2211, 2006.

[45] R. Laroia, J. Li, S. Rangan, and M. Srinivasan, "Enhanced opportunistic beamforming," in *IEEE Vehicular Technology Conference*, vol. 3, (USA), pp. 1762–1766, 2003.

[46] J. Chung, C. Hwang, K. Kim, and Y. Kim, "A random beamforming technique in MIMO systems exploiting multiuser diversity," *IEEE Journal on Selected Areas in Communications*, vol. 21, no. 6, pp. 848–855, 2003.

[47] R. Zhang, "Scheduling for maximum capacity in SDMA/TDMA networks," in *IEEE International Conference on Acoustics, Speech, and Signal Processing*, vol. 3, (USA), pp. 2141–2144, May 2002.

[48] H. Yin and H. Liu, "Performance of space-division multiple access (SDMA) with scheduling," *IEEE Transactions on Wireless Communication*, vol. 1, pp. 611–618, Oct. 2003.

[49] S. Serbetli and A. Yener, "Time slotted multiuser MIMO Systems: beamforming and scheduling strategies," *EURASIP Journal on Wireless Communications and Networking, Special Issue on Multiuser MIMO Networks*, 2004.

[50] 3rd Generation Partnership Project, *Quality of Service concept and architecture*, r23.107 ed., 2008.

[51] WiMAX Forum, *Mobile WiMAX-Part I: A technical overview and performance evaluation*, 2006.

[52] E. Biyikoglu, A. Gamal, and B. Prabhakar, "Adaptive transmission of variable-rate data over a fading channel for energy-efficiency," in *IEEE Global Telecommunications Conference*, vol. 1, (Taiwan), pp. 97–101, 2002.

[53] E. Biyikoglu, B. Prabhakar, and A. Gamal, "Energy-efficient packet transmission over a wireless link," *IEEE/ACM Transactions on Networking*, vol. 10, no. 4, pp. 487–499, 2002.

[54] B. Prabhakar, E. Biyikoglu, and A. Gamal, "Energy-efficient transmission over a wireless link via lazy packet scheduling," in *IEEE Conference on Computer Communications*, vol. 1, pp. 386–394, 2001.

[55] A. Gamal, C. Nair, B. Prabhakar, E. Biyikoglu, and S. Zahedi, "Energy-efficient scheduling of packet transmissions over wireless networks," in *IEEE Conference on Computer Communications*, vol. 3, (USA), pp. 1773–1782, 2002.

[56] Y. Yao and G. Giannakis, "Energy-efficient scheduling for wireless sensor networks," *IEEE Transactions on Communications*, vol. 53, pp. 1333–1342, Aug. 2005.

[57] P. Nuggehalli, V. Srinivasan, and R. Rao, "Delay constrained energy efficient transmission strategies for wireless devices," in *IEEE Conference on Computer Communications*, vol. 3, (USA), pp. 1765–1772, 2002.

[58] A. Fu, E. Modiano, and J. Tsitsiklis, "Optimal energy allocation for delay-constrained data transmission over a time-varying channel," in *IEEE Conference on Computer Communications*, vol. 2, pp. 1095–1105, 2003.

[59] E. Biyikoglu and A. Gamal, "On adaptive transmission for energy efficiency in wireless data networks," *IEEE Transactions on Information Theory*, vol. 50, pp. 3081–3094, Dec. 2004.

[60] M. Khojastepour and A. Sabharwal, "Delay-constrained scheduling: power efficiency, filter design, and bounds," in *IEEE Conference on Computer Communications*, vol. 3, (Hong Kong), pp. 1938–1949, 2004.

[61] B. Ata, "Dynamic power control in a wireless static channel subject to a Quality-of-Service constraint," *Operations Research*, vol. 53, no. 5, pp. 842–851, 2005.

[62] M. Zafer and E. Modiano, "A calculus approach to minimum energy transmission policies with quality of service guarantees," in *IEEE Conference on Computer Communications*, vol. 1, pp. 548–559, Mar. 2005.

[63] M. Zafer and E. Modiano, "Delay-constrained energy efficient data transmission over a wireless fading channel," in *IEEE Information Theory and Applications Workshop*, (USA), Jan. 2007.

[64] G. Caire, G. Taricco, and E. Biglieri, "Optimum power control over fading channels," *IEEE Transactions on Information Theory*, vol. 45, no. 9, pp. 1468–1489, 1999.

