University of Alberta

Efficiency and Security Analysis in Multi-User Wireless Communication Systems: Cooperation, Competition and Malicious Behavior

by

Jie Gao

A thesis submitted to the Faculty of Graduate Studies and Research
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy
in
Communications

Department of Electrical and Computer Engineering

©Jie Gao
Fall 2013
Edmonton, Alberta

Permission is hereby granted to the University of Alberta Libraries to reproduce single copies of this thesis and to lend or sell such copies for private, scholarly or scientific research purposes only. Where the thesis is converted to, or otherwise made available in digital form, the University of Alberta will advise potential users of the thesis of these terms.

The author reserves all other publication and other rights in association with the copyright in the thesis and, except as herein before provided, neither the thesis nor any substantial portion thereof may be printed or otherwise reproduced in any material form whatsoever without the author's prior written permission.

Dedicated to my parents and grand parents

Abstract

Efficiency and security are major concerns with increasingly higher importance in modern wireless communications. These two concerns are especially significant for multi-user wireless communications where different users share or compete for resources. Among different users, there are possibilities of *cooperation*, *competition*, and/or *malicious behavior*. Due to the possibility of cooperation among the users, the spectral and energy efficiency in multi-user wireless communications could be boosted. Due to the possibility of competition, the resource allocation in multi-user wireless systems may reach certain equilibrium. Due to the possibility of malicious behavior, the security and reliability of wireless communications can be undermined. In this thesis, a comprehensive analysis on the issues of efficiency and security in multi-user wireless communications is developed for three systems in four scenarios. The first multi-user system of multiple-input multiple-output two-way relaying has the feature of *cooperation* including limited coordination scenario and full coordination scenario. It is shown that high spectral efficiency can be achieved with efficient energy consumption in this system due to the cooperation among the users. Moreover, full coordination yields better results in both spectral and energy efficiency than limited coordination at the cost of higher overhead. The second multi-user system of legitimate transceiver(s) with jammer features the existence of *malicious behavior*. To measure the jamming threat, the worst-case jamming is studied for different cases according to the jammer's knowledge of the legitimate communication. The optimal/sub-optimal jamming strategy in each case is derived/analyzed. The third multi-user system of two-user interference channel features the *competition* of the users. The situation is modeled using noncooperative games with continuous mixed strategies. The outcomes of the games are analyzed

through the establishment of the conditions for the existence and uniqueness of mixed strategy Nash equilibrium.

~

Table of Contents

List of Tables

List of Figures

List of Abbreviations

Abbreviation	Definition
AF	amplify-and-forward
AWGN	additive white Gaussian noise
BC	broadcast
DF	decode-and-forward
EVD	eigenvalue decomposition
KKT	Karush-Kuhn-Tucker
MA	multiple access
MIMO	multiple-input multiple-output
MSNE	mixed-strategy Nash equilibrium
MSNEs	mixed-strategy Nash equilibria
NE	Nash equilibrium
NEs	Nash equilibria
PSD	positive semi-definite
SISO	single-input single-output
SNR	signal-to-noise ratio
SVD	singular value decomposition
TWR	two-way relaying
XOR	exclusive or

List of Symbols

x	scalar
$\mathbf{x}$	vector
$\mathbf{X}$	matrix
$\mathcal{X}$	set
$\mathbf{X}_{a\times b}$	matrix of size $a \times b$
$\mathbf{0}$	all-zero matrix of appropriate size
$\mathbf{I}$	identity matrix of appropriate size
$(\cdot)^{\mathrm{T}}$	transpose
$(\cdot)^{\mathrm{H}}$	Hermitian transpose
$\lvert x\rvert$	absolute value of number x
$\lvert\mathbf{X}\rvert$	determinant of matrix $\mathbf{X}$
$\lvert\mathcal{X}\rvert$	cardinality of set $\mathcal{X}$
$\mathbb{E}\{\cdot\}$	mathematical expectation
$\mathrm{Tr}\{\cdot\}$	trace of a matrix
$(\cdot)^{+}$	projection to positive orthant
$\mathbf{X} \succ 0$	matrix $\mathbf{X}$ is positive definite
$\mathbf{X} \succeq 0$	matrix $\mathbf{X}$ is positive semi-definite

Chapter 1

Introduction

1.1 Motivation

Wireless communications have experienced an incredibly fast development during the past decades. Take the cellular communication system as an example. While the peak bit-rate supported by the third generation (3G) mobile telecommunications technology was less than 1 megabits per second (Mbit/s) ten years ago, the current fourth generation (4G) standard supports a data rate up to 100 Mbits/s [1]. As wireless communication techniques and systems become ubiquitous and change people's lives in a unprecedented way, the issues of efficiency and security in wireless communications become more significant.

Efficiency turns out to be an increasingly more important issue due to the contrast between the ever-growing demand for wireless communication resources and the limited supply of such resources. The major resource for all wireless communications is spectrum. As the most part of the spectrum for wireless communications has been allocated, the problem of spectrum shortage appears [2], [3]. This problem will be even more severe in the future, since it is expected that the traffic volume of wireless communications will increase to above ten times of its current scale till 2016/2017 [4]. With the increase in the traffic volume and very limited supply of spectrum, the most reliable solution to the problem of spectrum shortage is improving spectral efficiency. Different techniques have been proposed to improve spectral efficiency including multiple-input multiple-output (MIMO), two-way relaying (TWR), etc. MIMO can significantly increase the data throughput of wire-

less communications without requiring extra spectrum bandwidth or transmission power [5], [6]. The high throughput is achieved by transmitting multiple streams of data through different spatial channels introduced by the multipath propagation. The improvement of spectral efficiency in TWR is achieved by allowing simultaneous data transmissions to different directions in the scenario of relay-assisted data exchange [7].

While spectrum is the common resource for all wireless communications, the transmitting sides of wireless communication systems are limited by their private resources, e.g., transmission power. In recently years, energy efficiency in wireless communications has attracted significant attention [8–10]. Energy-efficiency is important for two reasons. On a macro-scale, as the total energy consumption for wireless communications is growing at a high speed, improving energy-efficiency has both economical and environmental benefits. On a micro-scale, improving energy-efficiency is crucial for battery-powered wireless terminals or mobile devices. Higher efficiency is most beneficial when it is impossible or inconvenient to change or recharge the battery of the device such as a wireless sensor or a smartphone [11]. Energy efficiency of wireless communications can be improved through architecture design, resource management, the adoption of MIMO, etc.

Apart from efficiency, security is also a major concern in wireless communications. Due to the rapid development of wireless communications, the security issue rises while wireless communication systems of different scales and devices for different purposes become more common and popular. Major threats to wireless communications include passive wiretapping and active jamming. While the passive threat can be addressed by using well-designed security architectures, wireless communications are vulnerable to the active jamming attack [12]. Jamming aims at degrading the quality of or disrupting the information exchange in a communication system by directing energy toward the target receiver in a destructive manner [13]. A jamming attack is particularly effective because it is easy to launch using low-cost and small-sized devices while causing very significant results [14]. One example is that a jammer can drain the power of sensors in a wireless sensor network by disrupting their transmissions which results in the continuous repeti-

tions of their signal transmission [15]. Another example is that the jammer can cancel or significantly weaken a target signal at its intended receiver if the jammer knows that signal [16]. For estimating the jamming threat to a wireless communication system, the study of worst-case jamming can be adopted [17], [18].

In modern wireless communications, multi-user communications are of increasing importance. Due to the growth of the comparatively new applications in wireless communications such as wireless local area network, multi-user wireless communication systems are becoming very common and popular. As an example, the cell phones and laptops of costumers in a cafeteria connected to its Wi-Fi constitute a typical multi-user system. Basic models of multi-user systems include relay networks, interference channels, multiple access (MA) channels, broadcast (BC) channels, etc. In multi-user wireless communications, the existence of multiple users in the same wireless environment in general leads to the sharing of the spectral resource in the system. The wireless users can compete or coordinate with each other in sharing the resource. In turn, the performance of a multi-user system depends on the behavior of each and every wireless user in it. Given this unique feature, the investigation of efficiency and security in multi-user systems is of interest. Specifically, three questions come into sight:

1. Given that a TWR network is a spectral-efficient multi-user system, is it possible to achieve both spectral and energy efficiency in a TWR network?

2. In a multi-user system with both legitimate wireless communication(s) and a malicious jammer, what is the worst-case jamming threat if the jammer is able to optimize its jamming scheme?

3. When multiple users share and compete for the spectral resource in a multi-user system, is there a proper model to characterize the interactions of users and the resulting outcome of competition?

In order to answer the above three questions, the related literature is reviewed and discussed in detail in the following three subsections, respectively.

1.1.1 Efficiency in TWR

The performance of TWR depends on the transmit strategies of the participating nodes, i.e., the source nodes and the relay. Optimizing the transmit strategies such as power allocation and beamforming is one of the main research interests in TWR, especially when MIMO is considered [19–27]. The optimization of the transmit strategies in a TWR system helps to maximize the spectral efficiency in terms of sum-rate.

The transmit strategies of the relay and source nodes depend on the relaying scheme. Similar to one-way relaying, the relaying scheme in TWR can be amplify-and-forward (AF), decode-and-forward (DF), etc. Spectral efficiency for AF MIMO TWR is investigated in [28–31]. Transmit strategies for maximizing the weighted sum-rate of a TWR system are studied in [29], where the optimal solution is found through alternative optimization over the transmit strategies of the relay and source nodes. In [30], a low-complexity sub-optimal design of relay and source node transmit strategies is derived for either sum-rate maximization or power consumption minimization under quality-of-service requirements. The authors of [31] solve the robust joint source and relay optimization problem for a MIMO TWR system with imperfect channel state information. The joint optimization of transmit strategies for AF TWR is in general a nonconvex problem. Low-complexity sub-optimal solution can be obtained through diagonalizing the MIMO channel based on the singular value decomposition (SVD) or the generalized SVD and thereby transferring the transmit strategy of the participating nodes to power allocation problem. Finding the optimal solution, however, usually requires iterative algorithms with high complexity.

DF TWR has the advantage over AF TWR that it does not suffer from the problem of noise propagation. As a result, DF TWR may achieve a better performance than AF TWR, especially at low signal-to-noise ratio (SNR), at the cost of higher requirement on the relay due to the decoding and re-encoding [32]. Spectral efficiency for DF TWR has been studied in [33–35]. The optimal time division between the MA and BC phases and the optimal distribution of the relay's power for achieving weighted sum-rate maximization are studied in [33]. The achievable rate region

and the optimal transmit strategies of both the source nodes and the relay are studied in [34], where the relay's optimal transmit strategy is found by two water-filling based solutions coupled by the relay's power limit. The authors of [35] specifically investigate the optimal transmit strategy in the BC phase of the MIMO DF TWR. It is shown that there may exist different strategies that lead to the same point in the rate region.

Given the fact that DF TWR may achieve better performance than AF TWR especially at low SNR and the fact that different transmit strategies of the participating nodes in DF TWR can lead to the same spectral efficiency (in terms of the achieved sum-rate in the system), it is logical to ask how to find a strategy that has the best energy-efficiency among these strategies that lead to the maximum spectral efficiency. There is no answer to this question in the literature despite the above mentioned works regarding the spectral efficiency in TWR and those on the energy efficiency in TWR [30, 36, 37], which aim at minimizing the power consumption in TWR subject to quality of service constraints. This thesis work will fill this research gap.

1.1.2 Jamming threat in MIMO multi-user wireless communications

The jamming threat in wireless communications has been studied in many research works [38–44]. One of research interests in jamming is to investigate the feasibility and effect of jamming from the perspective of the jammer [12, 18, 39–41]. The research works adopting this perspective provide insights in understanding and measuring the threat of worst-case jamming to the target legitimate communications.

Different types of jamming are investigated in the literature. Noise jamming is a common type of jamming in the case that the jammer has no or limited information on the target signal [18]. Noise jamming impairs the legitimate communication through decreasing the SNR at the target receiver. The effect of noise jamming depends on the power of the jamming signal.

It is also possible that the jammer has the knowledge of the target signal, e.g.,

in the case that the jammer can also perform eavesdropping [42]. With such information, the jammer can use another type of jamming, i.e., the correlated jamming, to damage the legitimate communication by canceling or weakening the target signal at the target receiver [43]. The jammer's strategies for correlated jamming are studied in [16], [44].

It should be noted that with the development and application of MIMO wireless communications, a jammer equipped with multiple antennas will become common and pose a larger threat to the legitimate communication due to its ability to optimize its jamming strategy over the antennas. However, unfortunately, most of the aforementioned works on jamming threat focus on the single-input single-output (SISO) case. The results of jamming threat in the scenario that both the legitimate transceiver and the jammer have multiple antennas are limited. The jammer's strategy for worst-case noise jamming is investigated for MIMO MA and BC channels in [17, 45–47]. It is shown in [45] that without knowledge of the target signal or its covariance, the jammer can only use basic strategies of allocating power uniformly or maximizing the total power of the interference at the target receiver. In [46], the transmit strategies of a legitimate transmitter and a jammer on a Gaussian MIMO channel are investigated under a game-theoretic modeling with a general utility function. It is assumed that the jammer and the legitimate transmitter have the same level of channel state information (CSI), i.e., both uninformed, both with statistical CSI, or both with exact CSI. The optimal transmitted strategies of the legitimate transmitter and the jammer are represented as solutions to different problems versus different types of CSI. The worst-case jamming on MIMO multiple access and broadcast channels with the covariance of the target signal and all channel information available at the jammer is studied in [17] using game theory. Some properties of the optimal jamming strategies are characterized yet the optimal jamming solutions are not given. The necessary condition for optimal jamming on MIMO channels with arbitrary inputs when the covariance of the target signal and all channel information are available at the jammer is derived in [47]. For the case of Gaussian target signal, the solution of optimal jamming is given in closed-form. However, it is derived without considering the jamming channel. As a result, the

system model is simplified by implicitly assuming that the received jamming signal at the target receiver is exactly the same as the transmitted jamming signal at the jammer. The correlated jamming on MIMO Gaussian fading channel is studied in [16]. However, the study in [16] considers only one legitimate communication. Therefore, the measure of the worst-case noise jamming threat in general MIMO wireless communications and the correlated jamming threat in multi-target wireless communications remains an open problem, and will be investigated in this thesis.

1.1.3 Game theoretic study of wireless multi-user systems

Game theory studies the interactions of decision makers, in conflict or in cooperation, during their strategic decision making process [48], [49]. While it has been used to model problems in economics, political science, and many other areas, the application of game theory in wireless communications has attracted tremendous research interests during the past decade [50–60].

The problem of sharing and competing for spectral resources among users in a multi-user system with no central administration or coordination can be modeled in terms of noncooperative games [54, 55, 61, 62]. A mutual information game in the Gaussian interference channel is studied in [54], and the conditions guaranteeing the uniqueness of the Nash equilibrium (NE) are derived. In [55], a power control game for maximizing spectral efficiency is investigated and an algorithm is designed for achieving an efficient NE within multiple Nash equilibria (NEs). The common interest in these works is to study the existence and uniqueness of NE.

Most game theoretic studies of multi-user wireless communications, such as the above mentioned works, focus on pure strategies. However, there are strong motivations to investigate mixed strategies as an extension. Mixed strategies introduce deliberate randomness into the decision of a player such that the player can use more subtle strategies in the competition with other players. In consequence, the utilities obtained due to applying mixed strategies can be potentially improved for the users. Introducing mixed strategies is also instrumental for capturing the stochastic regularities of equilibria and players' strategies in noncooperative games [63].

There are just a few works on games with mixed strategies in the literature

[62, 64–66]. In [66], a two-user channel selection game is considered. Each user in the game assigns different probabilities to different power levels that it uses to communicate on each channel. A transmitter's channel selection game is studied in [62] and the mixed-strategy Nash equilibrium (MSNE) of the game is found. All the above works consider games in which the users' strategies are represented by discrete probability mass functions.

Considering the fact that discrete probability mass functions are only special cases of continuous probability distributions, a more general representation of mixed strategies is to use continuous distributions. The resulting games are continuous games. By introducing maximum flexibility in the users' choice of strategies, the continuous game modeling of multi-user systems captures the very essence of the strategy making and interaction. However, the modeling of multi-user wireless systems using continuous games is an open area of research, and will be addressed in this thesis for the two-user case.

1.2 Proposed research problems

Motivated by the aforementioned literature, this thesis proposes four problems, **P1**-**P4**, in multi-user wireless communications as stated in the following paragraphs.

P1: *Energy efficient sum-rate maximization in relay-oriented MIMO DF TWR.* The spectral efficiency of TWR is determined by the transmit strategies of the participating nodes. The maximum spectral efficiency that can be achieved depends on the level of coordination among the sources and the relay. The relay-oriented cooperation, in which the relay adjusts its own strategy according to the transmit strategies of the source nodes, has the advantage of low overhead. Given the transmit strategies of the source nodes, different transmit strategies of the relay may lead to the same spectral efficiency with different relay power consumptions in DF TWR. Therefore, it is of interest to find out the most energy-efficient transmit strategy of the relay that has minimum power consumption among all the strategies that maximize the spectral efficiency. The resulting strategy maximizes the spectral efficiency with the best

energy efficiency in the relay-oriented cooperation scenario.

P2: *Energy efficient sum-rate maximization in MIMO DF TWR with full cooperation.* While the relay-oriented cooperation has the advantage of low overhead, it does not achieve the maximum spectral efficiency due to the lack of cooperation of the source nodes. At the cost of higher overhead, higher spectral efficiency can be achieved if all the participating nodes (source nodes and relay) can jointly optimize their transmit strategies. Moreover, the energy efficiency can also be improved when all the nodes cooperate. Most research works on joint transmit strategy design in MIMO TWR focus on achieving the maximum spectral efficiency. However, different transmit strategies may lead to the same spectral efficiency with different total power consumption in the system. The investigation on the optimal transmit strategy of the source nodes and the relay that maximizes the spectral efficiency with minimum power consumption has been lacking. Therefore, finding the above energy efficient optimal transmit strategy is of interest.

P3: *Jamming and correlated jamming in multi-user wireless communications.* The security threat of jamming to a wireless communication system can be measured by studying the worst-case jamming to the legitimate communication. The damage that jamming can cause depends on the jammer's knowledge of the channels and the target signal. The less knowledge available to the jammer, the simpler the strategy that it can use. When the jammer has the knowledge of the channels and the statistics of the target signal, it can optimize its strategy to effectively degrade the information rate of the target channel. Furthermore, if the jammer knows the exact target signal, it can cancel or significantly weaken the target signal at the target receiver by performing correlated jamming under some conditions. The jammer's strategies and the resulting effects to the legitimate communication in the above two cases are important yet missing parts of the worst-case jamming in wireless communications.

P4: *Mixed strategy and MSNE in resource allocation games.* While the application of game theory in wireless communications is popular, the research works considering mixed strategies are limited. As mix strategies introduce deliberate randomness into the strategies of the players, it is suitable for modeling some scenarios in wireless communications. As an example, the resource allocation game with mixed strategy is considered in this thesis. Unlike most games with mixed strategies considered in the literature of wireless communications, the wireless users' strategies are represented by continuous probability distributions. The existence and uniqueness of the MSNE is then investigated given the above model.

1.3 Thesis outline

Chapter 2 presents the background of the thesis. Chapters 3–6 provide details for solving the problems **P1–P4**, respectively. The brief outline is as follows.

- Chapter 2 gives the background on the topics related to this thesis. The signal model, capacity, and power allocation of MIMO wireless communication channels are reviewed. The idea of TWR is explained while the signal model under AF and DF are given. The effect of jamming on the target signal at the receiver depending on the jammer's knowledge of the channels and the target signal is presented. The concept of correlated jamming is introduced. The basics of game theory, NE and MSNE are illustrated using examples.

- Chapter 3 addresses the problem **P1**. A sufficient and necessary condition for the optimal strategy of the relay in the scenario of relay-oriented cooperation in TWR is derived. Based on the above condition, an algorithm is designed for the relay to calculate its optimal strategy. The results of the power allocations are discussed versus the power limits of the relay and source nodes. It is shown that power could be wasted at the source nodes since they do not cooperate with the relay. Simulation results demonstrate the effectiveness of the obtained optimal strategy as well as the effect of asymmetry in the system.

- Chapter 4 addresses the problem **P2**. At the cost of higher overhead, the performance of the TWR considered in Chapter 3 can be improved by introducing full cooperation among the source nodes and relay. In this scenario, properties of the optimal strategies of the relay and the source nodes are derived in different cases. The optimal solution is given in each case, either in closed-form or through algorithms with a very limited number of steps. The results are simulated and compared with the optimal strategy in Chapter 3. The effect of asymmetry in the system is also shown in simulations.

- Chapter 5 addresses the problem **P3**. For noise jamming, the worst-case jamming to the communication over a MIMO target channel is derived. It is shown that the worst-case jamming can be given in closed-form under a certain condition. When the condition does not hold, an algorithm is provided to calculate the worst-case jamming while a closed-form approximation is given. For correlated jamming, the problem of multi-target correlated jamming is considered in the SISO case and proved to be convex.

- Chapter 6 addresses the problem **P4**. Modeling the channel selection and power allocation of two wireless users using games with mixed strategy, the results regarding the existence and uniqueness of MSNE are derived first for a two-user two-channel game and then for a two-user N-channel game. In the two-channel game, the MSNE which maximizes the utilities for both users is obtained, while for the N-channel game, an algorithm is provided to perform channel selection for users in order to achieve MSNE.

- Chapter 7 presents the conclusion of this thesis. Future research directions are also provided.

$\sim$

Chapter 2

Background

This chapter reviews the essentials of relevant topics for this thesis. The signal model and power allocation problem for MIMO systems is presented. The basic forms of multi-user MIMO, i.e., the MA and BC channels, are briefly reviewed. The concept, model, and relaying strategy is summarized for TWR. The mathematic model of jamming and noise jamming is provided with discussions. Finally, the basics of game theory, Nash equilibrium (NE) and mixed strategies are illustrated using examples.

2.1 Power allocation in MIMO wireless communications

2.1.1 Basic MIMO channel: signal model and capacity

Consider the data transmission from a transmitter with n_{t} antennas to a receiver with n_{r} antennas. Assuming additive white Gaussian noise (AWGN) at the receiver, the signal model for this basic MIMO channel can be expressed as

$$\overbrace{\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_{n_{\mathrm{r}}} \end{bmatrix}}^{\mathbf{y}} = \overbrace{\begin{bmatrix} h_{11} & \dots & h_{1n_{\mathrm{t}}} \\ \vdots & \ddots & \\ h_{n_{\mathrm{r}}1} & & h_{n_{\mathrm{r}}n_{\mathrm{t}}} \end{bmatrix}}^{\mathbf{H}} \overbrace{\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_{n_{\mathrm{t}}} \end{bmatrix}}^{\mathbf{x}} + \overbrace{\begin{bmatrix} n_1 \\ n_2 \\ \vdots \\ n_{n_{\mathrm{r}}} \end{bmatrix}}^{\mathbf{n}} \tag{2.1}$$

where $\mathbf{x}$, $\mathbf{y}$, $\mathbf{H}$, and $\mathbf{n}$ represent the transmitted signal, the received signal, the MIMO channel, and the noise, respectively. The element h_{kl} of the channel $\mathbf{H}$

represents the path gain from the lth antenna of the transmitter to the kth antenna of the receiver, which is determined by different factors including the path loss and fading effects. The elements of $\mathbf{x}$ and $\mathbf{y}$ represent the transmitted and received data streams, respectively. The noise is assumed to have zero mean and covariance $\sigma^2\mathbf{I}$ where $\mathbf{I}$ stands for the identity matrix.

If $\mathbf{H}$ is constant (valid under slow fading) and known at the transmitter and the receiver, the corresponding capacity of this MIMO channel is given as [67]

$$C = \max_{\mathbf{Q}_x} \log \left| \mathbf{I} + \frac{1}{\sigma^2} \mathbf{H} \mathbf{Q}_x \mathbf{H}^{\mathrm{H}} \right| \tag{2.2}$$

where $\mathbf{Q}_x$ is the covariance of the transmitted signal defined as $\mathbf{Q}_x \triangleq \mathbb{E}\{\mathbf{x}\mathbf{x}^{\mathrm{H}}\}$, $|\cdot|$ stands for the determinant, $(\cdot)^{\mathrm{H}}$ stands for the Hermitian transpose, and $\mathbb{E}\{\cdot\}$ represents the mathematical expectation.

From the expression (2.2), it can be seen that $\mathbf{Q}_x$ needs to be optimized to maximize the information rate of the considered MIMO channel. The optimization over $\mathbf{Q}_x$ is usually subject to the following trace constraint

$$\mathrm{Tr}\{\mathbf{Q}_x\} \leq P \tag{2.3}$$

in which $\mathrm{Tr}\{\cdot\}$ stands for the trace and P represents the power limit of the transmitter. The maximization of (2.2) over $\mathbf{Q}_x$ subject to (2.3) is the basic form of power allocation problem over MIMO channel. The optimal solution to this power allocation problem is the waterfilling-based solution described in detail as follows.

2.1.2 Waterfilling based power allocation

The first step for optimizing the power allocation, is the singular value decomposition (SVD) of the MIMO channel. Assume that the rank of $\mathbf{H}$ is r, where $r \leq \min(n_{\mathrm{r}}, n_{\mathrm{t}})$. The SVD of $\mathbf{H}$ can be written as

$$\mathbf{H} = \mathbf{U} \overbrace{\left[\begin{array}{ccc|c} \omega_1 & 0 & \dots & \\ 0 & \ddots & & \mathbf{0}_{r\times(n_{\mathrm{t}}-r)} \\ \vdots & & \omega_r & \\ \hline \multicolumn{3}{c|}{\mathbf{0}_{(n_{\mathrm{r}}-r)\times r}} & \mathbf{0}_{(n_{\mathrm{r}}-r)\times(n_{\mathrm{t}}-r)} \end{array}\right]}^{\mathbf{\Omega}} \mathbf{V}^{\mathbf{H}} \tag{2.4}$$

where $\mathbf{0}$ stands for all-zero matrix, and $\mathbf{U}$ and $\mathbf{V}$ are unitary matrices, i.e., $\mathbf{U}\mathbf{U}^{\mathrm{H}} = \mathbf{I}$ and $\mathbf{V}\mathbf{V}^{\mathrm{H}} = \mathbf{I}$, of sizes $n_{\mathrm{r}} \times n_{\mathrm{r}}$ and $n_{\mathrm{t}} \times n_{\mathrm{t}}$, respectively. The diagonal elements $\omega_1, \ldots, \omega_r$ are positive singular values of $\mathbf{H}$, or equivalently $\mathbf{\Omega}$. These positive singular values of $\mathbf{H}$ represent the effective parallel sub-channels of the MIMO channel. The number of sub-channels determines the maximum number of different data streams that can be transmitted and received over the MIMO channel.

Given the above SVD, the capacity of the MIMO channel can be rewritten as

$$C = \max_{\mathbf{Q}_x'} \log \left| \mathbf{I} + \frac{1}{\sigma^2} \mathbf{\Omega} \mathbf{Q}_x' \mathbf{\Omega}^{\mathrm{H}} \right| \tag{2.5}$$

where $\mathbf{Q}_x' \triangleq \mathbf{V}^{\mathrm{H}} \mathbf{Q}_x \mathbf{V}$ is of size $n_{\mathrm{t}} \times n_{\mathrm{t}}$. The constraint (2.3) now applies as $\mathrm{Tr}\{\mathbf{Q}_x'\} \leq P$. Since $\mathbf{Q}_x'$ is Hermitian positive semi-definite, according to the Hadamard's determinant inequality, the optimal $\mathbf{Q}_x^{'}$ that maximizes (2.5) should be diagonal [68]. Denote the kth diagonal element of $\mathrm{Tr}\{\mathbf{Q}_x^{'}\}$ as q_k. Then the equation (2.5) can be further rewritten as

$$C = \max_{\{q_1, \ldots q_r\}} \sum_{k=1}^{r} \log \left(1 + \frac{\omega_k^2 q_k}{\sigma^2} \right). \tag{2.6}$$

The trace constraint on $\mathbf{Q}_x^{'}$ simplifies as $\sum_{k=1}^{r} q_k \leq P$. From the above simplification, it can be seen that the power allocation problem over the MIMO channel is in essence the power allocation problem over the sub-channels.

The solution to the problem can be derived using Lagrangian methods. The optimal solution can be written as [69]

$$q_k = \begin{cases} \left(\frac{1}{\lambda} - \frac{\sigma_2}{\omega_k^2} \right)^+, & k = 1, \ldots, r \\ 0, & k > r \end{cases} \tag{2.7}$$

where λ is a constant chosen such that the power limit is satisfied with equality (i.e., $\sum_{k=1}^{r} q_k = P$) and $(\cdot)^+ = \max\{\cdot, 0\}$. The solution given by (2.7) is called the waterfilling-based power allocation, while $1/\lambda$ is called the water level. The term "waterfilling" is used because the resulting power allocation on sub-channel k fills the gap between the water level $/\lambda$ and σ_2/ω_k^2 (which reflects the quality of the kth sub-channel) if the water level is higher.

2.1.3 MIMO MA and BC channels

The previous results on capacity and power allocation apply to the MIMO channel with a single pair of transceiver. In this section, we are going to review two multi-user MIMO channels - the MIMO MA channel and the MIMO BC channel.

First consider the MIMO multiple access (MA) channel. Assume that there is one receiver with n_{r} antennas and M transmitters. The ith transmitter has $n_{\mathrm{t}i}$ antennas and transmits the signal $\mathbf{x}_i$ to the receiver. Denote the channel from the ith transmitter to the receiver as $\mathbf{H}_i$. Then the received signal at the receiver is expressed as

$$\mathbf{y} = \sum_{i=1}^{M} \mathbf{H}_i \mathbf{x}_i + \mathbf{n} \tag{2.8}$$

where $\mathbf{n}$ is the AWGN at the receiver. Denote $\mathbb{E}\{\mathbf{x}_i \mathbf{x}_i^{\mathrm{H}}\} = \mathbf{Q}_i$. The power of the ith transmitter is limited such that $\mathrm{Tr}\{\mathbf{Q}_i\} \leq P_i$. The sum-capacity of this MIMO multiple access channel is given as [70]

$$C = \max_{\{\mathbf{Q}_i, \forall i\}} \log \left| \mathbf{I} + \frac{1}{\sigma^2} \sum_{i=1}^{M} \mathbf{H}_i \mathbf{Q}_i \mathbf{H}_i^{\mathrm{H}} \right| \tag{2.9}$$

subject to the power limits of the transmitters. For deriving the optimal power allocation that achieves the sum-capacity, the procedure of iterative waterfilling can be used. Define

$$\mathbf{L}_i \triangleq \mathbf{I} + \frac{1}{\sigma^2} \sum_{j \neq i, j=1}^{M} \mathbf{H}_j \mathbf{Q}_j \mathbf{H}_j^{\mathrm{H}} \tag{2.10}$$

and denote the eigenvalue decomposition (EVD) of $\mathbf{L}_i$ as $\mathbf{L}_i = \mathbf{U}_i \mathbf{\Lambda}_i \mathbf{U}_i^{\mathrm{H}}$. It is proved [70] that the optimal power allocation can be achieved at convergence if the transmitters iteratively update their power allocation according to the single-user waterfilling procedure described in Section 2.1.2 but with the channel $\mathbf{H}$ replaced by $\mathbf{\Lambda}_i^{-\frac{1}{2}} \mathbf{U}_i^{\mathrm{H}} \mathbf{H}_i$.

In a MIMO broadcast (BC) channel, one transmitter sends information to M receivers. The transmitted information $\mathbf{x}$ is a summation of the messages ($\mathbf{x}_1$, ..., $\mathbf{x}_M$) intended for all the receivers. Each receiver needs to extract its own message from the received signal while the messages intended for other receivers become interference. For the MIMO BC channel, the sum-capacity can be reached using dirty paper coding. Details can be found in [67] and are omitted here.

2.2 Two-way relaying

For the case that two source nodes need to exchange messages via a relay, TWR improves spectral efficiency by allowing transmissions on both directions to proceed simultaneously.

2.2.1 Basic idea

Assume that there are two source nodes S1 and S2 that need to exchange information through a relay as shown in Fig. 2.1. The source node S1 needs to send its message x_1 to S2 while S2 needs to send its message x_2 to S1. There is no direct link between S1 and S2 and therefore all traffic goes through the relay.

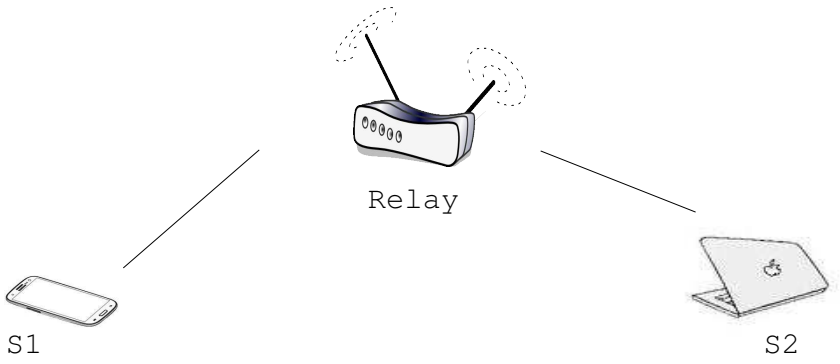


Figure 2.1: A basic relay system.

If the information exchange is achieved using one-way relaying, four time slots are required as illustrated in Fig. 2.2. From this figure, it can be seen that there is always one node idle in each of the four time slots. This is the consequence of the half-duplex mode of the nodes, i.e., they cannot transmit and receive message simultaneously. As a result, there is one idle channel (between one source node and the relay) in any time slot, which leads to low spectral efficiency.

However, using TWR, the spectral efficiency can be significantly increased by achieving the same information exchange in just two time slots. The idea is that the two source nodes S1 and S2 can transmit simultaneously and receive simultaneously as illustrated in Fig. 2.3 [7]. As a result, no channel is idle at any time despite the fact that all node still work in the half-duplex mode.

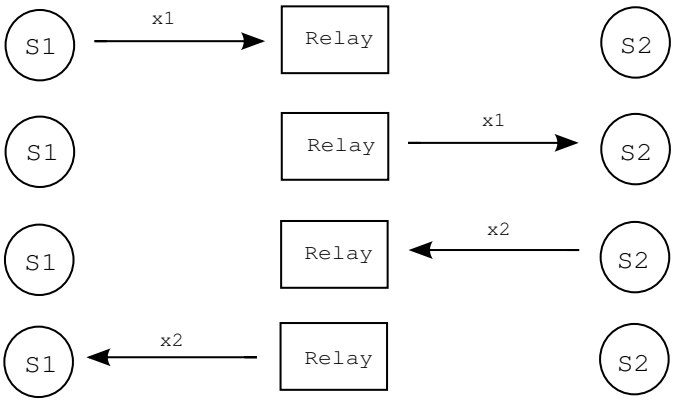


Figure 2.2: The procedure of information exchange in one-way relaying.

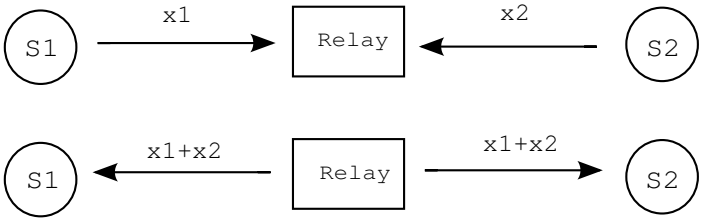


Figure 2.3: The procedure of information exchange in two-way relaying.

It can be seen that the two time slots of information exchange in TWR correspond to two phases. i.e., the MA phase in the first time slot and the BC phase in the second time slot.

2.2.2 Relaying strategy

The relaying strategy refers to the manner that the received message is processed at the relay before it is forwarded to the intended destination(s). Consider the basic case with single-antenna at all nodes as an example. Denote the channels from the source nodes S1 and S2 to the relay as h_1 and h_2, respectively. Assume channel reciprocity holds, i.e., the channels from S1 to the relay and from the relay to S1 are identical. The received signal at the relay is given as [7]

$$y_{\mathrm{r}} = h_1 x_1 + h_2 x_2 + n_{\mathrm{r}} \tag{2.11}$$

where n_{r} represents the AWGN at the receiver of the relay with variance σ_{r}^2.

The relay can choose from AF, DF, etc. relay strategies [32].

In the AF relaying strategy, the relay simply amplifies $y_{\rm r}$ by a gain α (determined by the relay's power limit $P_{\rm r}$) and broadcasts $\alpha y_{\rm r}$ to S1 and S2 in the BC phase. The received signals at the source nodes S1 and S2 are given as

$$y_1 = \alpha h_1 h_1 x_1 + \alpha h_1 h_2 x_2 + \alpha h_1 n_{\rm r} + n_1 \quad (2.12)$$

$$y_2 = \alpha h_2 h_2 x_2 + \alpha h_2 h_1 x_1 + \alpha h_2 n_{\rm r} + n_2 \quad (2.13)$$

where n_1 and n_2 are the noise at the receiver of S1 and S2 with variance σ_1^2 and σ_2^2, respectively. The first term in each of the above two expressions represents the back-propagating self-interference for the corresponding source node. The second term represents the intended signal. The third term represents the propagated noise from the relay. Since the source nodes know their own messages, i.e., S1 knows x_1 and S2 knows x_2, the self-interference can be subtracted from the received signal if S1 has the knowledge of h_1 and S2 has the knowledge of h_2. For constant channels, the sum-rate of the two source nodes is given as

$$R^{\rm s} = \frac{1}{2}\log\left(1 + \frac{\alpha^2 |h_1 h_2|^2 P_2}{\sigma_1^2 + \alpha^2 |h_1|^2 \sigma_{\rm r}^2}\right) + \frac{1}{2}\log\left(1 + \frac{\alpha^2 |h_2 h_1|^2 P_1}{\sigma_2^2 + \alpha^2 |h_2|^2 \sigma_{\rm r}^2}\right) \quad (2.14)$$

where $P_i = \mathbb{E}\{|x_i|^2\}, i = 1, 2$.

The AF relaying strategy features low complexity since the relay does not need to decode the received messages. However, the noise at the receiver of the relay is also amplified and forwarded to both source nodes. Therefore, the performance of AF-relaying can be poor at low SNRs [32], [71].

If the relay uses the DF relaying strategy, it first decodes x_1 and x_2 from $y_{\rm r}$ and then re-encodes them using either superposition or exclusive or (XOR) coding [22]. The superposition based re-encoding is the one illustrated in Fig. 2.3. In this method, the relay performs symbol-level superposition

$$x_{\rm r} = \sqrt{\beta P_{\rm r}} x_1 + \sqrt{(1-\beta) P_{\rm r}} x_2 \quad (2.15)$$

on the two messages where $\beta P_{\rm r}$ and $(1-\beta)P_{\rm r}$ represent a division of the relay's power. Then $x_{\rm r}$ is forwarded to both source nodes. The received signals at the

source nodes S1 and S2 are expressed as

$$y_1 = \sqrt{\beta P_{\rm r}} h_1 x_1 + \sqrt{(1-\beta) P_{\rm r}} h_1 x_2 + n_1 \quad (2.16)$$

$$y_2 = \sqrt{(1-\beta) P_{\rm r}} h_2 x_2 + \sqrt{\beta P_{\rm r}} h_2 x_1 + n_2 \quad (2.17)$$

where the first term in each expression represents self-interference and the second term represents the intended signal. With the channel information, each source node subtracts the self-interference from the received signal. The sum-rate is bounded by the MA phase sum-rate $R^{\rm ma}$ and BC phase sum-rate $R_1(\beta)+R_2(1-\beta)$ and is given as

$$R^{\rm s} = \frac{1}{2}\min\{R^{\rm ma}, R_1(\beta) + R_2(1-\beta)\} \quad (2.18)$$

where

$$R^{\rm ma} = \log\left(1 + \frac{P_1|h_1|^2 + P_2|h_2|^2}{\sigma_{\rm r}^2}\right) \quad (2.19)$$

$$R_1(\beta) = \min\left\{\log\left(1 + \frac{P_1|h_1|^2}{\sigma_{\rm r}^2}\right), \log\left(1 + \frac{\beta^2 P_{\rm r}|h_2|^2}{\sigma_2^2}\right)\right\} \quad (2.20)$$

$$R_2(1-\beta) = \min\left\{\log\left(1 + \frac{P_2|h_2|^2}{\sigma_{\rm r}^2}\right), \log\left(1 + \frac{(1-\beta)^2 P_{\rm r}|h_1|^2}{\sigma_1^2}\right)\right\}. \quad (2.21)$$

For the DF relaying strategy, the relay can also use the XOR coding. In this method, the relay decodes x_1 and x_2 into two bit streams and performs the XOR operation on the two streams to obtain a new bit stream. Then, this new bit stream is encoded into a symbol $x'_{\rm r}$ and forwarded to the source nodes. Each source node decodes $x'_{\rm r}$ from its received signal and obtains the corresponding bit stream. The source node can restore the message intended to it by performing the XOR on this bit stream and the bit stream of its own message. Details can be found in [72], [73] and are omitted here. Compared to superposition, XOR coding can achieve better performance in terms of sum-rate because the relay does not need to split power for transmitting two symbols. However, XOR coding has the limitation that it is appropriate only when the rates from the two source nodes to the relay are close to each other. For both superposition and XOR coding, there is no performance degradation due to noise propagation, especially at low SNRs .

2.2.3 Power allocation in TWR

Although the sum-rate expressions (2.14) and (2.18) are derived for TWR, it can be seen from these expressions that the performance of two-way relaying (TWR) depends on the power allocation of all participating nodes, especially for the DF relaying strategy. It follows from (2.14) that the relay and the source nodes should simply use their maximum power to maximize the sum-rate in AF TWR. However, it is different for the DF relaying strategy. It can be seen from the sum-rate expression in (2.18)-(2.21) that the sum-rate in DF TWR is bounded by both the MA phase and the BC phase sum-rates. The differences for the DF strategy as compared to the AF one include:

- the relay needs to find optimal division of its power,
- the relay may not need to use full power to maximize the sum-rate,
- the source nodes may not need to use full power to maximize the sum-rate.

Therefore, the optimal power allocation to achieve maximum sum-rate for DF TWR is not straightforward. The problem of finding the optimal power allocation for DF TWR is studied in [74]. A similar problem is considered in [75] with the assumption that the source nodes have equal power budgets. Including fairness as a consideration, the optimal power allocation for DF TWR is studied in [21]. Other studies on SISO DF TWR include the optimal time division between the MA and BC phases and the optimal distribution of the relay's power [33], and the minimization of the total transmit power consumption under the bit error rate constraints [76].

In the case of multiple antennas, the problem becomes more complicated. When the nodes have multiple antennas, the power allocation extends to the general term of transmit strategy including beamforming and precoding. Finding the optimal transmit strategies of the participating nodes is not straightforward in both AF and DF TWR. This problem has become the focus of many research efforts. Sum-rate maximization for MIMO AF TWR, in which the relay and the source nodes all have multiple antennas, is investigated in [28], [29], while a mean squared error minimizing scheme for MIMO AF TWR is studied in [19]. The main challenge in

investigating AF TWR is the coupling between the transmit strategies of the source nodes and the relay due to noise propagation. As a result of noise propagation, the optimization over the transmit strategies of the source nodes and the relay usually leads to nonconvex problems. For example, the information rate of the communication in either direction is a nonconvex function of the covariance/beamforming matrices of the source nodes and the relay [7].

DF TWR with multiple antennas has also been studied [23], [34], [35]. The achievable rate region and the optimal transmit strategies of both the source nodes and the relay are studied in [34], where the relay's optimal transmit strategy is found by two water-filling based solutions coupled by the relay's power limit. The authors of [35] specifically investigate the optimal transmit strategy in the BC phase of the MIMO DF TWR. It is shown that there may exist different strategies that lead to the same point in the rate region.

2.3 Jamming and correlated jamming

There are different types of jamming such as noise jamming, intelligent jamming, etc. [13]. To avoid confusion, we use the term "jamming" in this thesis to refer to noise jamming unless otherwise specified.

2.3.1 Noise jamming

With the presence of jamming, the target receiver sees extra noise in the received signal. As a result, the capacity of the target legitimate channel is reduced. If the target channel is SISO and the jammer has a single antenna, the information rate of the target channel under jamming is written as

$$R^{\mathrm{J}} = \log\left(1 + \frac{P_x|h|^2}{\sigma^2 + P_z|h_z|^2}\right) \tag{2.22}$$

where P_x and P_z are the transmission powers of the legitimate transmitter and jammer, respectively, while h represents the channel between the legitimate transceiver and h_z represents the channel from the jammer to the target receiver. In the above SISO case, the jammer does not need to know any information about the channel and has no advanced strategy. It simply increases its transmission power (subject to

its power limit) to further reduce the information rate of the target channel. Therefore, the effect of jamming in the SISO case is power-dominated.

The situation changes significantly when it comes to MIMO. Assume that both the legitimate transceiver and the jammer have multiple antennas. Denote the MIMO channel between the transceiver, i.e., the legitimate channel, as $\mathbf{H}$, and the channel from the jammer to the legitimate receiver as $\mathbf{H}_z$. Further denote the legitimate and jamming signals as $\mathbf{x}$ and $\mathbf{z}$, respectively. The information rate of the target channel under jamming is given as

$$R^{\mathrm{J}} = \log \left| \mathbf{I} + (\mathbf{H}\mathbf{Q}_x\mathbf{H}^{\mathrm{H}})(\sigma^2\mathbf{I} + \mathbf{H}_z\mathbf{Q}_z\mathbf{H}_z^{\mathrm{H}})^{-1} \right| \tag{2.23}$$

where $\mathbf{Q}_x = \mathbb{E}\{\mathbf{x}\mathbf{x}^{\mathrm{H}}\}$ and $\mathbf{Q}_z = \mathbb{E}\{\mathbf{z}\mathbf{z}^{\mathrm{H}}\}$. In this situation, the effect of jamming on the information rate of the legitimate channel depends on the following knowledge

- the knowledge of $\mathbf{H}_z$,
- the knowledge of $\mathbf{H}$,
- the knowledge of $\mathbf{Q}_x$.

In the case that the jammer has none of the above knowledge, the optimal strategy for the jammer is to use $\mathbf{Q}_z = \sigma_z^2\mathbf{I}$ where the constant σ_z^2 is determined by its power limit [46].

If the jammer knows $\mathbf{H}_z$, it can use some basic strategies. It can avoid wasting power by allocating transmission power only in the sub-channels corresponding to the positive eigen-values of $\mathbf{H}_z\mathbf{H}_z^{\mathrm{H}}$. It can also maximize the power of the effective noise, i.e. the noise power plus the jamming power, at the receiver by maximizing $\mathrm{Tr}\{\sigma^2\mathbf{I} + \mathbf{H}_z\mathbf{Q}_z\mathbf{H}_z^{\mathrm{H}}\}$ [45]. Note that maximizing the jamming power at the target receiver is not equivalent to minimizing the information rate of the target channel.