[65] E. Biglieri, G. Caire, and G. Taricco, "Limiting performance of block-fading channels with multiple antennas," *IEEE Transactions on Information Theory*, vol. 47, pp. 1468–1489, May 2001.

[66] Y. Liang, R. Zhang, and J. Cioffi, "Subchannel grouping and statistical waterfilling for vector block-fading channels," *IEEE Transactions on Communications*, vol. 36, no. 6, pp. 1131–1141, 2006.

[67] Y. Liang, R. Zhang, and J. Cioffi, "Transmit optimization for MIMO-OFDM with delay-constrained and no-delay-constrained traffic," *IEEE Transactions on Signal Processing*, vol. 54, no. 8, pp. 3190–3199, 2006.

[68] B. Collins and R. Cruz, "Transmission policies for time varying channels with average delay constraints," in *Allerton Conference on Communication, Control and Computing*, vol. 1, (Monticello), pp. 282–286, Sept. 1999.

[69] R. Berry and R. Gallager, "Communication over fading channels with delay constraints," *IEEE Transactions on Information Theory*, vol. 48, pp. 1135–1149, May 2002.

[70] H. Wang, "A simple packet-transmission scheme for wireless data over fading channels," *IEEE Transactions on Communications*, vol. 52, no. 7, pp. 1055–1059, 2004.

[71] D. Rajan, A. Sabharwal, and B. Aazhang, "Delay-bounded packet scheduling of bursty traffic over wireless channels," *IEEE Transactions on Information Theory*, vol. 50, no. 1, pp. 125–144, 2004.

[72] M. Neely, "Optimal energy and delay tradeoffs for multiuser wireless downlinks," *IEEE Transactions on Information Theory*, vol. 53, pp. 3095–3113, Sept. 2007.

[73] D. Wu, *Providing Quality-of-Service guarantees in wireless networks*. PhD thesis, Department of Electrical and Computer Engineering, Carnegie Mellon University, USA, 2003.

[74] E. Biglieri, J. Proakis, and S. Shamai, "Fading channels: information-theoretic and communications aspects," *IEEE Transactions on Information Theory*, vol. 44, no. 6, pp. 2619–2692, 1998.

[75] L. Ozarow, S. Shamai, , and A. D. Wyner, "Information theoretic consideration for cellular mobile radio," *IEEE Transactions on Vehicular Technology*, vol. 43, pp. 359–378, May 1994.

[76] V. Sivaraman and F. Chiussi, "Providing end-to-end statistical delay guarantees with earliest deadline first scheduling and per-hop traffic shaping," in *IEEE Conference on Computer Communications*, vol. 2, (Israel), pp. 631–640, 2000.

[77] D. Wu and R. Negi, "Effective capacity: a wireless link model for support of quality of service," *IEEE Transactions on Wireless Communications*, vol. 2, no. 6, pp. 630–643, 2003.

[78] S. Bhashyam, A. Sabharwal, and B. Aazhang, "Feedback gain in multiple antenna systems," *IEEE Transactions on Communications*, vol. 50, pp. 785–798, May 2002.

[79] J. Tang and X. Zhang, "Quality-of-Service driven power and rate adaptation over wireless links," *IEEE Transactions on Wireless Communications*, vol. 6, pp. 3058–3068, Aug. 2007.

[80] WiMAX Forum, *A Comparative analysis of mobile WiMAX deployment alternatives in the access network.*

[81] R. Negi and J. Cioffi, "Delay-constrained capacity with causal feedback," *IEEE Transactions on Information Theory*, vol. 48, no. 9, pp. 2487–2494, 2001.

[82] R. McEliece and W. Stark, "Channels with block interference," *IEEE Transactions on Information Theory*, vol. 30, no. 1, pp. 44–53, 1984.

[83] D. Bertsekas, *Dynamic programming and optimal control.* Athena Scientific, third ed., 2007.

[84] S. Boyd and L. Vandenberghe, *Convex optimization.* Cambridge, first ed., 2004.

[85] E. Chong and S. Zak, *An introduction to optimization.* John Wiley and Sons, second ed., 2001.

[86] J. Gablonsky, *Modifications of the DIRECT algorithm.* PhD thesis, North Carolina State University, USA, 2001.

[87] J. Nocedal and S. Wright, *Numerical optimization.* Springer, second ed., 1999.

[88] J. Luo, L. Lin, R. Yates, and P. Spasojevic, "Service outage based power and rate allocation," *IEEE Transactions on Information Theory*, vol. 49, pp. 323–330, Jan. 2003.