If the jammer also knows $\mathbf{H}$ in addition to $\mathbf{H}_z$, it can further focus its jamming power as if the legitimate channel is just $\mathbf{I}$ and the jamming is applied at the transmitter side of legitimate channel.

The above-mentioned situations are simple since the strategies available to the jammer are limited. Without the knowledge of the target signal or its covariance,

the jammer has no better strategy but maximizing the jamming power at the target receiver. Then, it is logical to consider what would be the result if the jammer knows $\mathbf{Q}_x$? In such a case, the jammer can optimize its transmit strategy to minimize the capacity of the target channel. It is the worst-case jamming for the legitimate transceiver. The resulting information rate of the legitimate transceiver in this case provides a lower bound of the rate under jamming. The above scenario is considered in the first part of Chapter 5 of this thesis.

It should be noted, however, that noise jamming is not the only jamming threat. The worst case of noise jamming described above is not the worst jamming for the target transceiver. If the jammer has the knowledge of the target signal, the jammer is capable of using a more powerful form of jamming, i.e., the correlated jamming.

2.3.2 Correlated jamming

A jammer aims at undermining the signal received at the target receiver. When the jammer knows the target signal, it can use correlated jamming. The basic idea of correlated jamming is that, instead of transmitting random noise signal to lower the SNR at the target receiver, the jammer transmits the minus version of the target signal which neutralizes the received signal at the receiver [16].

Consider the SISO case first. The received signal at the legitimate receiver without jamming is written as

$$y = hx + n. \tag{2.24}$$

If the jammer knows x, it can transmit the signal

$$z = -\alpha \frac{h}{h_z} x \tag{2.25}$$

where $\alpha \in (0, 1]$ is a constant determined by the jammer's power limit P_z.

In the presence of the above jamming signal, the received signal at the target receiver becomes

$$\begin{aligned} y' &= hx + h_z z + n & (2.26)\\ &= (1-\alpha)hx + n. & (2.27) \end{aligned}$$

If the jammer has sufficient power, i.e.,

$$P_z \geq \frac{|h|^2}{|h_z|^2} P_x \tag{2.28}$$

where $P_x \triangleq \mathbb{E}\{|x|^2\}$, it can completely cancel the target signal in (2.27) by setting $\alpha = 1$. Otherwise, the jammer can weaken the received signal by setting α as large as possible subject to its power limit.

For the MIMO case, a jammer can perform correlated jamming if two conditions hold. The first is that the jammer should have at least the same number of antennas as the target receiver. The second is that the jamming channel $\mathbf{H}_z$ has full rank. Assuming that the jammer transmits the signal $\mathbf{z}$, the received signal at the target receiver is given as

$$\mathbf{y}' = \mathbf{H}\mathbf{x} + \mathbf{H}_z\mathbf{z} + \mathbf{n}. \tag{2.29}$$

For correlated jamming, it should hold that

$$\mathbf{H}_z\mathbf{z} = -\alpha\mathbf{H}\mathbf{x}. \tag{2.30}$$

Assuming that $\mathbf{H}_z$ has full rank, the solution for the correlated jamming signal can be found as

$$\mathbf{z} = -\alpha\mathbf{H}_z^{\mathrm{H}}(\mathbf{H}_z\mathbf{H}_z^{\mathrm{H}})^{-1}\mathbf{H}\mathbf{x}. \tag{2.31}$$

Similarly to the SISO case, the power limit of the jammer determines if it can completely cancel the target signal.

Given the above basic model of correlated jamming, the problem of multi-target correlated jamming will be studied in the second part of Chapter 5 of this thesis. With multiple legitimate transceivers, the jammer needs to split its power for jamming the targets. The investigation will reveal how much damage could a single jammer do in a multi-user wireless system through correlated jamming.

2.4 Game theory, NE, and MSNE

In this section, the basics of game theory are introduced and illustrated using examples.

Table 2.1: Matrix representation of a strategic game

	s_1 =a	s_1 =b
s_2 =c	$(2, 2)$	$(3, 1)$
s_2 =d	$(3, 1)$	$(2, 2)$

2.4.1 A brief introduction to game theory

A game can be represented in different forms. A basic form of representation is the strategic form. An M-player game in strategic form has the following three parts [50]

- a set of players/users $\Delta = \{1, 2, \ldots, M\}$
- a set of strategies $\mathcal{S}_i = \{s_i\}$ for player/user $i, \forall i \in \Delta$.
- player/user $i(\forall i \in \Delta)$'s utility $u_i(s_1, \ldots, s_M)$ as a function of strategies of all players/users.

The above game can be briefly represented using a matrix, in which the relation between the players' utilities and their strategies is evident. An example, with $\Delta = \{1, 2\}$, $\mathcal{S}_1 = \{$a, b$\}$ and $\mathcal{S}_2 = \{$c, d$\}$, is given in Table 2.1, where the first and second items in the brackets in Table 2.1 represent the utility for player 1 and player 2, respectively, given their strategies.

Games can be classified into different types from different perspectives [77]. Depending on whether the players make their decisions simultaneously, games can be divided to simultaneous games and sequential games. Depending on whether the players' strategies and utilities are discrete and finite, games can be divided to discrete games and continuous games. Depending on whether the players' utilities always sum up to a constant, games can be divided to zero-sum games and non-zero-sum games. The most important classification of games, however, should be cooperative games and non-cooperative games depending on whether the players can coordinate their strategies for their own benefit. The following review will focus on non-cooperative games as it is adopted in the thesis.

Table 2.2: A game that has one NE

	$s_1 = 0$	$s_1 = 1$
$s_2 = 0$	$(0, 0)$	$(1, -1)$
$s_2 = 1$	$(-1, 1)$	$(1, 1)$

2.4.2 Non-cooperative games and NE

In a non-cooperative game, there is no coordination and the players make decisions independently. An important concept in non-cooperative games is equilibrium. The most well-known example of an equilibrium is the NE. An NE is such combination of strategies, one for each player, that no player can benefit from unilaterally deviating from its current strategy. Mathematically, it can be written as [49]

$$u_i(s_i^\star, s_{-i}^\star) \geq u_i(s_i', s_{-i}^\star),\ \forall s_i' \in \mathcal{S}_i,\ \forall i, \tag{2.32}$$

where $s_i^\star$ represents player i's strategy in the NE, $s_{-i}^\star$ represents the strategies of the set of all players but player i in the NE, and s_i' represents any available strategy but $s_i^\star$ for the ith player. The inequality (2.32) shows that each player's strategy in an NE is the *best response* to the strategies of other players.

Consider the following game as an example. Let $\Delta = \{1, 2\}$, $\mathcal{S}_1 = \{0, 1\}$, $\mathcal{S}_2 = \{0, 1\}$, $u_1 = s_1 - s_2 + s_1 s_2$ and $u_2 = s_2 - s_1 + s_1 s_2$. The utilities of the two players in the game are shown in Table 2.2. It can be seen that the strategy combination $s_1 = 1, s_2 = 1$ is the unique NE in this game. In such situation, there is no uncertainty in the outcome of the game.

However, not every game has an NE. The game shown in Table 2.1 does not have any NE [50]. It is also possible that multiple NEs exist in one game. Consider the following example. Two pairs of transceivers independently choose from one of two wireless channels to perform communication. The transmission power of the first and second transceiver pair are P_1 and P_2, respectively. The channel gain of the two channels are both h_1 for the first transceiver pair and both h_2 for the second transceiver pair. The interference channels (from the transmitter of one transceiver pair to the receiver of the other pair) have the same channel gain c. The noise is σ^2 at both receivers. Model the transceivers as players, their choices of channel as

Table 2.3: A channel selection game that has two NEs

	s_1 =c 1	s_1=c 2
s_2=c 1	(R_1', R_2')	(R_1, R_2)
s_2=c 2	(R_1, R_2)	(R_1', R_2')

strategies, and their information rates as utilities. Let $\mathcal{S}_1 = \mathcal{S}_2 =$\{channel 1(briefly, c 1), channel 2(briefly, c 2)\}. The game is shown in Table 2.3, where

$$R_1' = \log\left(1 + \frac{P_1|h_1|^2}{\sigma^2 + P_2|c|^2}\right) \tag{2.33}$$

$$R_1 = \log\left(1 + \frac{P_1|h_1|^2}{\sigma^2}\right) \tag{2.34}$$

$$R_2' = \log\left(1 + \frac{P_2|h_2|^2}{\sigma^2 + P_1|c|^2}\right) \tag{2.35}$$

$$R_2 = \log\left(1 + \frac{P_2|h_2|^2}{\sigma^2}\right). \tag{2.36}$$

Since $R_1' < R_1$ and $R_2' < R_2$, it can be seen that the above game has two Nash equilibria (NEs), i.e., $\{s_1 =$c 1, s_2=c 2$\}$ and $\{s_1 =$c 2, s_2=c 1$\}$.

2.4.3 Mixed strategy and MSNE

In all the games considered in Sections 2.4.1 and 2.4.2, the players use *pure strategies* only. Pure strategies represent definite choice with no uncertainty for the players. For example, player 1 must choose either c 1 or c 2 with probability 1 in the game represented by Table 2.3 if it uses pure strategy. However, if the players are able to assign probabilities to their pure strategies as they wish and thereby select the pure strategies with randomness, they can extend their strategy space to mixed strategies.

A *mixed strategy* of a player is a probability distribution over its pure strategies in $\mathcal{S}_i$, denoted as $p_i(\mathcal{S}_i)$, with the probability assigned to s_i being $p_i(s_i)$. For example, a mixed strategy for player 1 in the game represented by Table 2.3 is to choose c 1 with the probability of 0.4 and c 2 with the probability of 0.6. As there are infinitely many distributions over a set of pure strategies, the number of mixed strategies for each player is also infinite. In a game with mixed strategies, the utility

of player i is given as

$$U_i(p_i(\mathcal{S}_i), p_{-i}(\mathcal{S}_{-i})) = \sum_{(s_1,\ldots,s_M)\in\mathcal{S}} u_i(s_i, s_{-i}) \prod_{k\in\Delta} p_k(s_k) \tag{2.37}$$

where $p_{-i}(\mathcal{S}_{-i})$ is the combination of the probability distributions of all players but player i and $\mathcal{S}$ is the Cartesian product of $\mathcal{S}_1$,..., $\mathcal{S}_M$.

An MSNE is such combination of probability distributions, one from each user, that

$$U_i(p_i^\star(\mathcal{S}_i), p_{-i}^\star(\mathcal{S}_{-i})) \geq U_i(p_i'(\mathcal{S}_i), p_{-i}^\star(\mathcal{S}_{-i})),\ \forall p_i'(\mathcal{S}_i),\ \forall i, \tag{2.38}$$

where $p_i^\star(\mathcal{S}_i)$ represents the distribution of player i in the MSNE, $p_{-i}^\star(\mathcal{S}_{-i})$ represents the combination of the distributions of all players but player i in the MSNE, and $p_i'(\mathcal{S}_i)$ represents any distribution valid for the ith player but $p_i^\star(\mathcal{S}_i)$. Denote the set of player i's pure strategies with positive possibilities in $p_i^\star(\mathcal{S}_i)$ in the MSNE as $\mathcal{S}_i^+$. Note that all the pure strategies in $\mathcal{S}_i^+$ must lead to the same utility for player i given the distributions of other players, i.e.,

$$U_i(g(s_i), p_{-i}^\star(\mathcal{S}_{-i})) = U_i(p_i^\star(\mathcal{S}_i), p_{-i}^\star(\mathcal{S}_{-i})),\ \forall s_i \in \mathcal{S}_i^+,\ \forall i \tag{2.39}$$

where $g(x)$ is the distribution over $\mathcal{S}_i$ defined as

$$g(x) = \begin{cases} 0 & x \neq s_i \\ 1 & x = s_i. \end{cases} \tag{2.40}$$

The reason is that, otherwise, player i can increase its utility by assigning 0 probability to the pure strategy in $\mathcal{S}_i^+$ that leads to the minimum utility for it while increasing the probabilities of other pure strategies in $\mathcal{S}_i^+$, which contradicts the definition of MSNE.

The admission of mixed strategy and MSNE marks the beginning of the modern game theory [78]. The significance of mixed strategy lies in the fact that MSNE may exist with mixed strategies in a significant amount of games that do not have NE with pure strategies. Consider the example in Table 2.1 which has no NE in pure strategies. Assume that player 1 chooses $s_1 = a$ and $s_1 = b$ with probabilities p_a and p_b, respectively, while player 2 chooses $s_2 = c$ and $s_2 = d$ with probabilities p_c

and p_d, respectively. According to the table, the utilities of the players are obtained as

$$U_1 = p_a(2p_c + 3p_d) + p_b(3p_c + 2p_d) \tag{2.41}$$

$$U_2 = p_c(2p_a + p_b) + p_d(p_a + 2p_b). \tag{2.42}$$

Using the property in (2.39), we have

$$2p_c + 3p_d = 3p_c + 2p_d \tag{2.43}$$

$$2p_a + p_b = p_a + 2p_b. \tag{2.44}$$

Considering that $p_c + p_d = p_a + p_b = 1$, the above equations have the solution $p_c = p_d = p_a = p_b = 0.5$. Therefore, there exists a unique MSNE in this game that has no NE in pure strategy. The utilities for the two players in the MSNE are $U_1 = 2.5$ and $U_2 = 1.5$.

In the proposed research in Chapter 6, the problem of resource allocation in multi-user wireless communications is investigated using games with mixed strategies. The proposed game model is, however, much more complicated than the one described above in the following two aspects. First, the strategy of each player in the game will be represented by a continuous variable. Second, the utilities of the players are not simply given but defined by continuous functions of the player's strategies. With the above modeling, the proposed research aims at studying the existence and uniqueness of MSNEs for the resource allocation game in the considered scenario of two-user wireless communications.

$\sim$

Chapter 3

Relay-Oriented MIMO DF TWR: Maximizing Spectral Efficiency with Minimum Power

In this chapter, the MIMO DF TWR is investigated in the *relay optimization scenario,* in which the relay optimizes its own power allocation to achieve sum-rate maximization with minimum power consumption given the power allocation of the source nodes. The objective of this chapter is to find the optimal power allocation strategy of the relay in the relay optimization scenario. [1]

3.1 System model

Consider a TWR with two source nodes and one relay, where source node $i\,(i = 1, 2)$ and the relay have n_i and n_{r} antennas, respectively. In the MA phase, source node i transmits signal $\mathbf{W}_i\mathbf{s}_i$ to the relay. Here $\mathbf{W}_i$ is the precoding matrix of source node i and $\mathbf{s}_i$ is the complex Gaussian information symbol vector of source node i. The elements of $\mathbf{s}_i, \forall i$ are independent and identically distributed with zero mean and unit variance. The channels from source node i to the relay and from the relay to source node i are denoted as $\mathbf{H}_{i\mathrm{r}}$ and $\mathbf{H}_{\mathrm{r}i}$, respectively. Receiver channel state information is assumed to be known at both the relay and the source nodes, i.e., source node i knows $\mathbf{H}_{\mathrm{r}i}$ and the relay knows $\mathbf{H}_{i\mathrm{r}}, \forall i$. It is also assumed that the relay knows $\mathbf{H}_{\mathrm{r}i}, \forall i$ by using either channel reciprocity or channel feedback. The

[1] *A version of this chapter has been published in IEEE Trans. Signal Process., 61: 3563-3577 (2013).*

received signal at the relay in the MA phase is

$$\mathbf{y}_{\mathrm{r}} = \mathbf{H}_{1\mathrm{r}}\mathbf{W}_1\mathbf{s}_1 + \mathbf{H}_{2\mathrm{r}}\mathbf{W}_2\mathbf{s}_2 + \mathbf{n}_{\mathrm{r}} \tag{3.1}$$

where $\mathbf{n}_{\mathrm{r}}$ is the noise at the relay with covariance matrix $\sigma_{\mathrm{r}}^2\mathbf{I}$. The maximum transmission power of source node i is limited to $P_i^{\max}$. Define the transmit covariance matrices $\mathbf{D}_i \triangleq \mathbf{W}_i\mathbf{W}_i^{\mathrm{H}}, \forall i$, and $\mathbf{D} \triangleq [\mathbf{D}_1, \mathbf{D}_2]$. Then the sum-rate of the MA phase is bounded by [67]

$$R^{\mathrm{ma}}(\mathbf{D}) = \log\left|\mathbf{I}+(\mathbf{H}_{1\mathrm{r}}\mathbf{D}_1\mathbf{H}_{1\mathrm{r}}^{\mathrm{H}}+\mathbf{H}_{2\mathrm{r}}\mathbf{D}_2\mathbf{H}_{2\mathrm{r}}^{\mathrm{H}})(\sigma_{\mathrm{r}}^2)^{-1}\right|. \tag{3.2}$$

In the BC phase, the relay decodes $\mathbf{s}_1$ and $\mathbf{s}_2$ from the received signal, re-encodes messages using superposition coding and transmits the signal

$$\mathbf{x}_{\mathrm{r}} = \mathbf{T}_{\mathrm{r}2}\mathbf{s}_1 + \mathbf{T}_{\mathrm{r}1}\mathbf{s}_2 \tag{3.3}$$

where $\mathbf{T}_{\mathrm{r}i}$ is the $n_{\mathrm{r}} \times n_j$ relay precoding matrix for relaying the signal from source node j to source node i.[2] The maximum transmission power of the relay is limited to $P_{\mathrm{r}}^{\max}$. Note that in addition to the above superposition coding, the Exclusive-OR (XOR) based network coding is also used at the relay in the literature [71, 79, 80]. While XOR-based network coding may achieve a better performance than superposition coding, it relies on the symmetry of the traffic from the two source nodes. The asymmetry in the traffic in the two directions can lead to a significant degradation in the performance of XOR in TWR [79], [80]. As the general case of TWR is considered here and there is no guarantee of traffic symmetry, the approach of symbol-level superposition is assumed at the relay as it is considered in [7] and [33]. Moreover, for the MIMO case that we are considering, the superposition scheme can take advantage of the MIMO channels. In the superposition scheme, the relay uses separate beamformers for the signals towards two directions, which guarantees that each transmitted signal is optimal (subject to the transmission power constraints) given its MIMO channel. This cannot be achieved if the relay uses XOR-based network coding.

The received signal at source node i can be expressed as

$$\mathbf{y}_i' = \mathbf{H}_{\mathrm{r}i}\mathbf{x}_{\mathrm{r}} + \mathbf{n}_i \tag{3.4}$$

[2]It is assumed as default throughout this chapter that the user indices i and j satisfy $i \neq j$.

where $\mathbf{n}_i$ is the noise at source node i with covariance matrix $\sigma_i^2\mathbf{I}$. With the knowledge of $\mathbf{H}_{\mathrm{r}i}$ and $\mathbf{T}_{\mathrm{r}j}$, source node i subtracts the self-interference $\mathbf{H}_{\mathrm{r}i}\mathbf{T}_{\mathrm{r}j}\mathbf{s}_i$ from the received signal and the equivalent received signal at source node i is

$$\mathbf{y}_i = \mathbf{H}_{\mathrm{r}i}\mathbf{T}_{\mathrm{r}i}\mathbf{s}_j + \mathbf{n}_i. \tag{3.5}$$

Define $\mathbf{B}_i \triangleq \mathbf{T}_{\mathrm{r}i}\mathbf{T}_{\mathrm{r}i}^{\mathrm{H}}, \forall i$ and let $\mathbf{B} \triangleq [\mathbf{B}_1, \mathbf{B}_2]$. The sum-rate of the considered DF TWR can be written as [7], [33], [71]

$$R^{\mathrm{tw}}(\mathbf{B}, \mathbf{D}) = \frac{1}{2}\min\{R^{\mathrm{ma}}(\mathbf{D}), R(\mathbf{B}, \mathbf{D})\} \tag{3.6}$$

where

$$R(\mathbf{B}, \mathbf{D}) = \min\{\hat{R}_{\mathrm{r}1}(\mathbf{B}_1), \bar{R}_{2\mathrm{r}}(\mathbf{D}_2)\} + \min\{\hat{R}_{\mathrm{r}2}(\mathbf{B}_2), \bar{R}_{1\mathrm{r}}(\mathbf{D}_1)\} \tag{3.7}$$

in which

$$\bar{R}_{j\mathrm{r}}(\mathbf{D}_j) = \log|\mathbf{I} + (\mathbf{H}_{j\mathrm{r}}\mathbf{D}_j\mathbf{H}_{j\mathrm{r}}^{\mathrm{H}})(\sigma_{\mathrm{r}}^2)^{-1}| \tag{3.8}$$

and

$$\hat{R}_{\mathrm{r}i}(\mathbf{B}_i) = \log|\mathbf{I} + (\mathbf{H}_{\mathrm{r}i}\mathbf{B}_i\mathbf{H}_{\mathrm{r}i}^{\mathrm{H}})(\sigma_i^2)^{-1}|. \tag{3.9}$$

For brevity of presentation, we define the following sum-rate of the BC phase

$$R^{\mathrm{bc}}(\mathbf{B}) = \hat{R}_{\mathrm{r}1}(\mathbf{B}_1) + \hat{R}_{\mathrm{r}2}(\mathbf{B}_2). \tag{3.10}$$

For the relay optimization scenario considered here, the relay maximizes the sum-rate in (3.6) using minimum transmission power given the power allocation strategies of the source nodes.[3] Since the relay needs to know $\mathbf{W}_1$ and $\mathbf{W}_2$ for decoding $\mathbf{s}_1$ and $\mathbf{s}_2$, respectively, as well as for designing $\mathbf{T}_{\mathrm{r}1}$ and $\mathbf{T}_{\mathrm{r}2}$, the source nodes should send their respective precoding matrices to the relay after they decide their transmit strategies. Similarly, the relay should also send $\mathbf{T}_{\mathrm{r}1}$ and $\mathbf{T}_{\mathrm{r}2}$ to both source nodes.

Given the above system model, we next solve the relay optimization problem.

[3]The term 'sum-rate' by default means $R^{\mathrm{tw}}(\mathbf{B}, \mathbf{D})$ when we do not specify it to be the sum-rate of the BC or MA phase.

3.2 Relay optimization

In the relay optimization scenario, the relay and the source nodes do not coordinate in choosing their respective power allocation strategies. Instead, the relay aims at maximizing $R^{\mathrm{tw}}(\mathbf{B}, \mathbf{D})$ in (3.6) with minimum power consumption after the source nodes decide their strategies and inform the relay.

Denote the power allocation that the source nodes decide to use as $\mathbf{D}^0 \triangleq [\mathbf{D}_1^0, \mathbf{D}_2^0]$.[4] For maximizing the sum-rate given $\mathbf{D}^0$, the relay solves the following optimization problem[5]

$$\max_{\mathbf{B}} \quad R^{\mathrm{tw}}(\mathbf{B}, \mathbf{D}^0) \tag{3.11a}$$

$$\mathbf{s.t.} \quad \mathrm{Tr}\{\mathbf{B}_1 + \mathbf{B}_2\} \leq P_{\mathrm{r}}^{\max}. \tag{3.11b}$$

The problem (3.11) is convex. However, in order to find the optimal $\mathbf{B}$ with minimum $\mathrm{Tr}\{\mathbf{B}_1 + \mathbf{B}_2\}$ among all possible $\mathbf{B}$'s that achieve the maximum of the objective function in (3.11), extra constraints need to be considered. Two necessary constraints are given below

$$\hat{R}_{\mathrm{r}i}(\mathbf{B}_i) \leq \bar{R}_{j\mathrm{r}}(\mathbf{D}_j^0), \forall i \tag{3.12a}$$

$$R(\mathbf{B}, \mathbf{D}^0) \leq R^{\mathrm{ma}}(\mathbf{D}^0). \tag{3.12b}$$

The considered relay optimization problem (3.11) with additional necessary constraints (3.12a) and (3.12b) becomes nonconvex. The above necessary constraints are introduced here to show that the considered relay optimization problem is nonconvex. For a sufficient and necessary condition for a power allocation strategy to be optimal in terms of maximizing sum-rate with minimum power consumption, please see Theorem 3.2 later in this section.

The constraint (3.12a) is necessary because, given $\mathbf{D}^0$, due to the expression of $R(\mathbf{B}, \mathbf{D})$ in (3.7), the power consumption of the relay can be reduced while

[4] The source nodes may determine their power allocation strategies using different objectives. Note that different source node power allocation strategies lead to different solutions of the relay optimization problem. However, the approach adopted in this chapter for solving the relay optimization problem is valid for arbitrary source node power allocation.

[5] The positive semi-definite (PSD) constraints $\mathbf{D}_i \succeq 0, \forall i$ and $\mathbf{B}_i \succeq 0, \forall i$ are assumed as default and omitted for brevity in all formations of optimization problems in this chapter.

the sum-rate $R^{\mathrm{tw}}(\mathbf{B},\mathbf{D})$ in (3.6) can be kept unchanged by reducing $\mathrm{Tr}\{\mathbf{B}_i\}$ if $\hat{R}_{\mathrm{r}i}(\mathbf{B}_i) > \bar{R}_{j\mathrm{r}}(\mathbf{D}_j^0)$. Note that (3.12a) is not necessarily satisfied with equality at optimality. In fact, it can be shown using subsequent results in Section 3.2.2 that (3.12a) should be satisfied with inequality for at least one i at optimality. It can also be shown that (3.12a) can be satisfied with inequalities for both i's at optimality even if the relay has an unlimited power budget. We stress that (3.12a) is not sufficient for obtaining the optimal solution. Other constraints are also needed including (3.12b). The constraint (3.12b) is also necessary because, given $\mathbf{D}^0$, if (3.12b) is not satisfied, then the power consumption of the relay can be reduced while the sum-rate $R^{\mathrm{tw}}(\mathbf{B},\mathbf{D}^0)$ can be kept unchanged by decreasing $R(\mathbf{B},\mathbf{D}^0)$ so that $R(\mathbf{B},\mathbf{D}^0) = R^{\mathrm{ma}}(\mathbf{D}^0)$.

The constraints in (3.12) make the considered problem nonconvex. The objective in this section is to find an efficient method of deriving the optimal power allocation of the relay in the considered scenario of relay optimization. It is straightforward to see that the power allocation of the relay should be based on waterfilling for relaying the signal in either direction regardless of how the relay distributes its power in the two directions. This is due to the fact that the BC phase is interference free since both source nodes are able to subtract their self-interference. If the objective were to maximize $R^{\mathrm{bc}}(\mathbf{B})$ instead of $R^{\mathrm{tw}}(\mathbf{B},\mathbf{D}^0)$, the optimal strategy of the relay could be found via a simple search. Indeed, in that case, we could find the optimal power allocation of the relay and consequently the optimal $\mathbf{B}$ by searching for the optimal proportion that the relay distributes its power in the two directions. However, such approach is infeasible for the considered problem. The reason is that first of all it is unknown what is the total power that the relay uses in the optimal solution. As power efficiency is also considered, the relay may not use full power in its optimal strategy. Moreover, from the expression of $R^{\mathrm{tw}}(\mathbf{B},\mathbf{D})$ in (3.6), it can be seen that the maximum achievable $R^{\mathrm{tw}}(\mathbf{B},\mathbf{D}^0)$ also depends on $\bar{R}_{1\mathrm{r}}(\mathbf{D}_1^0)$, $\bar{R}_{2\mathrm{r}}(\mathbf{D}_2^0)$, and $R^{\mathrm{ma}}(\mathbf{D}^0)$. Due to this dependence, the two constraints in (3.12) are necessary for the considered problem of sum-rate maximization with minimum power consumption. However, these two constraints are implicit in the sense that they are constraints on the rates instead of on the power allocation of the relay. Such con-

straints offer no insight in finding the optimal $\mathbf{B}$. In order to transform the above mentioned dependence of $R^{\mathrm{tw}}(\mathbf{B}, \mathbf{D}^0)$ on $\bar{R}_{1\mathrm{r}}(\mathbf{D}_1^0)$, $\bar{R}_{2\mathrm{r}}(\mathbf{D}_2^0)$, and $R^{\mathrm{ma}}(\mathbf{D}^0)$ into an explicit form, and to discover the insight behind the constraints in (3.12), we next propose the idea of relative water-levels and develop a method based on this idea.

3.2.1 Relative water-levels

Denote the rank of $\mathbf{H}_{\mathrm{r}i}$ as $r_{\mathrm{r}i}$ and the singular value decomposition (SVD) of $\mathbf{H}_{\mathrm{r}i}$ as $\mathbf{U}_{\mathrm{r}i}\mathbf{\Omega}_{\mathrm{r}i}\mathbf{V}_{\mathrm{r}i}^{\mathrm{H}}$. Assume that the first $r_{\mathrm{r}i}$ diagonal elements of $\mathbf{\Omega}_{\mathrm{r}i}$ are non-zero, sorted in descending order and denoted as $\omega_{\mathrm{r}i}(1), \ldots, \omega_{\mathrm{r}i}(r_{\mathrm{r}i})$, while the last $\min\{n_i, n_{\mathrm{r}}\} - r_{\mathrm{r}i}$ diagonal elements are zeros. Define $\mathcal{I}_i \triangleq \{1, \ldots, r_{\mathrm{r}i}\}, \forall i$ and $\alpha_i(k) \triangleq |\omega_{\mathrm{r}i}(k)|^2/\sigma_i^2$, $\forall k \in \mathcal{I}_i, \forall i$. For a given $\mathbf{D} = [\mathbf{D}_1, \mathbf{D}_2]$, define $\mu_1(\mathbf{D}_1)$, $\mu_2(\mathbf{D}_2)$, and $\mu_{\mathrm{ma}}(\mathbf{D})$ such that

$$\sum_{k\in\mathcal{I}_2} \log\left(1 + \big(\frac{1}{\mu_1(\mathbf{D}_1)}\alpha_2(k) - 1\big)^+\right) = \bar{R}_{1\mathrm{r}}(\mathbf{D}_1) \tag{3.13a}$$

$$\sum_{k\in\mathcal{I}_1} \log\left(1 + \big(\frac{1}{\mu_2(\mathbf{D}_2)}\alpha_1(k) - 1\big)^+\right) = \bar{R}_{2\mathrm{r}}(\mathbf{D}_2) \tag{3.13b}$$

$$\sum_i \sum_{k\in\mathcal{I}_i} \log\left(1 + \big(\frac{1}{\mu_{\mathrm{ma}}(\mathbf{D})}\alpha_i(k) - 1\big)^+\right) = R^{\mathrm{ma}}(\mathbf{D}) \tag{3.13c}$$

where $(\cdot)^+$ stands for the projection to the positive orthant. The physical meaning of $\mu_i(\mathbf{D}_i)$ is that if waterfilling is performed on $\omega_{\mathrm{r}j}(k)$'s, $\forall k \in \mathcal{I}_j$ using the water-level $1/\mu_i(\mathbf{D}_i)$, then the information rate of the transmission from the relay to source node j using the resulting waterfilling-based power allocation achieves precisely $\bar{R}_{i\mathrm{r}}(\mathbf{D}_i)$. The physical meaning of $\mu_{\mathrm{ma}}(\mathbf{D})$ is that if waterfilling is performed on $\omega_{\mathrm{r}i}(k)$'s, $\forall k \in \mathcal{I}_i, \forall i$ using the water-level $1/\mu_{\mathrm{ma}}(\mathbf{D})$, then the sum-rate of the transmission from the relay to the two source nodes using the resulting waterfilling-based power allocation achieves precisely $R^{\mathrm{ma}}(\mathbf{D})$. Note that $1/\mu_i(\mathbf{D}_i), \forall i$ and $1/\mu_{\mathrm{m}}(\mathbf{D})$ are not the actual water-levels for the MA or the BC phases. They are just relative water-levels introduced to transform and simplify the constraints in (3.12). Denote the actual water-levels used by the relay for relaying the signal from source node j to source node i as $1/\lambda_i, \forall i$. With water-level $1/\lambda_i$, $\mathbf{B}_i$ can be given as $\mathbf{B}_i = \mathbf{V}_{\mathrm{r}i}\mathbf{P}_{\mathrm{r}i}(\lambda_i)\mathbf{V}_{\mathrm{r}i}^{\mathrm{H}}$ where

$$\mathbf{P}_{\mathrm{r}i}(\lambda_i) = \begin{bmatrix} \left(\frac{1}{\lambda_i} - \frac{1}{\alpha_i(1)}\right)^+ & & \\ & \ddots & \\ & & \left(\frac{1}{\lambda_i} - \frac{1}{\alpha_i(\mathrm{r}_{\mathrm{r}i})}\right)^+ \\ & & \mathbf{0}_{n_{\mathrm{r}}-r_{\mathrm{r}i}}, \end{bmatrix} \forall i \tag{3.14}$$

in which $\mathbf{0}_{n_{\mathrm{r}}-r_{\mathrm{r}i}}$ stands for all-zero matrix of size $(n_{\mathrm{r}}-r_{\mathrm{r}i}) \times (n_{\mathrm{r}}-r_{\mathrm{r}i})$. The power allocated on $\omega_{\mathrm{r}i}(k)$ is $p_{\mathrm{r}i}(k) = \left(1/\lambda_i - 1/\alpha_i(k)\right)^+, \forall k \in \mathcal{I}_i, \forall i$. The resulting rate $\hat{R}_{\mathrm{r}i}(\mathbf{B}_i)$ is given by $\sum\limits_{k\in\mathcal{I}_i} \log 1 + \left(\alpha_i(k)/\lambda_i - 1\right)^+$. Using $\mu_1(\mathbf{D}_1)$, $\mu_2(\mathbf{D}_2)$, and $\mu_{\mathrm{ma}}(\mathbf{D})$, the constraints in (3.12a) can be rewritten as

$$\lambda_i \geq \mu_j(\mathbf{D}_j^0), \forall i \tag{3.15a}$$

$$\sum_i \sum_{k\in\mathcal{I}_i} \log\left(1 + \left(\frac{1}{\lambda_i}\alpha_i(k) - 1\right)^+\right) \leq \sum_i \sum_{k\in\mathcal{I}_i} \log\left(1 + \left(\frac{1}{\mu_{\mathrm{ma}}(\mathbf{D}^0)}\alpha_i(k) - 1\right)^+\right). \tag{3.15b}$$

Given (4.2a) and (4.2b), it is easy to see that (3.12a) is equivalent to (3.15a). Moreover, the equivalence between (3.12b) and (3.15b) can be explained as follows. Given $\mathbf{D}^0$ and (3.12b), $R^{\mathrm{tw}}(\mathbf{B}, \mathbf{D}^0)$ in (3.11a) becomes $R(\mathbf{B}, \mathbf{D}^0)/2$. Given (3.12a), or equivalently (3.15a), $R(\mathbf{B}, \mathbf{D})$ in (3.7) with $\mathbf{D} = \mathbf{D}^0$ becomes $\hat{R}_{\mathrm{r}1}(\mathbf{B}_1) + \hat{R}_{\mathrm{r}2}(\mathbf{B}_2)$. Then, substituting the left-hand side of (3.12b) with $\hat{R}_{\mathrm{r}1}(\mathbf{B}_1) + \hat{R}_{\mathrm{r}2}(\mathbf{B}_2)$, i.e., $R^{\mathrm{bc}}(\mathbf{B})$ in (3.10), and using (4.2c), the constraint (3.15b) is obtained.

The procedure for the relay optimization can be summarized in the following three steps:

1. Obtain $\mu_1(\mathbf{D}_1^0)$, $\mu_2(\mathbf{D}_2^0)$, and $\mu_{\mathrm{ma}}(\mathbf{D}^0)$ from $\mathbf{D}^0$;

2. Determine the optimal λ_i;

3. Obtain $\mathbf{P}_{\mathrm{r}i}(\lambda_i)$ and $\mathbf{B}_i$ from λ_i.

The first and the third steps are straightforward given the definitions (4.2a)-(4.2c) and (3.14). Therefore, finding the optimal $\lambda_i, \forall i$ in the second step is the essential part to be dealt with later in this section.

From hereon, $\mu_1(\mathbf{D}_1)$, $\mu_2(\mathbf{D}_2)$, and $\mu_{\mathrm{ma}}(\mathbf{D})$ are denoted as μ_1, μ_2 and μ_{ma}, respectively, for brevity. The same markers/superscripts on $\mathbf{D}_i$ and/or $\mathbf{D}$ are used on μ_i and/or μ_{ma} to represent the connection. For example, $\mu_i(\mathbf{D}_i^0)$ and $\mu_{\mathrm{ma}}(\tilde{\mathbf{D}})$ are briefly denoted as μ_i^0 and $\tilde{\mu}_{\mathrm{ma}}$, respectively. The rate $\hat{R}_{\mathrm{r}i}(\mathbf{B}_i)$ obtained using water-level $1/\lambda_i$ is also denoted as $\hat{R}_{\mathrm{r}i}(\lambda_i)$.

3.2.2 Algorithm for relay optimization

Using the relative water-levels $\mu_i, \forall i$ and $\mu_{\rm ma}$, we can now develop the algorithm for relay optimization. In order to do that, the following lemmas are presented.

Lemma 3.1: $1/\mu_{\rm ma} < \max\{1/\mu_1, 1/\mu_2\}$.

Proof: The proof for Lemma 3.1 is straightforward. Using (4.2a)-(4.2c), it can be seen that $R^{\rm ma}(\mathbf{D}) \geq \sum_i \bar{R}_{i\rm r}(\mathbf{D}_i)$ if $1/\mu_{\rm ma} \geq \max\{1/\mu_1, 1/\mu_2\}$. However, given the definitions in (3.2) and (3.8), it can be seen that $R^{\rm ma}(\mathbf{D}) \geq \sum_i \bar{R}_{i\rm r}(\mathbf{D}_i)$ is impossible [67]. Therefore, $1/\mu_{\rm ma} < \max\{1/\mu_1, 1/\mu_2\}$. ■

Lemma 3.2: Assume that there exist $\{\lambda_i, \lambda_j\}$ and $\{\lambda_i', \lambda_j'\}$ such that $\lambda_i' < \lambda_i \leq \lambda_j < \lambda_j'$. If $\sum_l {\rm Tr}\{\mathbf{P}_{{\rm r}l}(\lambda_l)\} = \sum_l {\rm Tr}\{\mathbf{P}_{{\rm r}l}(\lambda_l')\}$, then $\sum_l \hat{R}_{{\rm r}l}(\lambda_l) > \sum_l \hat{R}_{{\rm r}l}(\lambda_l')$ as long as $1/\lambda_j > \min_{k\in\mathcal{I}_j}\{1/\alpha_j(k)\}$.

Proof: See Subsection A.1 in Appendix. ■

Essentially, Lemma 3.2 states that, for any given $\{\lambda_1, \lambda_2\}$ such that $1/\lambda_2 > \min_{k\in\mathcal{I}_2}\{1/\alpha_2(k)\}$ assuming $\lambda_1 \leq \lambda_2$, decreasing $\min\{\lambda_1, \lambda_2\}$ and increasing $\max\{\lambda_1, \lambda_2\}$ while fixing the total power consumption leads to a smaller BC phase sum-rate than that achieved by using $\{\lambda_1, \lambda_2\}$.

Lemma 3.3: Assume that there exist $\{\lambda_i, \lambda_j\}$ and $\{\lambda_i', \lambda_j'\}$ such that $\lambda_i < \lambda_j$, $\lambda_i' > \lambda_i$ and $\lambda_j' > \lambda_j$, and

$$\hat{R}_{{\rm r}i}(\lambda_i') + \hat{R}_{{\rm r}j}(\lambda_j) = \hat{R}_{{\rm r}i}(\lambda_i) + \hat{R}_{{\rm r}j}(\lambda_j') \tag{3.16}$$

then as long as $\lambda_i' \leq \lambda_j$, it holds true that

$${\rm Tr}\{\mathbf{P}_{{\rm r}i}(\lambda_i')\} + {\rm Tr}\{\mathbf{P}_{{\rm r}j}(\lambda_j)\} < {\rm Tr}\{\mathbf{P}_{{\rm r}i}(\lambda_i)\} + {\rm Tr}\{\mathbf{P}_{{\rm r}j}(\lambda_j')\}. \tag{3.17}$$

Proof: See Subsection A.2 in Appendix. ■

In other words, Lemma 3.3 states that, for any given $\{\lambda_1, \lambda_2\}$, decreasing $\min\{\lambda_1, \lambda_2\}$ and increasing $\max\{\lambda_1, \lambda_2\}$ such that the BC phase sum-rate is unchanged, the power consumption increases.

Theorem 3.1: The optimal solution of the considered relay optimization prob-

lem always satisfies the following properties

$$\min\left\{\frac{1}{\lambda_1},\frac{1}{\lambda_2}\right\}=\min\left\{\frac{1}{\mu_1^0},\frac{1}{\mu_2^0}\right\} \quad \text{if} \quad \lambda_1\neq\lambda_2 \tag{3.18a}$$

$$\frac{1}{\lambda_1}=\frac{1}{\lambda_2}=\min\left\{\frac{1}{\mu_{\rm ma}^0},\frac{1}{\lambda^0}\right\} \quad \text{if} \quad \lambda_1=\lambda_2 \tag{3.18b}$$

where $1/\lambda^0$ is the water-level obtained by waterfilling $P_{\rm r}^{\max}$ on $\omega_{{\rm r}i}(k),\forall k\in\mathcal{I}_i,\forall i$.

Proof: See Subsection A.3 in Appendix. ■

According to the proof of Theorem 3.1, it can be seen that $\lambda_1\neq\lambda_2$ at optimality and consequently the equation in (3.18a) holds when both of the following two conditions are satisfied: (i) the relay has sufficient power, i.e., $1/\lambda^0>\min\{1/\mu_1^0,1/\mu_2^0\}$, and (ii) there is asymmetry between μ_1^0 and μ_2^0, i.e., $\min\{1/\mu_1^0,1/\mu_2^0\}<1/\mu_{\rm ma}^0<\max\{1/\mu_1^0,1/\mu_2^0\}$. If either of the above two conditions is not satisfied, $\lambda_1=\lambda_2$ at optimality and consequently the equation in (3.18b) holds.

Theorem 3.2: In the relay optimization scenario, the conditions (3.15a), (3.15b), (3.18a), and (3.18b) are sufficient and necessary to determine the optimal $\{\lambda_1,\lambda_2\}$ with minimum power consumption among all $\{\lambda_1,\lambda_2\}$'s that maximize the sum-rate $R^{\rm tw}(\mathbf{B},\mathbf{D}^0)$.

Proof: See Subsection A.4 in Appendix. ■

It should be noted that the power constraint (3.11b) is not always tight at optimality due to the constraints in (3.15a), (3.15b) (or equivalently (3.12a), (3.12b)), (3.18a), and (3.18b). Each of (3.15a), (3.15b), (3.18a), and (3.18b) may refrain the relay from using its full power at optimality. The reason can be found from the proofs of Theorems 1 and 2. Specifically, (3.15a) and (3.18a) make sure that there is no superfluous power spent for relaying the signal in each direction while (3.15b) and (3.18b) guarantee that the power consumption of the relay cannot be further reduced without reducing the sum-rate.

Based on the above results in Theorems 1 and 2, the algorithm summarized in Table 3.1 is proposed to find the optimal relay power allocation for the relay optimization problem. In order to make sure that the sum-rate is maximized while no power is wasted, the algorithm balances $\hat{R}_{\rm r1}(\mathbf{B}_1)$ and $\hat{R}_{\rm r2}(\mathbf{B}_2)$ via adjusting λ_1 and λ_2 according to $\bar{R}_{\rm 1r}(\mathbf{D}_1^0)$, $\bar{R}_{\rm 2r}(\mathbf{D}_2^0)$, and $R^{\rm ma}(\mathbf{D}^0)$. The algorithm uses relative water-levels, which are not explicitly related to corresponding rates.

Table 3.1: The algorithm for relay optimization.

1. Initial waterfilling: allocate $P_{\rm r}^{\max}$ on $\omega_{{\rm r}i}(k), \forall k \in \mathcal{I}_i, \forall i$ using waterfilling. Denote the initial water level as $1/\lambda^0$. Set $1/\lambda_1 = 1/\lambda_2 = 1/\lambda^0$. The power allocated on $\omega_{{\rm r}i}(k)$ is $p_{{\rm r}i}(k) = \big(1/\lambda_i - 1/\alpha_i(k)\big)^+, \forall k \in \mathcal{I}_i, \forall i$.
2. Check if $1/\lambda_i \leq 1/\mu_j^0$ for both $i = 1, 2$. If yes, proceed to Step 6. Otherwise, assume that $1/\lambda_1 > 1/\mu_2^0$, proceed to Step 3.
3. Set $\lambda_1 = \mu_2^0$. Check if $1/\lambda_2 > 1/\mu_1^0$. If not, proceed to Step 4. Otherwise, proceed to Step 5.
4. Calculate $P_{\rm r}' = P_{\rm r}^{\max} - \sum_{k \in \mathcal{I}_1} p_{\rm r1}(k)$. Allocate $P_{\rm r}'$ on $\omega_{\rm r2}(k)$'s,$\forall k \in \mathcal{I}_2$ via waterfilling. Obtain the water level $1/\lambda_2$. If $1/\lambda_2 > 1/\mu_1^0$, proceed to Step 5. Otherwise, go to Step 6.
5. Set $\lambda_2 = \mu_1^0$ and proceed to Step 6.
6. If $1/\lambda_i \geq 1/\mu_{\rm ma}^0, \forall i$, set $\lambda_i = \mu_{\rm ma}^0, \forall i$. Check if $1/\lambda_i \leq 1/\mu_{\rm ma}^0, \forall i$. If yes, output $\lambda_i, \forall i$ and break. Otherwise, check if $\sum_i \hat{R}_{{\rm r}i}(\lambda_i) \leq R^{\rm ma}(\mathbf{D}^0)$. If yes, output $\lambda_i, \forall i$ and break. Otherwise, proceed to Step 7.
7. Assuming that $\lambda_j < \lambda_i$, find λ_j' such that $\lvert\mathcal{M}_{{\rm r}j}^+\rvert \log\lambda_j' = \sum_{k \in \mathcal{M}_{{\rm r}j}^+} \log\alpha_j(k) - R^{\rm ma}(\mathbf{D}^0) + \bar{R}_{j{\rm r}}(\mathbf{D}_j^0)$, where $p_{{\rm r}j}(k) = \big(1/\lambda_j' - 1/\alpha_j(k)\big)^+, \forall k \in \mathcal{I}_j$, $\mathcal{M}_{{\rm r}j}^+ \triangleq \{k \lvert p_{{\rm r}j}(k) > 0\}$ and $\lvert\mathcal{M}_{{\rm r}j}^+\rvert$ is the cardinality of the set $\mathcal{M}_{{\rm r}j}^+$. Set $\lambda_j = \lambda_j'$ and output λ_i and λ_j.