[89] J. Luo, R. Yates, and P. Spasojevic, "Service outage based power and rate allocation for parallel fading channels," *IEEE Transactions on Information Theory*, vol. 51, no. 7, pp. 2594–2611, 2005.

[90] R. Negi, M. Charikar, and J. Cioffi, "Minimum outage transmission over fading channels with delay constraint," in *IEEE International Conference on Communication*, vol. 1, pp. 282–286, 2000.

[91] X. Liu and A. Goldsmith, "Optimal power allocation over fading channels with stringent delay constraints," in *IEEE International Conference on Communications*, vol. 3, pp. 1413–1418, May 2002.

[92] K. Wong, "Stochastic power allocation using causal channel information for delay-limited communications," *IEEE Communications Letters*, vol. 10, pp. 748–750, Nov. 2006.

[93] A. James, "Distributions of matrix variates and latent roots derived from normal samples," *Annals of Mathematical Statistics*, vol. 35, pp. 475–501, 1964.

[94] P. Smith and M. Shafi, "On a Gaussian approximation to the capacity of wireless MIMO systems," in *IEEE International Conference on Communications*, 2002.

[95] I. Gradshteyn and I. Ryzhik, *Table of integrals, series, and products*. Academic Press, sixth ed., 2000.

[96] D. Jones, C. Perttunen, and B. Stuckman, "Lipschitzian optimization without the Lipschitz constant," *Optimization Theory and Applications*, vol. 79, pp. 157–181, Oct. 1993.

[97] J. Chen and K. Wong, "Channel-statistics assisted power minimization for central Wishart MIMO block-fading channels with an outage probability

constraint," in *Student Paper Contest in IEEE Sarnoff Symposium*, (USA), May 2007.

[98] P. Dighe, R. Mallik, and S. Jamuar, "Analysis of transmit-receive diversity in Rayleigh fading," *IEEE Transactions on Communications*, vol. 51, no. 4, pp. 694–703, 2003.

[99] V. Girko, "A refinement of the central limit theorem for random determinants," *Theory of Probability and its Applications*, vol. 42, pp. 121–129, 1997.

[100] K. Wong and J. Chen, "Near-optimal power allocation and multiuser scheduling with outage capacity constraints exploiting only channel statistics," *IEEE Transactions on Wireless Communications*, vol. 7, pp. 812–818, Mar. 2008.

[101] M. Kang and M. Alouini, "Largest eigenvalue of complex Wishart matrices and performance analysis of MIMO MRC systems," *IEEE Journal on Selected Areas Communications*, vol. 21, no. 3, pp. 418–426, 2003.

[102] D. Wu and R. Negi, "Effective capacity channel model for frequency-selective fading channels," *ACM Wireless Networks*, vol. 13, no. 3, pp. 299–310, 2007.

[103] J. Tang and X. Zhang, "Cross-layer modeling for Quality of Service guarantees over wireless links," *IEEE Transactions on Wireless Communications*, vol. 6, pp. 4504–4512, Dec. 2007.

[104] J. Tang and X. Zhang, "Quality-of-Service driven power and rate adaptation for multichannel communications over wireless links," *IEEE Transactions on Wireless Communications*, vol. 6, pp. 4349–4360, Dec. 2007.

[105] J. Tang and X. Zhang, "Cross-layer-model based adaptive resource allocation for statistical QoS guarantees in mobile wireless networks," *IEEE Transactions on Wireless Communications*, vol. 6, pp. 1–11, Dec. 2007.

[106] J. Tang and X. Zhang, "Power-delay tradeoff over wireless networks," in *IEEE International Symposium on a World of Wireless, Mobile, and Multimedia Networks*, (USA), 2008.

[107] A. Tulino and S. Verdu, *Random matrix theory and wireless communications*. Now, first ed., 2004.

[108] Z. Wang and G. Giannakis, "Outage mutual information of space-time MIMO channels," *IEEE Transactions on Information Theory*, vol. 50, no. 4, pp. 657–662, 2004.

[109] C. So-In, R. Jain, and A. Tamimi, "Scheduling in IEEE 802.16e mobile WiMAX networks: key issues and a survey," *IEEE Journal on Selected Areas in Communications*, vol. 27, no. 2, pp. 156–171, 2009.

[110] M. McKay and I. Collings, "General capacity bounds for spatially correlated Rician MIMO channels," *IEEE Transactions on Information Theory*, vol. 51, pp. 3121–3145, Sept. 2005.