By relating the relative water-levels to the corresponding rates and power allocation, the algorithm can be explained more intuitively as follows. Step 1 performs initial power allocation and obtains the initial water level λ^0. The water-levels $\lambda_i = \lambda^0, \forall i$ maximize $R^{\rm bc}(\mathbf{B})$ among all possible $\{\lambda_1, \lambda_2\}$ combinations subject to the power limit of the relay. Step 2 checks whether $\min\{\hat{R}_{{\rm r}i}(\mathbf{B}_i), \bar{R}_{j{\rm r}}(\mathbf{D}_j)\}$ is upper-bounded by $\bar{R}_{j{\rm r}}(\mathbf{D}_j^0), \forall i$. If $\hat{R}_{\rm r1}(\lambda_1^0) > \bar{R}_{2{\rm r}}(\mathbf{D}_2^0)$, the relay reduces its transmission power allocated for relaying the signal from source node 2 to source node 1 so that $\hat{R}_{\rm r1}(\lambda_1) = \bar{R}_{2{\rm r}}(\mathbf{D}_2^0)$ in Step 3. In the case that $\hat{R}_{\rm r1}(\lambda_1)$ is reduced in Step 3, in terms of increasing λ_1, extra power becomes available for relaying the signal from source node 1 to source node 2. Therefore, if $\hat{R}_{\rm r2}(\lambda_2^0) < \bar{R}_{1{\rm r}}(\mathbf{D}_1^0)$, the remaining power of the relay is allocated for relaying the signal from source node 1 to source node 2 at first in Step 4. Later in Step 4, it is checked whether $\hat{R}_{\rm r2}(\lambda_2) > \bar{R}_{1{\rm r}}(\mathbf{D}_1^0)$ under the new power allocation. If $\hat{R}_{\rm r2}(\lambda_2) > \bar{R}_{1{\rm r}}(\mathbf{D}_1^0)$, the relay reduces its transmission power allocated for relaying the signal from source

node 1 to source node 2 so that $\hat{R}_{\mathrm{r}2}(\lambda_2) = \bar{R}_{1\mathrm{r}}(\mathbf{D}_1^0)$ in Step 5. Steps 6 checks whether $\hat{R}_{\mathrm{r}1}(\lambda_1) + \hat{R}_{\mathrm{r}2}(\lambda_2) \leq R^{\mathrm{ma}}(\mathbf{D}^0)$ is satisfied. In the case that this constraint is not satisfied, Step 6 or 7 revise the power allocation so that $\hat{R}_{\mathrm{r}1}(\lambda_1) + \hat{R}_{\mathrm{r}2}(\lambda_2) = R^{\mathrm{ma}}(\mathbf{D}^0)$ and the power consumption of the relay is minimized. We stress that the above procedure in the proposed algorithm, which terminates after Step 6 or 7, is not iterative.

The following theorem regarding the proposed algorithm is in order.

Theorem 3.3: The water-levels obtained using the algorithm for relay optimization in Table 3.1 achieve the optimal relay power allocation for the considered relay optimization problem of sum-rate maximization with minimum relay power consumption.

Proof: See Subsection A.5 in Appendix. ■

Depending on the source node power allocation strategies and the power limit at the relay, different results can be obtained at the output of the algorithm in Table 3.1. Define the following power thresholds $P_{\mathrm{ma}} \triangleq \sum_i \sum_{k\in\mathcal{I}_i} \left(1/\mu_{\mathrm{ma}}^0 - 1/\alpha_i(k)\right)^+$, $P_{\mathrm{sm}} \triangleq \sum_i \sum_{k\in\mathcal{I}_i} \left(1/\max\{\mu_1^0,\mu_2^0\} - 1/\alpha_i(k)\right)^+$, $P_{\mathrm{md}} \triangleq \sum_i \sum_{k\in\mathcal{I}_i} \left(1/\mu_i^0 - 1/\alpha_i(k)\right)^+$ and $P_{\mathrm{lg}} \triangleq \sum_i \sum_{k\in\mathcal{I}_i} \left(1/\min\{\mu_1^0,\mu_2^0\} - 1/\alpha_i(k)\right)^+$ where the subscripts 'sm', 'md', and 'lg' mean 'small', 'medium' and 'large', respectively. Recall from Lemma 3.1 that $\mu_{\mathrm{ma}}^0 > \min\{\mu_1^0,\mu_2^0\}$. Denote the situation that $\mu_{\mathrm{ma}}^0 \geq \max\{\mu_1^0,\mu_2^0\}$ as Case I and the situation that $\mu_{\mathrm{ma}}^0 < \max\{\mu_1^0,\mu_2^0\}$ as Case II, we next analyze the optimal solution in these two cases in detail.

For Case I, it can be seen that $P_{\mathrm{ma}} \leq P_{\mathrm{sm}} \leq P_{\mathrm{md}} \leq P_{\mathrm{lg}}$. According to the value of $P_{\mathrm{r}}^{\max}$, there are five subcases which are discussed one by one in the following text.

Subcase I-1: $P_{\mathrm{r}}^{\max}$ is small such that $P_{\mathrm{r}}^{\max} < P_{\mathrm{ma}}$. In the Subcase I-1, the algorithm proceeds through Steps 1-2-6 and

$$\lambda_i = \lambda^0 > \mu_{\mathrm{ma}}^0, \forall i \tag{3.19a}$$

$$\sum_i \mathrm{Tr}\{\mathbf{P}_{\mathrm{r}i}(\lambda_i)\} = P_{\mathrm{r}}^{\max} \tag{3.19b}$$

at the output of the algorithm, while (3.15a) and (3.15b) are satisfied with inequality.

Note that some power of the source nodes is wasted in this subcase. Since the sum-rate $R^{\mathrm{tw}}(\mathbf{B},\mathbf{D})$ is bounded by $\hat{R}_{\mathrm{r}1}(\lambda_1)+\hat{R}_{\mathrm{r}2}(\lambda_2)$ due to the small power limit of the relay, the source nodes could use less power without reducing $R^{\mathrm{tw}}(\mathbf{B},\mathbf{D})$ if there would be coordination in the system. Indeed, if the source nodes could be coordinated to optimize their power allocation as well, they only need to use the power of $\mathrm{Tr}\{\mathbf{D}_1^\dagger\}$+$\mathrm{Tr}\{\mathbf{D}_2^\dagger\}$ where $\mathbf{D}^\dagger \triangleq [\mathbf{D}_1^\dagger,\mathbf{D}_2^\dagger]$ is the optimal solution to the following problem

$$\min_{\mathbf{D}} \ \mathrm{Tr}\{\mathbf{D}_1\}+\mathrm{Tr}\{\mathbf{D}_2\} \tag{3.20a}$$

$$\mathbf{s.t.}\ R^{\mathrm{ma}}(\mathbf{D}) \geq \hat{R}_{\mathrm{r}1}(\lambda^0)+\hat{R}_{\mathrm{r}2}(\lambda^0) \tag{3.20b}$$

$$\bar{R}_{1\mathrm{r}}(\mathbf{D}_1) \geq \hat{R}_{\mathrm{r}2}(\lambda^0) \tag{3.20c}$$

$$\bar{R}_{2\mathrm{r}}(\mathbf{D}_2) \geq \hat{R}_{\mathrm{r}1}(\lambda^0). \tag{3.20d}$$

It can be shown that $\mathrm{Tr}\{\mathbf{D}_1^0\}+\mathrm{Tr}\{\mathbf{D}_2^0\} > \mathrm{Tr}\{\mathbf{D}_1^\dagger\}+\mathrm{Tr}\{\mathbf{D}_2^\dagger\}$ in this subcase. Therefore, the power of $\mathrm{Tr}\{\mathbf{D}_1^0\}+\mathrm{Tr}\{\mathbf{D}_2^0\}-\mathrm{Tr}\{\mathbf{D}_1^\dagger\}-\mathrm{Tr}\{\mathbf{D}_2^\dagger\}$ is wasted at the source nodes because of the lack of coordination.

Subcase I-2: increase $P_{\mathrm{r}}^{\max}$ such that $P_{\mathrm{ma}} \leq P_{\mathrm{r}}^{\max} \leq P_{\mathrm{sm}}$. Then the algorithm proceeds through Steps 1-2-6.

Subcase I-3: increase $P_{\mathrm{r}}^{\max}$ such that $P_{\mathrm{sm}} < P_{\mathrm{r}}^{\max} \leq P_{\mathrm{md}}$. Then the algorithm proceeds through Steps 1-2-3-4-6.

Subcase I-4: further increase $P_{\mathrm{r}}^{\max}$ such that $P_{\mathrm{md}} < P_{\mathrm{r}}^{\max} \leq P_{\mathrm{lg}}$. Then the algorithm proceeds through Steps 1-2-3-4-5-6.

Subcase I-5: further increase $P_{\mathrm{r}}^{\max}$ such that $P_{\mathrm{r}}^{\max} > P_{\mathrm{lg}}$. Then the algorithm proceeds through Steps 1-2-3-5-6. In the above subcases when $P_{\mathrm{r}}^{\max} \geq P_{\mathrm{ma}}$, it holds that

$$\lambda_i = \mu_{\mathrm{ma}}^0 \geq \lambda^0, \forall i \tag{3.21a}$$

$$\sum_i \mathrm{Tr}\{\mathbf{P}_{\mathrm{r}i}(\lambda_i)\} \leq P_{\mathrm{r}}^{\max} \tag{3.21b}$$

at the output of the algorithm, while (3.15a) is satisfied with inequality for each i such that $1/\mu_i^0 > 1/\mu_{\mathrm{ma}}^0$ and (3.15b) is satisfied with equality. For these subcases, the sum-rate $R^{\mathrm{tw}}(\mathbf{B},\mathbf{D})$ is bounded by $R^{\mathrm{ma}}(\mathbf{D}^0)$ and there is no waste of power at the source nodes.

For Case II, it holds that $\min\{\mu_1^0, \mu_2^0\} < \mu_{\rm ma}^0 < \max\{\mu_1^0, \mu_2^0\}$ according to Lemma 3.1. Assume that $\mu_2^0 > \mu_1^0$ and find $\bar{\lambda}_2$ such that $\hat{R}_{\rm r2}(\bar{\lambda}_2) = R^{\rm ma}(\mathbf{D}^0) - \bar{R}_{\rm 2r}(\mathbf{D}_2^0)$. Let $\bar{\lambda}_1 = \mu_2^0$ and define $P'_{\rm ma} \triangleq \sum_i \sum_{k\in\mathcal{I}_i} \left(1/\bar{\lambda}_i - 1/\alpha_i(k)\right)^+$. It can be seen from Lemma 3.3 that $P'_{\rm ma} > P_{\rm ma}$. Since $\mu_{\rm ma}^0 < \max\{\mu_1^0, \mu_2^0\}$, it holds that $P'_{\rm ma} > P_{\rm sm}$. Therefore, for Case II, the power thresholds satisfy $P_{\rm sm} < P'_{\rm ma} < P_{\rm md} < P_{\rm lg}$. The following subcases appear as $P_{\rm r}^{\max}$ increases.

Subcase II-1: $P_{\rm r}^{\max}$ is small such that $P_{\rm r}^{\max} < P_{\rm sm}$. Then, the algorithm proceeds through Steps 1-2-6 and

$$\lambda_i = \lambda^0 > \max\{\mu_1^0, \mu_2^0\}, \forall i \tag{3.22a}$$

$$\sum_i {\rm Tr}\{\mathbf{P}_{{\rm r}i}(\lambda_i)\} = P_{\rm r}^{\max} \tag{3.22b}$$

at the output of the algorithm, while (3.15a) and (3.15b) are satisfied with inequality.

Subcase II-2: increase $P_{\rm r}^{\max}$ such that $P_{\rm sm} \le P_{\rm r}^{\max} \le P'_{\rm ma}$. Then the algorithm proceeds through Steps 1-2-3-4-6 and

$$\lambda_1 = \mu_2^0 \ge \lambda^0 \tag{3.23a}$$

$$\sum_i {\rm Tr}\{\mathbf{P}_{{\rm r}i}(\lambda_i)\} = P_{\rm r}^{\max} \tag{3.23b}$$

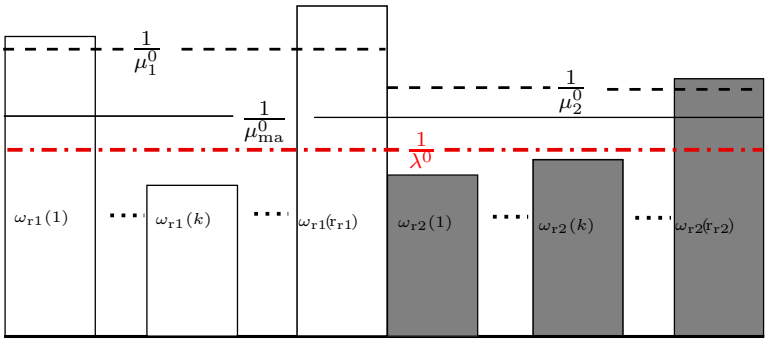
at the output of the algorithm, while (3.15a) is satisfied with equality for $i = 1$ and inequality for $i = 2$. Note that there is waste of power at the source nodes for the above two subcases as long as $P_{\rm r}^{\max} < P'_{\rm ma}$ because the sum-rate $R^{\rm tw}(\mathbf{B}, \mathbf{D})$ is bounded by $\hat{R}_{\rm r1}(\lambda_1) + \hat{R}_{\rm r2}(\lambda_2)$.

Subcase II-3: increase $P_{\rm r}^{\max}$ such that $P'_{\rm ma} < P_{\rm r}^{\max} \le P_{\rm md}$. Then the algorithm proceeds through Steps 1-2-3-4-6-7.

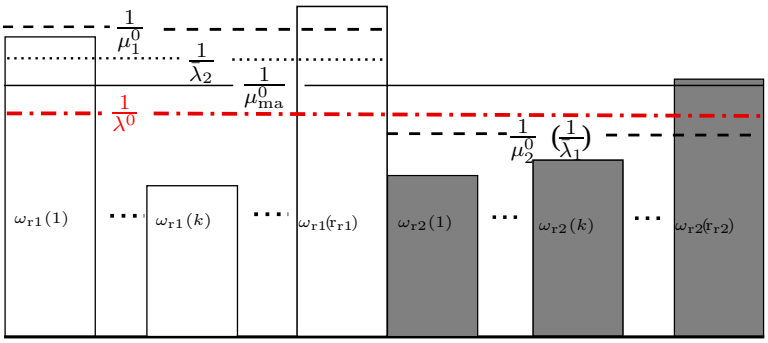
Subcase II-4: further increase $P_{\rm r}^{\max}$ such that $P_{\rm md} < P_{\rm r}^{\max} \le P_{\rm lg}$. Then the algorithm proceeds through Steps 1-2-3-4-5-6-7. Subcase II-5: further increase $P_{\rm r}^{\max}$ such that $P_{\rm r}^{\max} > P_{\rm lg}$. Then the algorithm proceeds through Steps 1-2-3-5-6-7. In the subcases when $P_{\rm r}^{\max} \ge P'_{\rm ma}$, it holds that

$$\lambda_1 = \mu_2^0 > \lambda^0 \tag{3.24a}$$

$$\sum_i {\rm Tr}\{\mathbf{P}_{{\rm r}i}(\lambda_i)\} \le P_{\rm r}^{\max} \tag{3.24b}$$



(a) Subcase I-1: $P_{\rm r}^{\max} < P_{\rm ma}$, $\mu_{\rm ma}^0 \geq \max\{\mu_1^0, \mu_2^0\}$.



(b) Subcase II-2: $P_{\rm sm} \leq P_{\rm r}^{\max} < P'_{\rm ma}$, $\mu_{\rm ma}^0 < \max\{\mu_1^0, \mu_2^0\}$.

Figure 3.1: Illustration of μ_1^0, μ_2^0, $\mu_{\rm ma}^0$, and λ^0 for the scenario of relay optimization.

at the output of the algorithm, while (3.15a) is satisfied with equality for $i = 1$ and inequality for $i = 2$, and (3.15b) is satisfied with equality. The optimal λ_2 is found in Step 7 of the proposed algorithm. For these subcases, there is no waste of power at the source nodes.

Two of the above subcases, i.e., Subcases I-1 and II-2, are illustrated in Fig. 3.1.

From the above discussion, it can be seen that the algorithm in Table 3.1 obtains the optimal power allocation in at most seven steps without iterations.

Recall that the sum-rate of DF TWR is bounded by both the sum-rate of the MA phase and the sum-rate of the BC phase. In the scenario of relay optimization, the relay optimizes its power allocation which affects the sum-rate of the BC phase. Since the relay may or may not use all its available power at optimality (i.e., for the optimal power allocation), the sum-rate of the BC phase is not necessarily maximized at optimality. Moreover, it is also possible that the sum-rate of the BC

phase at optimality is not even the maximum sum-rate of the BC phase that can be achieved using the power consumed by the relay at optimality. We specify the term *efficient* to describe such optimal power allocation of the relay that maximizes the BC phase sum-rate $R^{\mathrm{bc}}(\mathbf{B})$ with the actually consumed power at the relay. Thus, the relay's power allocation is efficient if it generates the maximum sum-rate for broadcasting the messages of the source nodes given its power consumption. For example, when the relay uses all its available power at optimality, the optimal power allocation of the relay is efficient if it maximizes the sum-rate of the BC phase, and inefficient otherwise. When the relay uses the power $P_{\mathrm{r}} < P_{\mathrm{r}}^{\max}$ at optimality, the optimal power allocation is efficient if the achieved sum-rate of the BC phase is the maximum achievable sum-rate of the BC phase with power consumption P_{r}, and inefficient otherwise. Then the following two conclusions can be drawn for the scenario of relay optimization.

First, the optimal relay power allocation in the relay optimization scenario is always efficient for Case I (i.e., $\mu_{\mathrm{ma}}^0 \geq \max\{\mu_1^0, \mu_2^0\}$). In such a case, it can be seen from (3.19a) and (3.21a) that $1/\lambda_1 = 1/\lambda_2$ at optimality regardless of whether the relay uses all its available power. Therefore, the BC phase sum-rate $R^{\mathrm{bc}}(\mathbf{B})$ is always maximized given the relay's power consumption. However, the optimal relay power allocation is inefficient for Case II (i.e., $\mu_{\mathrm{ma}}^0 < \max\{\mu_1^0, \mu_2^0\}$) as long as $P_{\mathrm{r}}^{\max} > P_{\mathrm{l}}$. Moreover, the larger the difference between $\max\{\mu_1^0, \mu_2^0\}$ and μ_{ma}^0 in this case, the more inefficient the optimal relay power allocation becomes when $P_{\mathrm{r}}^{\max} > P_{\mathrm{l}}$. Given the definitions (4.2a)-(4.2c) and Lemma 3.1, $\mu_{\mathrm{ma}}^0 < \max\{\mu_1^0, \mu_2^0\}$ in Case II indicates that one source node uses more power, has more antennas and/or better channel condition compared to those of the other source node. Indeed, if the power budget, number of antennas, and channel conditions are the same for the two source nodes, as an extreme example, it leads to $\mu_{\mathrm{ma}}^0 > \mu_1^0 = \mu_2^0$. Therefore, it can be seen that the asymmetry between the power budget, number of antennas, and/or channel conditions can degrade the relay power allocation efficiency in the scenario of relay optimization.

Second, the considered relay optimization scenario may result in the waste of power at the source nodes. However, the relay never wastes any power. This is due

to the fact that the relay is aware of the source node power allocation strategies and optimizes its own power allocation based on them. As a result, it can use only part of the available power if its power limit $P_{\mathrm{r}}^{\max}$ is large. However, the relay power allocation strategy is unknown to the source nodes when the source nodes decide their power allocation strategies. Therefore, the possibility of wasting power in the relay optimization scenario can be viewed as the tradeoff for low complexity. Indeed, in the relay optimization scenario, there is no coordination between the relay and the source nodes. As a result, it is almost impossible to achieve the maximum sum-rate with minimum total power consumption referred to as network-level optimality. In order to achieve the network-level optimality, the scenario of network optimization, in which the relay and the source nodes jointly maximize the sum-rate of the TWR with minimum total power consumption, is considered in Chapter 4 of this thesis.

3.3 Numerical and simulation results

In this section, we provide simulation examples for some results presented earlier and demonstrate the proposed algorithm for relay optimization in Table 3.1. The general setup is as follows. The elements of the channels $\mathbf{H}_{\mathrm{r}i}$ and $\mathbf{H}_{i\mathrm{r}}, \forall i$ are generated from complex Gaussian distribution with zero mean and unit variance unless otherwise specified. The noise variances $\sigma_i^2, \forall i$ and σ_{r}^2 are equal to each other and denoted uniformly as σ^2. While the source node power allocation strategy $\mathbf{D}^0$ can be arbitrary, we use for simulations the $\mathbf{D}^0$ that maximizes the MA phase sum-rate $R^{\mathrm{ma}}(\mathbf{D})$. The rates $R^{\mathrm{ma}}(\mathbf{D})$, $\bar{R}_{i\mathrm{r}}(\mathbf{D}_i)$, and $\hat{R}_{\mathrm{r}i}(\mathbf{B}_i)$ are briefly denoted as R^{ma}, $\bar{R}_{i\mathrm{r}}$, and $\hat{R}_{\mathrm{r}i}$, respectively, in the figures in this section.

3.3.1 A demonstration of Lemma 3.2

It is assumed that the number of antennas at the relay n_{r} is 8 while source node 1 has $n_1 = 6$ antennas and source node 2 has $n_2 = 5$ antennas. Each curve in Fig. 3.2 shows the sum-rate $\hat{R}_{\mathrm{r}1} + \hat{R}_{\mathrm{r}2}$ versus the water-level $1/\lambda_1$ for a given ratio of $P_{\mathrm{r}}^{\max}$ over σ^2. In each curve, for each given $1/\lambda_1$, the relay consumes all the remaining

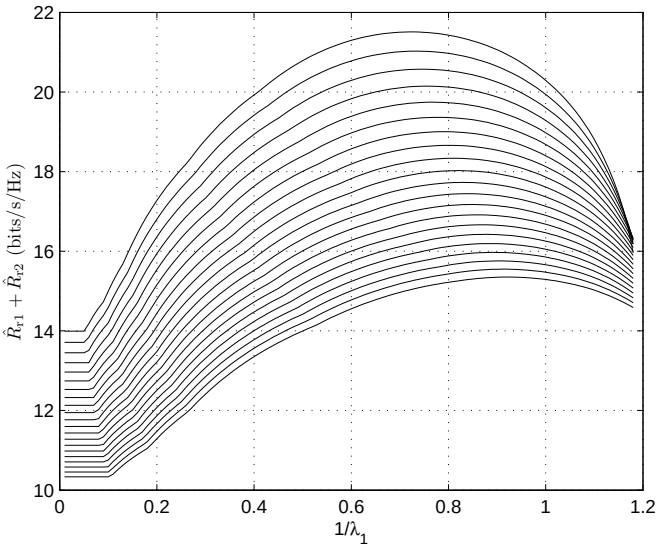
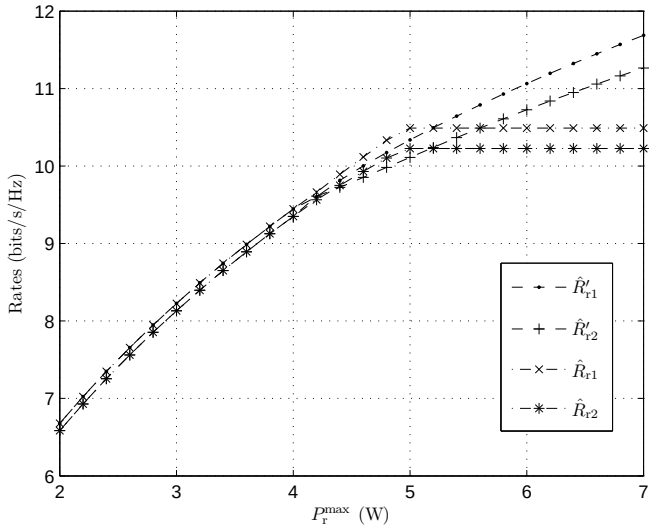


Figure 3.2: $\hat{R}_{\mathrm{r}1} + \hat{R}_{\mathrm{r}2}$ versus $1/\lambda_1$ under different $P_{\mathrm{r}}^{\max}/\sigma^2$.

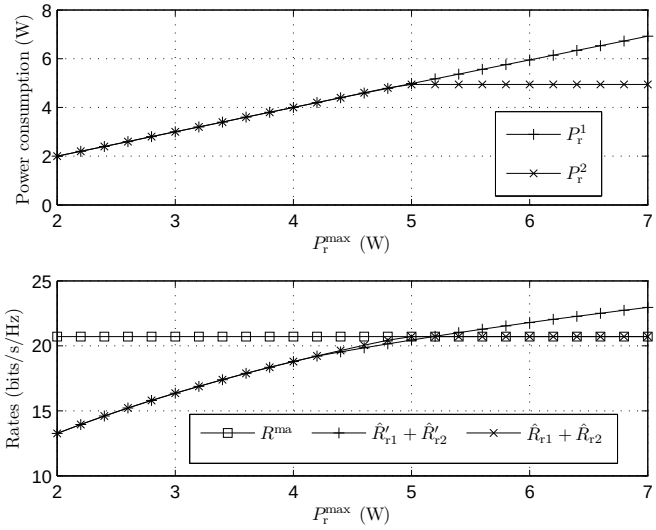
power to maximize $1/\lambda_2$. Therefore, the power consumption of the relay is fixed and equals $P_{\mathrm{r}}^{\max}$. For each curve, σ^2 is different. The curve at the bottom corresponds to the ratio $P_{\mathrm{r}}^{\max}/\sigma^2$ equal to 4 dB. For each time, when the ratio of $P_{\mathrm{r}}^{\max}$ over σ^2 increases, a new curve of $\hat{R}_{\mathrm{r}1} + \hat{R}_{\mathrm{r}2}$ versus $1/\lambda_1$, which lies above the previous curve, is plotted. The curve at the top corresponds to the ratio $P_{\mathrm{r}}^{\max}/\sigma^2$ equal to 7 dB. It can be seen from Fig. 3.2 that the sum-rate $\hat{R}_{\mathrm{r}1} + \hat{R}_{\mathrm{r}2}$ is a nonconvex function of $1/\lambda_1$. However, $\hat{R}_{\mathrm{r}1} + \hat{R}_{\mathrm{r}2}$ is non-decreasing before reaching the maximum and non-increasing after that. Note that $1/\lambda_1 = 1/\lambda_2 = 1/\lambda^0$ when the BC phase sum-rate is maximized. As a result, it can be seen that increasing $\max\{1/\lambda_1, 1/\lambda_2\}$ and decreasing $\min\{1/\lambda_1, 1/\lambda_2\}$ while fixing the total power consumption leads to a smaller BC phase sum-rate for any given $\{1/\lambda_1, 1/\lambda_2\}$. Therefore, Fig. 3.2 verifies the result presented in Lemma 3.2.

3.3.2 The relay optimization problem

Fig. 3.3a compares the BC phase rates at optimality of the relay optimization problem, which considers power consumption minimization, with the BC phase rates at optimality of the problem (3.11), which does not minimize the power consumption, under different $P_{\mathrm{r}}^{\max}$. One channel realization is shown. The specific setup



(a) $\hat{R}_{{\rm r}i}$ in the optimal solution of the sum-rate maximization problems with and without minimization of power consumption, respectively, versus $P_{\rm r}^{\max}$.



(b) Relay power consumption, $R^{\rm ma}(\mathbf{D}^0)$ and $\sum_i \hat{R}_{{\rm r}i}$ in the optimal solution of the sum-rate maximization problems with and without minimization of power consumption, respectively, versus $P_{\rm r}^{\max}$.

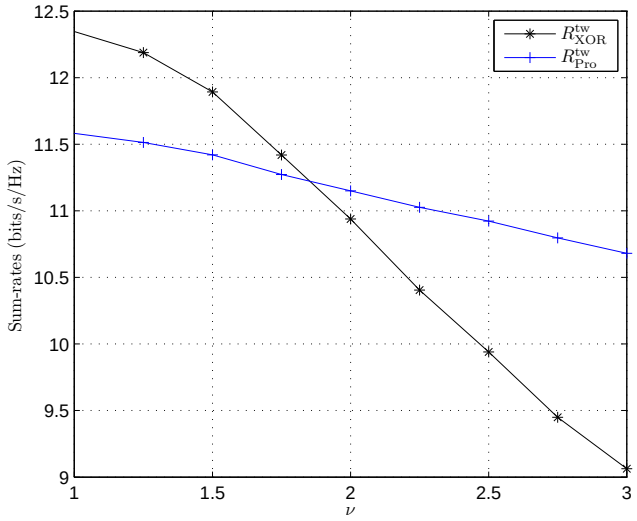
Figure 3.3: Illustration of relay optimization.

for this simulation is as follows. The number of antennas n_1, n_2, and $n_{\rm r}$ are set to be $6, 5$, and 8, respectively. The power limits for the source nodes are set to be $P_1^{\max} = P_2^{\max} = 3$ W. The noise variance is normalized so that $\sigma^2 = 1$. The MA phase rates for this channel realization are 20.7 for $R^{\rm ma}(\mathbf{D}^0)$, 11.2 for $\bar{R}_{1{\rm r}}(\mathbf{D}_1^0)$, and

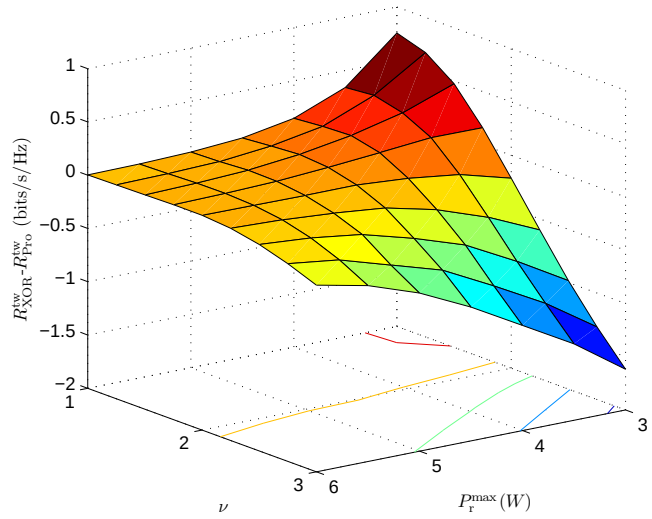
11.0 for $\bar{R}_{2\mathrm{r}}(\mathbf{D}_2^0)$. In Fig. 3.3a, $\hat{R}'_{\mathrm{r}i}$ represents $\hat{R}_{\mathrm{r}i}(\mathbf{B}'_i)$ where $\mathbf{B}'_i$'s, $\forall i$ are the optimal solution (obtained using CVX [81]) to the problem (3.11) which does not minimize the power consumption, and $\hat{R}_{\mathrm{r}i}$ represents $\hat{R}_{\mathrm{r}i}(\mathbf{B}_i)$ where $\mathbf{B}_i$'s, $\forall i$ are the optimal solution to the relay optimization problem considering power consumption minimization obtained using the algorithm in Table 3.1. It can be seen from Fig. 3.3a that $\hat{R}'_{\mathrm{r}i} = \hat{R}_{\mathrm{r}i}$ when $P_{\mathrm{r}}^{\max}$ is small. The reason is that $\hat{R}'_{\mathrm{r}i}$ is small when $P_{\mathrm{r}}^{\max}$ is below certain threshold. As a result, the constraints in (3.12) and (3.18b) are always satisfied and the solutions to the problem (3.11) and the relay optimization problem are the same. As $P_{\mathrm{r}}^{\max}$ increases, $R^{\mathrm{tw}}(\mathbf{B}, \mathbf{D}^0)$ becomes larger and is finally bounded by $R^{\mathrm{ma}}(\mathbf{D}^0)$, while the relay power consumption is not necessarily minimized in the solution of the problem (3.11) which does not consider power consumption minimization. This can be seen from the first subplot of Fig. 3.3b, which shows that the power consumption in the solution derived using the proposed algorithm, denoted as P_{r}^2, saturates when $P_{\mathrm{r}}^{\max} \geq 4.9$ W, while the power consumption in the solution to the problem (3.11) which does not consider power consumption minimization, denoted as P_{r}^1, keeps increasing. As a result, as can be seen from the second subplot of Fig. 3.3b, $\sum_i \hat{R}_{\mathrm{r}i}$ never exceeds $R^{\mathrm{ma}}(\mathbf{D}^0)$, while $\sum_i \hat{R}'_{\mathrm{r}i}$ grows beyond $R^{\mathrm{ma}}(\mathbf{D}^0)$ when $R^{\mathrm{tw}}(\mathbf{B}, \mathbf{D}^0)$ is bounded by $R^{\mathrm{ma}}(\mathbf{D}^0)$. Meanwhile, it can also be seen from the second subplot of Fig. 3.3b that the maximum sum-rates $R^{\mathrm{tw}}(\mathbf{B}, \mathbf{D}^0)$ for the two compared solutions are the same, both of which equal to $\sum_i \hat{R}'_{\mathrm{r}i} = \sum_i \hat{R}_{\mathrm{r}i}$ when $\sum \hat{R}'_{\mathrm{r}i} \leq R^{\mathrm{ma}}(\mathbf{D}^0)$ and equal to $R^{\mathrm{ma}}(\mathbf{D}^0)$ when $\sum \hat{R}'_{\mathrm{r}i} > R^{\mathrm{ma}}(\mathbf{D}^0)$. Thus, this example demonstrates that the proposed algorithm in Table 3.1 achieves maximum sum-rate in the scenario of relay optimization with minimum power consumption.

3.3.3 Comparison with XOR-based relay scheme

We must first clarify that there is no XOR-based scheme for us to conduct a fair comparison with the proposed scheme. The reason is that no XOR-based scheme has been proposed to maximize the sum-rate of the TWR and at the same time minimize the power consumption of the relay as the proposed scheme does. Therefore, to perform this comparison, we need to use the XOR-based scheme that maximizes the sum-rate of MIMO DF TWR without considering the power consumption as



(a) Sum-rates vs. ν, comparison of the XOR-based scheme of [82] (without considering power consumption minimization) and the proposed scheme averaged over 5000 channel realizations ($P_{\mathrm{r}}^{\max} = 3$).



(b) Difference of sum-rates vs. $P_{\mathrm{r}}^{\max}$ and ν, comparison of the XOR-based scheme of [82] (without considering power consumption minimization) and the proposed scheme averaged over 5000 channel realizations.

Figure 3.4: Comparison with XOR based relay scheme.

in [82]. First, we compare the maximum sum-rates achieved by the XOR-based scheme of [24] and the proposed scheme versus the channel asymmetry. In this simulation, we set the number of antennas such that $n_1 = 4, n_2 = 3$, and $n_{\mathrm{r}} = 6$. Power limits are $P_1^{\max} = P_2^{\max} = 2$ W, $P_{\mathrm{r}}^{\max} = 3$ W. Noise power σ^2 is set to

1. The elements of the channels $\mathbf{H}_{\mathrm{r}1}$ and $\mathbf{H}_{\mathrm{r}2}$ are complex Gaussian distributed with zero mean and variances ν and $1/\nu$, respectively. Therefore, when ν becomes larger, the channels become more asymmetric. For each value of ν, the sum-rates obtained by the XOR-based scheme of [24] and the proposed scheme are averaged over 5000 channel realizations and are shown in Fig. 3.4a, denoted as $R_{\mathrm{XOR}}^{\mathrm{tw}}$ and $R_{\mathrm{Pro}}^{\mathrm{tw}}$, respectively. From this figure, it can be seen that the XOR-based scheme is better than the proposed scheme when the channel asymmetry is not very large. On the other hand, the proposed scheme becomes superior when the channel asymmetry is large, i.e., $\nu > 1.9$. Moreover, it can be seen that the XOR-based scheme is much more sensitive to channel asymmetry as its performance decreases much faster than that of the proposed scheme when the asymmetry increases.

We also compare the maximum sum-rates achieved by the XOR-based and the proposed schemes versus both $P_{\mathrm{r}}^{\max}$ and ν. In this simulation, the number of antennas, noise power, and power limits of the source nodes are the same as in the previous simulation. We vary $P_{\mathrm{r}}^{\max}$ and ν so that $P_{\mathrm{r}}^{\max}$ increases from 3 W to 6 W and ν increases from 1 to 3. For each combination of $P_{\mathrm{r}}^{\max}$ and ν, we obtain the sum-rates of the XOR-based scheme and the proposed scheme (averaged over 5000 channel realizations) and show their difference in Fig. 3.4b. From this figure, it can be seen that, the difference of the two compared schemes is small in terms of achieved sum-rate when $P_{\mathrm{r}}^{\max}$ is large. Indeed, even for the very symmetric case ($\nu = 1$), the advantage of the XOR-based scheme vanishes as the power limit $P_{\mathrm{r}}^{\max}$ increases. Similarly, for the asymmetric case, the advantage of the proposed scheme also decreases when $P_{\mathrm{r}}^{\max}$ increases. Therefore, it shows that neither of the proposed scheme and the XOR-based scheme is definitely superior. The XOR-based scheme achieves higher sum-rate than the proposed scheme when the channel is symmetric. The proposed scheme, on the other hand, is better for the case of asymmetric channels. Nevertheless, when the relay power limit increases, the difference of the two schemes vanishes.

3.3.4 The effect of asymmetry in source node power limits and number of antennas

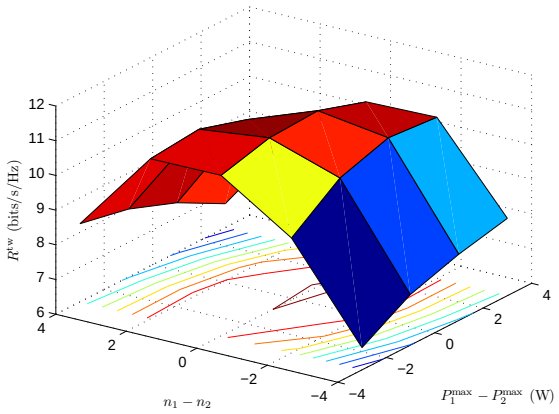
The specific setup for this example is as follows. The noise variance is normalized so that $\sigma^2 = 1$. The number of antennas at the relay, i.e., n_{r}, is set to be 6. The power limit of the relay, i.e., $P_{\mathrm{r}}^{\max}$ is set to be 3 W. The total number of antennas at both source nodes is fixed such that $n_1 + n_2 = 6$. The total available power at both source nodes is also fixed such that $P_1^{\max} + P_2^{\max} = 5$ W. Given the above total number of antennas and total available power at the source nodes, the relay optimization problem is solved for different n_1, n_2, $P_1^{\max}$, and $P_2^{\max}$ for 1000 channel realizations. The resulting average sum-rate and average power consumption of the relay, and the percentage of efficient power allocation at optimality are plotted in Figs. 3.5a, 3.5b, and 3.5c, respectively, versus the difference between the number of antennas and the difference between the power limits at the source nodes. From Fig. 3.5a, it can be seen that the sum-rate at optimality of the relay optimization is the largest when there is no asymmetry in the number of antennas at the source nodes and no asymmetry or only small asymmetry in the power limits of the source nodes. As the asymmetry becomes larger in either number of antennas or power limits, the sum-rate at optimality of the relay optimization decreases. Therefore, it can be seen from this figure that the asymmetry in the above aspects leads to smaller sum-rate at optimality of the considered relay optimization problem. Relating Figs. 3.5b and 3.5c to Fig. 3.5a, two more observations can be made. First, the relay does not necessarily use all the available power for sum-rate maximization in the relay optimization scenario. Second, the asymmetry in the number of antennas and power limits leads to low power allocation efficiency. It can be seen from Fig. 3.5b that when one of $P_1^{\max} - P_2^{\max}$ and $n_1 - n_2$ is positive while the other is negative, the relay uses a part of its available power. However, the achieved sum-rate is smaller compared to the sum-rate in the case when $P_1^{\max} - P_2^{\max} = 0$ and $n_1 - n_2 = 0$ (see Fig. 3.5a). In this situation, since the average power consumption and the average sum-rate are both low, the percentage of efficient power allocation is larger than 0 but less than the percentage when $P_1^{\max} - P_2^{\max} = 0$ and $n_1 - n_2 = 0$, as can be seen from Fig. 3.5c. When $P_1^{\max} - P_2^{\max}$ and $n_1 - n_2$ are

both positive or both negative, the relay uses more power than the power used in the case when $P_1^{\max} - P_2^{\max} = 0$ and $n_1 - n_2 = 0$ while the achieved sum-rate is smaller than that in the latter case. In this situation, since the average power consumption is high while the average sum-rate is low, the percentage of efficient power allocation is very low, if not zero, as can be seen from Fig. 3.5c. The above facts become more obvious when the asymmetry becomes larger. Therefore, it can be seen from Figs. 3.5b and 3.5c that the asymmetry on the power limits and the number of antennas can lead to low power allocation efficiency.

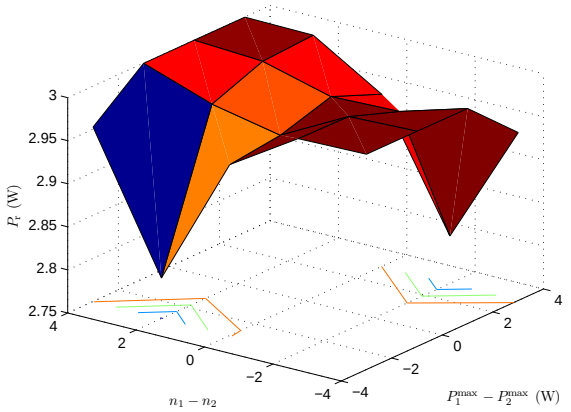
3.4 Conclusion

In this chapter, we have solved the problem of sum-rate maximization with minimum power consumption for MIMO DF TWR in the scenario of relay optimization. For finding the optimal solution, we have found a sufficient and necessary optimality condition for power allocation. Based on this condition, we have proposed an algorithm to find the optimal solution. The proposed algorithm allows the relay to obtain its optimal power allocation in several steps. We have shown that, as a trade-off for low complexity, there can be waste of power at the source nodes in the relay optimization scenario because of the lack of coordination. We have also shown that the asymmetry in the number of antennas and power limits at the source nodes can result in the sum-rate performance degradation and the power allocation inefficiency in MIMO DF TWR.

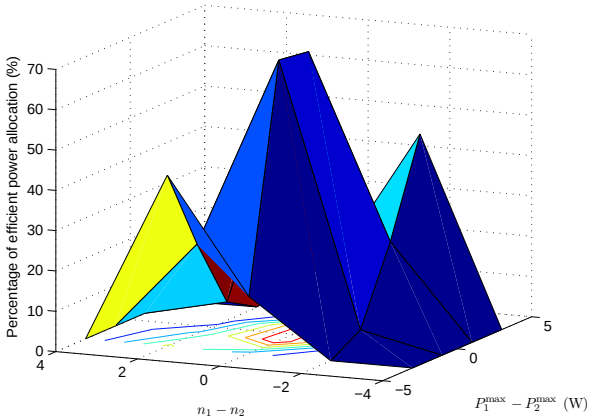
$\sim$



(a) Sum-rate at optimality.



(b) Relay power consumption at optimality.



(c) Percentage of efficient power allocation at optimality.

Figure 3.5: Effect of asymmetry: the average sum-rate, average relay power consumption, and percentage of efficient power allocation at optimality of relay optimization versus the difference between number of antennas and the difference between power limits at the source nodes in 1000 channel realizations.

Chapter 4

Maximizing Spectral Efficiency with Minimum Power in MIMO DF TWR with Full Cooperation

The solution of the relay optimization scenario derived in Chapter 3 gives the optimal power allocation of the relay in a MIMO DF TWR system in the case when there is no coordination between the relay and the source nodes. However, if the participating nodes have sufficient computational capability and can jointly optimize their power allocation strategies, a better performance than that in the relay optimization scenario can be achieved. This chapter studies the problem of sum-rate maximization with minimum power consumption for MIMO DF TWR in the *network optimization scenario* in which the relay and the source nodes jointly optimize their power allocations. The objective of this chapter is to find the jointly optimal power allocation of the relay and the source nodes while reducing the complexity of finding the optimal solution. [1]

4.1 System model

The system model used in this chapter is the same as described in Section 3.1 of Chapter 3. Therefore, expressions (3.1)-(3.10) still hold here. The detailed model is omitted here. However, it is important to recall that with the actual water-levels used by the relay for relaying the signal from source node j to source node i denoted

[1] *A version of this chapter has been published in IEEE Trans. Signal Process., 61: 3578–3591 (2013).*

as $1/\lambda_i, \forall i$, it holds that

$$\hat{R}_{\mathrm{r}1}(\mathbf{B}_1) = \sum_{k\in\mathcal{I}_1} \log\left(1 + \left(\frac{1}{\lambda_1}\alpha_1(k) - 1\right)^+\right) \tag{4.1a}$$

$$\hat{R}_{\mathrm{r}2}(\mathbf{B}_2) = \sum_{k\in\mathcal{I}_2} \log\left(1 + \left(\frac{1}{\lambda_2}\alpha_2(k) - 1\right)^+\right) \tag{4.1b}$$

where $\mathcal{I}_i = \{1, \ldots, r_{\mathrm{r}i}\}$. Therefore, the rate $\hat{R}_{\mathrm{r}i}(\mathbf{B}_i)$ obtained using water-level $1/\lambda_i$ is alternatively denoted as $\hat{R}_{\mathrm{r}i}(\lambda_i)$.

It is also necessary to recall that, same as in Chapter 3, the relative water levels $1/\mu_1(\mathbf{D}_1)$, $1/\mu_2(\mathbf{D}_2)$, and $1/\mu_{\mathrm{ma}}(\mathbf{D})$ are defined as

$$\sum_{k\in\mathcal{I}_2} \log\left(1 + \left(\frac{1}{\mu_1(\mathbf{D}_1)}\alpha_2(k) - 1\right)^+\right) = \bar{R}_{1\mathrm{r}}(\mathbf{D}_1) \tag{4.2a}$$

$$\sum_{k\in\mathcal{I}_1} \log\left(1 + \left(\frac{1}{\mu_2(\mathbf{D}_2)}\alpha_1(k) - 1\right)^+\right) = \bar{R}_{2\mathrm{r}}(\mathbf{D}_2) \tag{4.2b}$$

$$\sum_i \sum_{k\in\mathcal{I}_i} \log\left(1 + \left(\frac{1}{\mu_{\mathrm{ma}}(\mathbf{D})}\alpha_i(k) - 1\right)^+\right) = R^{\mathrm{ma}}(\mathbf{D}). \tag{4.2c}$$

With the same system model, the considered scenario in this chapter is different. In Chapter 3 with the relay optimization scenario, given $\mathbf{W}_1$ and $\mathbf{W}_2$ as the transmit strategies of the source nodes, the relay optimize its own transmit strategy $\mathbf{B}$. In this chapter with network optimization scenario, the source nodes and relay jointly optimize $\mathbf{W}_1$, $\mathbf{W}_2$, and $\mathbf{B}$ such that the maximal spectrum efficiency is achieved with minimum total power consumption in the system.

For the network optimization scenario considered here, the relay and the source nodes jointly maximize the sum-rate in (3.6) with minimum total transmission power in the network.[2] Similar to the relay optimization scenario, the relay needs to know $\mathbf{W}_1$ and $\mathbf{W}_2$ while both source nodes need to know $\mathbf{T}_{\mathrm{r}1}$ and $\mathbf{T}_{\mathrm{r}2}$. In the network optimization scenario, it is preferable that the TWR is able to operate in a centralized mode in which the relay can serve as a central node that carries out the computations. If the system works in a decentralized mode, it may lead to high

[2]The term 'sum-rate' by default means $R^{\mathrm{tw}}(\mathbf{B}, \mathbf{D})$ when we do not specify it to be the sum-rate of the BC or MA phase.

overhead because of the information exchange during the iterative optimization process.

We next solve the network optimization problem.

4.2 Network optimization

In the network optimization scenario, the relay and the source nodes jointly optimize their power allocation to achieve sum-rate maximization with minimum total power consumption in the system for the MIMO DF TWR. Compared to the optimal solution of the relay optimization problem in Chapter 3, the optimal solution of the network optimization problem achieves larger sum-rate and/or less power consumption at the cost of higher computational complexity.

The sum-rate maximization part can be formulated as the following optimization problem[3]

$$\max_{\{\mathbf{B},\mathbf{D}\}} \quad R^{\mathrm{tw}}(\mathbf{B},\mathbf{D}) \tag{4.3a}$$

$$\mathbf{s.t.} \quad \mathrm{Tr}\{\mathbf{D}_i\} \leq P_i^{\max}, \forall i \tag{4.3b}$$

$$\mathrm{Tr}\{\mathbf{B}_1+\mathbf{B}_2\} \leq P_{\mathrm{r}}^{\max}. \tag{4.3c}$$

where $P_i^{\max}$ and $P_{\mathrm{r}}^{\max}$ are the power limits for source node i and the relay, respectively. The above problem is a convex problem which can be rewritten into the standard form by introducing variables t, t_1, t_2 as follows

$$\max_{\{t,t_1,t_2,\mathbf{B},\mathbf{D}\}} \quad t \tag{4.4a}$$

$$\mathbf{s.t.} \quad t \leq R^{\mathrm{ma}}(\mathbf{D}),\ t \leq t_1+t_2 \tag{4.4b}$$

$$t_i \leq \hat{R}_{\mathrm{r}j}(\mathbf{B}_j),\ t_i \leq \bar{R}_{i\mathrm{r}}(\mathbf{D}_i), \forall i \tag{4.4c}$$

$$\mathrm{Tr}\{\mathbf{D}_i\} \leq P_i^{\max}, \forall i,\ \mathrm{Tr}\{\mathbf{B}_1+\mathbf{B}_2\} \leq P_{\mathrm{r}}^{\max}. \tag{4.4d}$$

If transmission power minimization is also taken into account, the following

[3]The positive semi-definite (PSD) constraints $\mathbf{D}_i \succeq 0, \forall i$ and $\mathbf{B}_i \succeq 0, \forall i$ are assumed as default and omitted for brevity in all formations of optimization problems in this chapter.

constraints become necessary

$$\hat{R}_{\mathrm{r}i}(\mathbf{B}_i) \leq \bar{R}_{j\mathrm{r}}(\mathbf{D}_j), \forall i \tag{4.5a}$$

$$R^{\mathrm{ma}}(\mathbf{D}) = R(\mathbf{B}, \mathbf{D}). \tag{4.5b}$$

The reason why the above constraints are necessary if transmission power minimization also needs to be taken into account is as follows. Given the fact that $R^{\mathrm{ma}}(\mathbf{D}) < \bar{R}_{1\mathrm{r}}(\mathbf{D}_1) + \bar{R}_{2\mathrm{r}}(\mathbf{D}_2)$ whenever $\mathrm{Tr}\{\mathbf{D}_1\} + \mathrm{Tr}\{\mathbf{D}_2\} > 0$, it can be seen that the power consumption of the relay can be reduced by reducing $\mathrm{Tr}\{\mathbf{B}_i\}$ without decreasing the sum-rate $R^{\mathrm{tw}}(\mathbf{B}, \mathbf{D})$ in (3.6) if $\hat{R}_{\mathrm{r}i}(\mathbf{B}_i) > \bar{R}_{j\mathrm{r}}(\mathbf{D}_j)$. Therefore, the constraint (4.5a) is necessary. Subject to (4.5a), $R^{\mathrm{tw}}(\mathbf{B}, \mathbf{D})$ in (3.6) can be written as $\min\{R^{\mathrm{ma}}(\mathbf{D}), \hat{R}_{\mathrm{r}1}(\mathbf{B}_1) + \hat{R}_{\mathrm{r}2}(\mathbf{B}_2)\}/2$. Using the fact that $R^{\mathrm{ma}}(\mathbf{D}) < \hat{R}_{\mathrm{r}1}(\mathbf{B}_1) + \hat{R}_{\mathrm{r}2}(\mathbf{B}_2)$ when $\hat{R}_{\mathrm{r}1}(\mathbf{B}_1) = \bar{R}_{2\mathrm{r}}(\mathbf{D}_2)$ and $\hat{R}_{\mathrm{r}2}(\mathbf{B}_2) = \bar{R}_{1\mathrm{r}}(\mathbf{D}_1)$, it can be shown that the power consumption of at least one source node can be reduced without decreasing $R^{\mathrm{tw}}(\mathbf{B}, \mathbf{D})$ if $R^{\mathrm{ma}}(\mathbf{D}) > R(\mathbf{B}, \mathbf{D})$ while the power consumption of the relay can be reduced without decreasing $R^{\mathrm{tw}}(\mathbf{B}, \mathbf{D})$ if $R^{\mathrm{ma}}(\mathbf{D}) < R(\mathbf{B}, \mathbf{D})$. Thus, the constraint (4.5b) is also necessary.

Considering the constraints (4.5a) and (4.5b), the problem of finding the optimal power allocation becomes nonconvex. Relating (4.2a)-(4.2c) with (4.1a)-(4.1b), the above two constraints (4.5a) and (4.5b) can be rewritten as

$$\lambda_i \geq \mu_j, \forall i \tag{4.6a}$$

$$\sum_i \sum_{k \in \mathcal{I}_i} \log\left(1 + \left(\frac{1}{\lambda_i}\alpha_i(k) - 1\right)^+\right) = \sum_i \sum_{k \in \mathcal{I}_i} \log\left(1 + \left(\frac{1}{\mu_{\mathrm{ma}}}\alpha_i(k) - 1\right)^+\right). \tag{4.6b}$$

It should be noted that the constraints (4.5a) and (4.5b), or equivalently (4.6a) and (4.6b), are not sufficient in general. Due to the intrinsic complexity of the considered problem, it is too complicated to formulate a general sufficient and necessary condition for optimality of the original problem of sum-rate maximization with minimum power consumption. Instead, we will show sufficient and necessary optimality conditions for the equivalent problems in the subcases in which the original problem can be transferred into equivalent convex problems. For other subcases, we will develop important properties based on the above necessary conditions (4.5)

or equivalently (4.6) which can significantly reduce the computational complexity of searching for the optimal solution.

The following lemma that applies for all subcases is introduced for subsequent analysis.

Lemma 4.1: Given $\mathbf{D}_1$ and $\mathbf{D}_2$ with $P_1^{\max} \geq \text{Tr}\{\mathbf{D}_1\} > 0$ and $P_2^{\max} \geq \text{Tr}\{\mathbf{D}_2\} > 0$, if $1/\mu_i > 1/\mu_{\text{ma}} > 1/\mu_j$, then the following two results hold true: 1) $1/\mu_{\text{ma}}(\widetilde{\mathbf{D}}) \leq 1/\mu_j$ where $\widetilde{\mathbf{D}} = [\widetilde{\mathbf{D}}_1, \widetilde{\mathbf{D}}_2]$ with $\widetilde{\mathbf{D}}_i = \mathbf{0}$ and $\widetilde{\mathbf{D}}_j = \mathbf{D}_j$, 2) there exists $t \in [0, 1)$ such that with $\widehat{\mathbf{D}}_i = t\mathbf{D}_i$ and $\widehat{\mathbf{D}}_j = \mathbf{D}_j$, we have $1/\mu_i(\widehat{\mathbf{D}}_i) > 1/\mu_{\text{ma}}(\widehat{\mathbf{D}}) = 1/\mu_j$ where $\widehat{\mathbf{D}} = [\widehat{\mathbf{D}}_1, \widehat{\mathbf{D}}_2]$.

Proof: See Subsection B.1 in Appendix. ■

Lemma 4.1 relates the source nodes transmit strategy $\mathbf{D}$ with the relative water-levels $1/\mu_1, 1/\mu_2$, and $1/\mu_{\text{ma}}$. It shows a range that the relative water-level $1/\mu_{\text{ma}}$ can achieve by fixing $\mathbf{D}_j$ and changing $\mathbf{D}_i$ given that $1/\mu_i > 1/\mu_{\text{ma}} > 1/\mu_j$.

Lemma 4.2: The optimal solution of the network optimization problem has the following property

$$\lambda_j = \mu_i > \mu_{\text{ma}} \quad \text{if } \lambda_i < \lambda_j \text{ or } \mu_i > \mu_{\text{ma}}. \tag{4.7}$$

Proof: See Subsection B.2 in Appendix. ■

Lemma 4.2 develops a property of the optimal solution that follows from the constraints (4.6a) and (4.6b). This property is needed for future analysis.

In the scenario of network optimization, the three nodes aim at finding the optimal matrices $\mathbf{D}$ and $\mathbf{B}$ that minimize $\text{Tr}\{\mathbf{D}_1\} + \text{Tr}\{\mathbf{D}_2\} + \text{Tr}\{\mathbf{B}_1 + \mathbf{B}_2\}$ among all $\mathbf{D}$ and $\mathbf{B}$ that achieve the maximum of the objective function in (4.3). Considering the fact that the optimal $\mathbf{B}$ and $\mathbf{D}$ depend on each other, solving this problem generally involves alternative optimization of $\mathbf{B}$ and $\mathbf{D}$. It is, however, of interest to avoid such alternative process when it is possible due to its high complexity. Next we use an initial power allocation[4] to classify the problem of finding the optimal $\mathbf{B}$ and $\mathbf{D}$ for network optimization into two cases, each with several subcases.

Consider the following initial power allocation of the source nodes and the relay, which decides the maximum achievable sum-rates of the MA and BC phases,

[4] Note that the initial power allocation is not the solution to the considered problem and it is only used for enabling classification.

respectively. The initial power allocation of the source nodes is the solution to the following problem

$$\max_{\mathbf{D}} \quad R^{\mathrm{ma}}(\mathbf{D}) \tag{4.8a}$$

$$\mathbf{s.t.} \quad \mathrm{Tr}\{\mathbf{D}_i\} \leq P_i^{\max}, \forall i \tag{4.8b}$$

which is a power allocation problem on multiple-access channels studied in [70], while the initial power allocation of the relay is to allocate $P_{\mathrm{r}}^{\max}$ on $\alpha_i(k)$'s, $\forall k \in \mathcal{I}_i, \forall i$ based on the waterfilling procedure. Denote the optimal solution of (4.8) as $\mathbf{D}^0 = [\mathbf{D}_1^0, \mathbf{D}_2^0]$ and the water level corresponding to the relay's initial power allocation as $1/\lambda^0$. The case when $R^{\mathrm{ma}}(\mathbf{D}^0) \geq \hat{R}_{\mathrm{r}1}(\lambda^0) + \hat{R}_{\mathrm{r}2}(\lambda^0)$, i.e., when the maximum achievable sum-rate of the MA phase is lager than or equal to that of the BC phase, is denoted as Case I and the case when $R^{\mathrm{ma}}(\mathbf{D}^0) < \hat{R}_{\mathrm{r}1}(\lambda^0) + \hat{R}_{\mathrm{r}2}(\lambda^0)$, i.e., when the maximum achievable sum-rate of the MA phase is less than that of the BC phase, is denoted as Case II. We next study the problem of maximizing $R^{\mathrm{tw}}(\mathbf{B}, \mathbf{D})$ with minimum power consumption and find the optimal power allocation for Cases I and II, respectively, in the following subsections.

4.2.1 Finding the optimal solution in Case I, i.e., $R^{\mathrm{ma}}(\mathbf{D}^0) \geq \hat{R}_{\mathrm{r}1}(\lambda^0) + \hat{R}_{\mathrm{r}2}(\lambda^0)$

Since $R^{\mathrm{ma}}(\mathbf{D}^0) \geq \hat{R}_{\mathrm{r}1}(\lambda^0) + \hat{R}_{\mathrm{r}2}(\lambda^0)$, it can be inferred that $1/\lambda_0 \leq 1/\mu_{\mathrm{ma}}^0$. In this case, the sum-rate $R^{\mathrm{tw}}(\mathbf{B}, \mathbf{D})$ in (3.6) is upper-bounded by the sum-rate $\hat{R}_{\mathrm{r}1}(\lambda^0) + \hat{R}_{\mathrm{r}2}(\lambda^0)$. The following two subcases should be considered separately.

Subcase I-1: The following convex optimization problem is feasible

$$\min_{\mathbf{D}} \quad \mathrm{Tr}\{\mathbf{D}_1\} + \mathrm{Tr}\{\mathbf{D}_2\} \tag{4.9a}$$

$$\mathbf{s.t.} \quad R^{\mathrm{ma}}(\mathbf{D}) \geq \hat{R}_{\mathrm{r}1}(\lambda^0) + \hat{R}_{\mathrm{r}2}(\lambda^0) \tag{4.9b}$$

$$\bar{R}_{1\mathrm{r}}(\mathbf{D}_1) \geq \hat{R}_{\mathrm{r}2}(\lambda^0) \tag{4.9c}$$

$$\bar{R}_{2\mathrm{r}}(\mathbf{D}_2) \geq \hat{R}_{\mathrm{r}1}(\lambda^0) \tag{4.9d}$$

$$\mathrm{Tr}(\mathbf{D}_i) \leq P_i^{\max}, \forall i. \tag{4.9e}$$

In this subcase, the maximum sum-rate $R^{\mathrm{tw}}(\mathbf{B}, \mathbf{D})$ can achieve $\hat{R}_{\mathrm{r}1}(\lambda^0) + \hat{R}_{\mathrm{r}2}(\lambda^0)$. In order to achieve this maximum sum-rate, it is necessary that $\lambda_1 = \lambda_2 = \lambda^0$.

Therefore, the relay should use up all available power $P_{\rm r}^{\max}$ at optimality, and the optimal $\mathbf{B}_i, \forall i$ are equal to $\mathbf{V}_{{\rm r}i}\mathbf{P}_{{\rm r}i}(\lambda^0)\mathbf{V}_{{\rm r}i}^{\rm H}, \forall i$ where $\mathbf{P}_{{\rm r}i}(\lambda_i)$ is given in Section 4.1. As a result, the original problem simplifies to finding the optimal $\mathbf{D}_1$ and $\mathbf{D}_2$ such that $R^{\rm tw}(\mathbf{B}, \mathbf{D})$ achieves $\hat{R}_{\rm r1}(\lambda^0) + \hat{R}_{\rm r2}(\lambda^0)$ with minimum power consumption. Using equations (3.6) and (3.7), it can be shown that a sufficient and necessary condition for $\mathbf{D}$ to be optimal in this subcase is that $\mathbf{D}$ is the optimal solution to the convex optimization problem (4.9). Denoting the optimal solution to the problem (4.9) as $\mathbf{D}^\star = [\mathbf{D}_1^\star, \mathbf{D}_2^\star]$, the total power consumption in this subcase is $P_{\rm r}^{\max} + {\rm Tr}\{\mathbf{D}_1^\star\} + {\rm Tr}\{\mathbf{D}_2^\star\}$.

It can be seen that the optimal solution of $\mathbf{B}$ and $\mathbf{D}$ in the above specific subcase, i.e., Subcase I-1, as described above satisfies the general constraint (4.6a), or equivalently (4.5a), for the original problem since the constraints (4.9c) and (4.9d) are considered in the problem (4.9). It can also be shown that the above optimal solution in Subcase I- 1 also satisfies the general constraint (4.6b), or equivalently (4.5b), for the original problem as stated in the following theorem.

Theorem 4.1: The optimal solution in Subcase I-1 satisfies $\mu_{\rm ma}^\star = \lambda^0$, and thereby satisfies (4.6b) given that $\lambda_1 = \lambda_2 = \lambda^0$ at optimality.

Proof: See Subsection B.3 in Appendix.

Considering the constraints (4.9b)-(4.9e), it can be seen that the problem (4.9) is feasible if and only if the following problem

$$\begin{aligned}
\max_{\mathbf{D}} \quad & \bar{R}_{j{\rm r}}(\mathbf{D}_j) && \text{(4.10a)}\\
\mathbf{s.t.} \quad & \bar{R}_{i{\rm r}}(\mathbf{D}_i) \geq \hat{R}_{{\rm r}j}(\lambda^0) && \text{(4.10b)}\\
& R^{\rm ma}(\mathbf{D}) \geq \hat{R}_{\rm r1}(\lambda^0) + \hat{R}_{\rm r2}(\lambda^0) && \text{(4.10c)}\\
& {\rm Tr}(\mathbf{D}_1) \leq P_1^{\max} && \text{(4.10d)}\\
& {\rm Tr}(\mathbf{D}_2) \leq P_2^{\max} && \text{(4.10e)}
\end{aligned}$$

is feasible and its optimal solution, denoted as $\mathbf{D}^*$, satisfies $\bar{R}_{j{\rm r}}(\mathbf{D}_j^*) \geq \hat{R}_{{\rm r}i}(\lambda^0), \forall j$.[5] However, it is possible that $\bar{R}_{j{\rm r}}(\mathbf{D}_j^*) < \hat{R}_{{\rm r}i}(\lambda^0)$ for some i and j. It is also possible that the problem (4.10) is not even feasible. In both of the above two situations the

[5]Note that if $\bar{R}_{j{\rm r}}(\mathbf{D}_j^*) \geq \hat{R}_{{\rm r}i}(\lambda^0)$ for $i = 1, j = 2$ in (4.10) then it also holds that $\bar{R}_{j{\rm r}}(\mathbf{D}_j^*) \geq \hat{R}_{{\rm r}i}(\lambda^0)$ for $i = 2, j = 1$ and vice versa.

problem (4.9) is infeasible. This leads to the second subcase of Case I.

Subcase I-2: The problem (4.9) is infeasible.

Unlike Subcase I-1, the maximum sum-rate $R^{\mathrm{tw}}(\mathbf{B},\mathbf{D})$ in this subcase cannot achieve $\hat{R}_{\mathrm{r}1}(\lambda^0)+\hat{R}_{\mathrm{r}2}(\lambda^0)$. As mentioned above, there are two possible situations when the problem (4.9) is infeasible: (i) $\bar{R}_{j\mathrm{r}}(\mathbf{D}_j^*)<\hat{R}_{\mathrm{r}i}(\lambda^0),\forall j$ and (ii) the problem (4.10) is infeasible for specific valued of i and j. Using Lemma 3.1 in Chapter 3 and the fact that $R^{\mathrm{ma}}(\mathbf{D}^0)\geq\hat{R}_{\mathrm{r}1}(\lambda^0)+\hat{R}_{\mathrm{r}2}(\lambda^0)$ for Case I, it can be shown that if the problem (4.10) is infeasible for specific values of i and j, then it is feasible (but $\bar{R}_{j\mathrm{r}}(\mathbf{D}_j^*)<\hat{R}_{\mathrm{r}i}(\lambda^0)$) when the values of i and j are switched. Therefore, the problem (4.9) is infeasible if and only if there exists at least one specific value of j in $\{1,2\}$ such that problem (4.10) is feasible but $\bar{R}_{j\mathrm{r}}(\mathbf{D}_j^*)<\hat{R}_{\mathrm{r}i}(\lambda^0)$. Denote this specific value of j as l and denote the corresponding i as $\bar{l}$. It infers, based on the definitions (4.2a)-(4.2c), that $1/\mu_l<1/\lambda^0$ whenever $1/\mu_{\mathrm{ma}}\geq 1/\lambda^0$ and $1/\mu_{\bar{l}}\geq 1/\lambda^0$. As a result, whenever $1/\mu_{\mathrm{ma}}\geq 1/\lambda^0$, or equivalently, $R^{\mathrm{ma}}(\mathbf{D})\geq\hat{R}_{\mathrm{r}1}(\lambda^0)+\hat{R}_{\mathrm{r}2}(\lambda^0)$, the sum-rate $R^{\mathrm{tw}}(\mathbf{B},\mathbf{D})$ is bounded by $\hat{R}_{\mathrm{r}1}(\lambda_1)+\hat{R}_{\mathrm{r}2}(\lambda_2)$ according to equation (3.6), which is less than $\hat{R}_{\mathrm{r}1}(\lambda^0)+\hat{R}_{\mathrm{r}2}(\lambda^0)$ when $1/\mu_l<1/\lambda^0$ (according to the constraint (4.6a) and Lemma 3.2 in Chapter 3). Moreover, whenever $1/\mu_{\mathrm{ma}}<1/\lambda^0$, or equivalently, $R^{\mathrm{ma}}(\mathbf{D})<\hat{R}_{\mathrm{r}1}(\lambda^0)+\hat{R}_{\mathrm{r}2}(\lambda^0)$, the sum-rate $R^{\mathrm{tw}}(\mathbf{B},\mathbf{D})$ is bounded by $R^{\mathrm{ma}}(\mathbf{D})$ according to equation (3.6), which is also less than $\hat{R}_{\mathrm{r}1}(\lambda^0)+\hat{R}_{\mathrm{r}2}(\lambda^0)$. Therefore, the maximum sum-rate $R^{\mathrm{tw}}(\mathbf{B},\mathbf{D})$ in this subcase always cannot achieve $\hat{R}_{\mathrm{r}1}(\lambda^0)+\hat{R}_{\mathrm{r}2}(\lambda^0)$.

With the above denotation of l and $\bar{l}$, the following theorem characterizes the optimal solution in this subcase.

Theorem 4.2: Denote the optimal $\mathbf{D}_i$ in Subcase I-2 as $\mathbf{D}_i^*,\forall i$ and the optimal λ_i as $\lambda_i^*,\forall i$. The optimal strategies for the source nodes and the relay satisfy the following properties:

1. $\min\limits_i\{1/\mu_i^*\}<1/\mu_{\mathrm{ma}}^*<1/\lambda^0$;

2. The relay uses full power $P_{\mathrm{r}}^{\max}$;

3. $\mathbf{D}^*$ maximizes $\min_i\{1/\mu_i\}$ among all $\mathbf{D}$'s that satisfy

$$R^{\mathrm{ma}}(\mathbf{D}) \geq R^{\mathrm{ma}}(\mathbf{D}^*) \tag{4.11a}$$

$$\mathrm{Tr}(\mathbf{D}_i) \leq P_i^{\max}, \forall i \tag{4.11b}$$

4. $1/\mu_l^* < 1/\mu_{\bar{l}}^*$.

Proof: Please see Subsection B.4 in Appendix.

While the original problem cannot be simplified into an equivalent form in this subcase, the properties in Theorem 4.2 help to significantly reduce the complexity of searching for the optimal solution by narrowing down the set of qualifying power allocations. Denote $\mathbf{D}_l^{\mathrm{l}}$ as the $\mathbf{D}_l$ that maximizes $\bar{R}_{l\mathrm{r}}(\mathbf{D}_l)$ subject to the constraints $\mu_l \geq \mu_{\mathrm{ma}}$ and $\mathrm{Tr}\{\mathbf{D}_l\} \leq P_l^{\max}$ and denote μ_l^{l} as the corresponding μ_l. According to Theorem 4.2, if $\hat{R}_{\mathrm{r}\bar{l}}(\lambda_{\bar{l}}^\dagger) + \hat{R}_{\mathrm{r}l}(\lambda_l^\dagger) \leq R^{\mathrm{ma}}(\mathbf{D}^\dagger)$, where

$$\lambda_{\bar{l}}^\dagger = \mu_l^{\mathrm{l}} \tag{4.12a}$$

$$\mathrm{Tr}\{\mathbf{P}_{\mathrm{r}\bar{l}}(\lambda_{\bar{l}}^\dagger)\} + \mathrm{Tr}\{\mathbf{P}_{\mathrm{r}l}(\lambda_l^\dagger)\} = P_{\mathrm{r}}^{\max} \tag{4.12b}$$

and $\mathbf{D}^\dagger$ is the optimal solution of the following problem

$$\max_{\mathbf{D}} \quad R^{\mathrm{ma}}(\mathbf{D}) \tag{4.13a}$$

$$\mathbf{s.t.} \quad \bar{R}_{l\mathrm{r}}(\mathbf{D}_l) \geq \bar{R}_{l\mathrm{r}}(\mathbf{D}_l^{\mathrm{l}}) \tag{4.13b}$$

$$\mathrm{Tr}(\mathbf{D}_i) \leq P_i^{\max}, \forall i \tag{4.13c}$$

then the maximum achievable sum-rate in Subcase I-2 is $\hat{R}_{\mathrm{r}\bar{l}}(\lambda_{\bar{l}}^\dagger) + \hat{R}_{\mathrm{r}l}(\lambda_l^\dagger)$ (according to Lemma 3.2 in Chapter 3), the optimal $\mathbf{B}_i, \forall i$ in this subcase is given by $\mathbf{B}_i = \mathbf{V}_{\mathrm{r}i}\mathbf{P}_{\mathrm{r}i}(\lambda_i^\dagger)\mathbf{V}_{\mathrm{r}i}^{\mathrm{H}}$, and the optimal $\mathbf{D}$ is the solution to the following power minimization problem

$$\min_{\mathbf{D}} \quad \mathrm{Tr}\{\mathbf{D}_1\} + \mathrm{Tr}\{\mathbf{D}_2\} \tag{4.14a}$$

$$\mathbf{s.t.} \quad R^{\mathrm{ma}}(\mathbf{D}) \geq \sum_i \hat{R}_{\mathrm{r}i}(\lambda_i^\dagger) \tag{4.14b}$$

$$\bar{R}_{\bar{l}\mathrm{r}}(\mathbf{D}_{\bar{l}}) \geq \hat{R}_{\mathrm{r}l}(\lambda_l^\dagger) \tag{4.14c}$$

$$\bar{R}_{l\mathrm{r}}(\mathbf{D}_l) \geq \hat{R}_{\mathrm{r}\bar{l}}(\lambda_{\bar{l}}^\dagger) \tag{4.14d}$$

$$\mathrm{Tr}(\mathbf{D}_i) \leq P_i^{\max}, \forall i. \tag{4.14e}$$

If $\hat{R}_{\mathrm{r}\bar{l}}(\lambda_{\bar{l}}^{\dagger})+\hat{R}_{\mathrm{r}l}(\lambda_{l}^{\dagger}) > R^{\mathrm{ma}}(\mathbf{D}^{\dagger})$, denote the objective $R^{\mathrm{tw}}(\mathbf{B},\mathbf{D})$ as R^{obj}. According to Theorem 4.2, the optimal solution can be found by maximizing R^{obj} so that it can be achieved by both $R^{\mathrm{ma}}(\mathbf{D})$ and $\sum_i \hat{R}_{\mathrm{r}i}(\lambda_i)$ subject to the following two constraints: 1) $1/\lambda_{\bar{l}} = 1/\tilde{\mu}_l$ which is obtained according to Lemma 4.2, Properties 1 and 4 of Theorem 4.2; 2) $1/\lambda_l$ is obtained by waterfilling the remaining power on $\alpha_l(k), \forall k \in \mathcal{I}_l$ (Property 2 of Theorem 4.2), where $1/\tilde{\mu}_l = 1/\mu_l(\tilde{\mathbf{D}})$ is the optimal value of the objective function in the following optimization problem (Property 3 of Theorem 4.2)

$$\max_{\mathbf{D}} \quad \frac{1}{\mu_l} \tag{4.15a}$$

$$\mathbf{s.t.} \quad R^{\mathrm{ma}}(\mathbf{D}) \geq R^{\mathrm{obj}} \tag{4.15b}$$

$$\mathrm{Tr}(\mathbf{D}_i) \leq P_i^{\max}, \forall i. \tag{4.15c}$$

Here $\tilde{\mathbf{D}}$ denotes the optimal solution of (4.15) for the given R^{obj}. Since maximizing $1/\mu_l$ is equivalent to maximizing $\bar{R}_{l\mathrm{r}}(\mathbf{D}_l)$, the objective function of the above problem can be substituted by $\bar{R}_{l\mathrm{r}}(\mathbf{D}_l)$ and $1/\tilde{\mu}_l$ can be obtained from the optimal value of $\bar{R}_{l\mathrm{r}}(\mathbf{D}_l)$ in the above problem using (4.2a) or (4.2b). As mentioned at the beginning of Subcase I-2, the optimal $R^{\mathrm{tw}}(\mathbf{B},\mathbf{D})$ is less than $\sum_i \hat{R}_{\mathrm{r}i}(\lambda^0)$. Therefore, starting from the point by setting $R^{\mathrm{obj}} = \sum_i \hat{R}_{\mathrm{r}i}(\lambda^0)$, we can adjust R^{obj} as follows, to achieve the optimal $R^{\mathrm{tw}}(\mathbf{B},\mathbf{D})$. We first solve the following problem given R^{obj}

$$\max_{\mathbf{D}} \quad \bar{R}_{l\mathrm{r}}(\mathbf{D}_l) \tag{4.16a}$$

$$\mathbf{s.t.} \quad R^{\mathrm{ma}}(\mathbf{D}) \geq R^{\mathrm{obj}} \tag{4.16b}$$

$$\mathrm{Tr}(\mathbf{D}_i) \leq P_i^{\max}, \forall i \tag{4.16c}$$

to get the optimal $\tilde{\mathbf{D}}$ for the given R^{obj} and obtain the resulting $1/\tilde{\mu}_l = 1/\mu_l(\tilde{\mathbf{D}})$. Then, we set $1/\lambda_{\bar{l}} = 1/\tilde{\mu}_l$ and allocate all the remaining power on $\alpha_l(k)$'s, $\forall k \in \mathcal{I}_l$. If the resulting $\sum_i \hat{R}_{\mathrm{r}i}(\lambda_i)$ is less than R^{obj}, it infers that R^{obj} should be decreased in (4.16) and the above process should be repeated. On the other hand, if the resulting $\sum_i \hat{R}_{\mathrm{r}i}(\lambda_i)$ is larger than R^{obj}, then R^{obj} should be increased in (4.16) and the above process should be repeated. The optimal solution is found when the resulting

$\sum_i \hat{R}_{\mathrm{r}i}(\lambda_i)$ is equal to R^{obj}. With an appropriate step size of increasing/decreasing R^{obj}, R^{obj} in the above procedure converges to the optimal $R^{\mathrm{tw}}(\mathbf{B}, \mathbf{D})$.

After obtaining the optimal R^{obj}, $1/\tilde{\mu}_l$ and $\lambda_{\bar{l}}$, the source nodes solve the problem of power minimization, which is

$$\min_{\mathbf{D}} \quad \mathrm{Tr}\{\mathbf{D}_1\} + \mathrm{Tr}\{\mathbf{D}_2\} \tag{4.17a}$$

$$\mathbf{s.t.} \quad R^{\mathrm{ma}}(\mathbf{D}) \geq R^{\mathrm{obj}} \tag{4.17b}$$

$$\bar{R}_{\bar{l}\mathrm{r}}(\mathbf{D}_{\bar{l}}) \geq \hat{R}_{\mathrm{r}l}(\lambda_l) \tag{4.17c}$$

$$\bar{R}_{l\mathrm{r}}(\mathbf{D}_l) \geq \hat{R}_{\mathrm{r}\bar{l}}(\lambda_{\bar{l}}) \tag{4.17d}$$

$$\mathrm{Tr}(\mathbf{D}_i) \leq P_i^{\max}, \forall i. \tag{4.17e}$$

However, it can be shown that if $\bar{R}_{l\mathrm{r}}(\tilde{\mathbf{D}}_l)$ is not the maximum that $\bar{R}_{l\mathrm{r}}(\mathbf{D}_l)$ can achieve subject to the constraint (4.15c) (without the constraint (4.15b)), then $\mathbf{B}$ and $\mathbf{D}$ remain the same after solving the above problem.

Using Property 2 of Theorem 4.2, it can be seen from (4.14) and (4.17) that the minimization of total power consumption becomes the minimization of the source node power consumption in Subcase I-2 since the relay always needs to consume all its available power for achieving optimality.

The complete procedure of finding the optimal solution in Case I is summarized in the algorithm in Table 4.1. The algorithm finds the optimal solution either in one shot (Steps 1 and 2) or through a bisection search for the optimal R^{obj} (Steps 3 to 5). Denoting $\Delta = R^{\max} - R^{\min}$, the worst case number of iterations in the bisection search is $\log(\Delta/\epsilon)$. Within each iteration, a convex problem, i.e., problem (4.16), is solved followed by a simple waterfilling procedure which has linear complexity for the given R^{obj}. Therefore, the complexity of the proposed algorithm is low.

Subcases I-1 and I-2 cover all possible situations for Case I that $R^{\mathrm{ma}}(\mathbf{D}^0) \geq \hat{R}_{\mathrm{r}1}(\lambda^0) + \hat{R}_{\mathrm{r}2}(\lambda^0)$.

4.2.2 Finding the optimal solution in Case II, i.e., $R^{\mathrm{ma}}(\mathbf{D}^0) < \hat{R}_{\mathrm{r}1}(\lambda^0) + \hat{R}_{\mathrm{r}2}(\lambda^0)$

Since $R^{\mathrm{ma}}(\mathbf{D}^0) < \hat{R}_{\mathrm{r}1}(\lambda^0) + \hat{R}_{\mathrm{r}2}(\lambda^0)$, it can be seen using (4.1a), (4.1b), and (4.2c) that $1/\lambda_0 > 1/\mu_{\mathrm{ma}}^0$. The following four subcases are possible.

Table 4.1: Algorithm for finding the optimal solution for Case I.

1. Check if the problem (4.9) is feasible. If yes, find the optimal $\mathbf{D}$ from the problem (4.9). The optimal $\mathbf{B}$ is given by $\mathbf{B}_i = \mathbf{V}_{\mathrm{r}i}\mathbf{P}_{\mathrm{r}i}(\lambda^0)\mathbf{V}_{\mathrm{r}i}^{\mathrm{H}}, \forall i$. Otherwise, specify l and $\bar{l}$ so that the problem (4.10) is feasible but $\bar{R}_{l\mathrm{r}}(\mathbf{D}_l^*) < \hat{R}_{\mathrm{r}\bar{l}}(\lambda^0)$ and proceed to Step 2.
2. Obtain $\mathbf{D}_l^{\mathrm{l}}$ and μ_l^{l}. Calculate $\lambda_i^{\dagger}, \forall i$ using (4.12). Check if $\sum_i \hat{R}_{\mathrm{r}i}(\lambda_i^{\dagger}) \leq R^{\mathrm{ma}}(\mathbf{D}^{\dagger})$. If yes, the optimal $\mathbf{B}$ is given by $\mathbf{B}_i = \mathbf{V}_{\mathrm{r}i}\mathbf{P}_{\mathrm{r}i}(\lambda_i^{\dagger})\mathbf{V}_{\mathrm{r}i}^{\mathrm{H}}, \forall i$. Find the optimal $\mathbf{D}$ from (4.14). Otherwise, proceed to Step 3.
3. Set $R^{\max} = \sum_i \hat{R}_{\mathrm{r}i}(\lambda^0)$ and $R^{\min} = 0$. Initialize $R^{\mathrm{obj}} = R^{\max}$ and proceed to Step 4.
4. Solve the problem (4.16) and obtain $\mathbf{D}$ and $1/\tilde{\mu}_l$. Set $1/\lambda_{\bar{l}} = 1/\tilde{\mu}_l$. Allocate all the remaining power on $\alpha_l(k)$'s, $\forall k \in \mathcal{I}_l$ using waterfilling and obtain $1/\lambda_l$. Check if $\lvert \sum_i \hat{R}_{\mathrm{r}i}(\lambda_i) - R^{\mathrm{ma}}(\mathbf{D}) \rvert < \epsilon$, where ϵ is the positive tolerance. If yes, proceed to Step 6 with R^{obj} and $\lambda_i, \forall i$. Otherwise, proceed to Step 5.
5. If $R^{\mathrm{ma}}(\mathbf{D}) - \sum_i \hat{R}_{\mathrm{r}i}(\lambda_i) > \epsilon$, set $R^{\max} = R^{\mathrm{obj}}$. If $\sum_i \hat{R}_{\mathrm{r}i}(\lambda_i) - R^{\mathrm{ma}}(\mathbf{D}) > \epsilon$, set $R^{\min} = R^{\mathrm{obj}}$. Let $R^{\mathrm{obj}} = (R^{\max} + R^{\min})/2$ and go back to Step 4.
6. Solve the power minimization problem (4.17). Output $\mathbf{D}$ and $\mathbf{B}_i = \mathbf{V}_{\mathrm{r}i}\mathbf{P}_{\mathrm{r}i}(\lambda_i)\mathbf{V}_{\mathrm{r}i}^{\mathrm{H}}, \forall i$.

Subcase II-1: $1/\mu_{\mathrm{ma}}^0 \leq \min\{1/\mu_1^0, 1/\mu_2^0\}$. In this subcase, the maximum sum-rate $R^{\mathrm{tw}}(\mathbf{B}, \mathbf{D})$ is bounded by $R^{\mathrm{ma}}(\mathbf{D}^0)$. The optimal $\mathbf{D}$ is $\mathbf{D}^0$, and consequently both source nodes use all their available power at optimality. It can be seen that a sufficient and necessary condition for $\mathbf{B}$ to be optimal in this subcase is that $\mathbf{B}$ is the optimal solution to the following convex optimization problem

$$\min_{\mathbf{B}} \quad \mathrm{Tr}\{\mathbf{B}_1 + \mathbf{B}_2\} \tag{4.18a}$$

$$\mathbf{s.t.} \quad \hat{R}_{\mathrm{r}1}(\mathbf{B}_1) + \hat{R}_{\mathrm{r}2}(\mathbf{B}_2) \geq R^{\mathrm{ma}}(\mathbf{D}^0). \tag{4.18b}$$

The solution of (4.18) can be given in closed-form as $\mathbf{B}_i = \mathbf{V}_{\mathrm{r}i}\mathbf{P}_{\mathrm{r}i}(\mu_{\mathrm{ma}}^0)\mathbf{V}_{\mathrm{r}i}^{\mathrm{H}}, \forall i$.

Subcase II-2: there exist l and $\bar{l}$ such that $1/\mu_l^0 \leq 1/\mu_{\mathrm{ma}}^0 < 1/\mu_{\bar{l}}^0 \leq 1/\lambda^0$.[6] In this subcase, the maximum achievable $R^{\mathrm{tw}}(\mathbf{B}, \mathbf{D})$ is also $R^{\mathrm{ma}}(\mathbf{D}^0)$. Therefore, the

[6]For the consistency of denotation, the constrained indices $l \in \{1, 2\}$ and $\bar{l} \in \{1, 2\}\backslash\{l\}$ are also used here in Case II. However, it should be noted that they are not determined by the same constraint as in Case I.

optimal $\mathbf{D}$ is $\mathbf{D}^0$ and both source nodes use all their available power at optimality. It can be shown that a sufficient and necessary condition for $\mathbf{B}$ to be optimal in this subcase is that $\mathbf{B}$ is the optimal solution to the following convex optimization problem

$$\min_{\mathbf{B}} \quad \mathrm{Tr}\{\mathbf{B}_1 + \mathbf{B}_2\} \tag{4.19a}$$

$$\mathbf{s.t.} \quad \hat{R}_{\mathrm{r}1}(\mathbf{B}_1) + \hat{R}_{\mathrm{r}2}(\mathbf{B}_2) \geq R^{\mathrm{ma}}(\mathbf{D}^0) \tag{4.19b}$$

$$\hat{R}_{\mathrm{r}\bar{l}}(\mathbf{B}_{\bar{l}}) = \bar{R}_{l\mathrm{r}}(\mathbf{D}_l^0). \tag{4.19c}$$

The solution of (4.19) can also be expressed in closed-form. The optimal $\mathbf{B}_{\bar{l}}$ is given by $\mathbf{B}_{\bar{l}} = \mathbf{V}_{\mathrm{r}\bar{l}}\mathbf{P}_{\mathrm{r}\bar{l}}(\mu_l^0)\mathbf{V}_{\mathrm{r}\bar{l}}^{\mathrm{H}}$ and the optimal $\mathbf{B}_l$ is given by $\mathbf{B}_l = \mathbf{V}_{\mathrm{r}l}\mathbf{P}_{\mathrm{r}l}(\lambda_l)\mathbf{V}_{\mathrm{r}l}^H$, where λ_l satisfies $\hat{R}_{\mathrm{r}l}(\lambda_l) = R^{\mathrm{ma}}(\mathbf{D}^0) - \bar{R}_{l\mathrm{r}}(\mathbf{D}_l^0)$.

Subcase II-3: there exist l and $\bar{l}$ such that $1/\mu_l^0 \leq 1/\mu_{\mathrm{ma}}^0 < 1/\lambda^0 < 1/\mu_{\bar{l}}^0$ and there exists λ_l such that

$$\hat{R}_{\mathrm{r}l}(\lambda_l) \geq R^{\mathrm{ma}}(\mathbf{D}^0) - \bar{R}_{l\mathrm{r}}(\mathbf{D}_l^0) \tag{4.20a}$$

$$\mathrm{Tr}\{\mathbf{P}_{\mathrm{r}l}(\lambda_l)\} \leq P_{\mathrm{r}}^{\max} - \mathrm{Tr}\{\mathbf{P}_{\mathrm{r}\bar{l}}(\mu_l^0)\}. \tag{4.20b}$$

The optimal solutions of $\mathbf{B}$ and $\mathbf{D}$ in this subcase are the same as those given in Subcase II-2.

In the above three subcases, the maximum achievable $R^{\mathrm{tw}}(\mathbf{B}, \mathbf{D})$ is $R^{\mathrm{ma}}(\mathbf{D}^0)$. Therefore, the original problem of maximizing $R^{\mathrm{tw}}(\mathbf{B}, \mathbf{D})$ with minimum total power consumption in the network simplifies to the problem that the relay uses minimum power consumption to achieve the BC phase sum-rate $\hat{R}_{\mathrm{r}1}(\mathbf{B}_1) + \hat{R}_{\mathrm{r}2}(\mathbf{B}_2)$ that is equal to $R^{\mathrm{ma}}(\mathbf{D}^0)$.

Subcase II-4: there exist l and $\bar{l}$ such that $1/\mu_l^0 \leq 1/\mu_{\mathrm{ma}}^0 < 1/\lambda^0 < 1/\mu_{\bar{l}}^0$ and there is no λ_l that satisfies the conditions in (4.20). In this subcase, the maximum $R(\mathbf{B}, \mathbf{D})$ cannot achieve $R^{\mathrm{ma}}(\mathbf{D}^0)$ although $R^{\mathrm{ma}}(\mathbf{D}^0) < \hat{R}_{\mathrm{r}1}(\lambda^0) + \hat{R}_{\mathrm{r}2}(\lambda^0)$.

Theorem 4.3: Denote the optimal $\mathbf{D}_i$ as $\mathbf{D}_i^*, \forall i$ and the optimal λ_i as $\lambda_i^*, \forall i$. In Subcase II-4, the optimal strategies for the source nodes and the relay satisfy the following properties:

1. $\min_i\{1/\mu_i^*\} < 1/\mu_{\mathrm{ma}}^* < 1/\mu_{\mathrm{ma}}^0$;

2. Properties 2-4 in Theorem 4.2 also apply for Subcase II-4.

Proof: See Subsection B.5 in Appendix.

According to Theorem 4.3, the original problem of maximizing $R^{\mathrm{tw}}(\mathbf{B}, \mathbf{D})$ with minimum total power consumption becomes the problem of finding the maximum achievable $R^{\mathrm{tw}}(\mathbf{B}, \mathbf{D})$ with the relay using all its available power and the source nodes using minimum power. From Theorem 4.3, it can be seen that the optimal solutions in the Subcases I-2 and II-4 share very similar properties. There is also an intuitive way to understand the similarity. Although Subcases I-2 and II-4 are classified to opposite cases according to the initial power allocation, it is the same for both of them that $R(\mathbf{B}, \mathbf{D})$ cannot achieve $R^{\mathrm{ma}}(\mathbf{D}^0)$. As a result, the relay needs to use as much power as possible and the source nodes need to decrease $R^{\mathrm{ma}}(\mathbf{D})$ from $R^{\mathrm{ma}}(\mathbf{D}^0)$ until the maximum $R(\mathbf{B}, \mathbf{D})$ can achieve $R^{\mathrm{ma}}(\mathbf{D})$. This similarity leads to the common properties of the above two subcases. Moreover, due to this similarity between Theorems 2 and 3, Steps 2 to 6 of the algorithm in Table 4.1 can be used to derive the optimal solution in Subcase II-4 if the part of $R^{\max} = \sum_i \hat{R}_{ri}(\lambda^0)$ in Step 3 is substituted by $R^{\max} = R^{\mathrm{ma}}(\mathbf{D}^0)$.

Concluding Cases I and II, the complete procedure of deriving the optimal solution to the problem of sum-rate maximization with minimum total transmission power for the scenario of network optimization is summarized in Table 4.2.

4.2.3 Discussion: efficiency and the effect of asymmetry

In the previous two subsections, we have found the solutions of the network optimization problem for different subcases. Given these solutions, the subcases can now be compared and related to each other for more insights.

The solutions found in all subcases are *optimal* in the sense that they achieve the maximum achievable sum-rate with the minimum possible total power consumption. However, the optimal solutions in different subcases may not be equally good from another viewpoint which is power efficiency at the relay and the source nodes. Specifically, although the power allocation of the source nodes and the relay jointly maximize the sum-rate of the TWR over the MA and BC phases at optimality, the power allocation of these nodes may not be optimal in their individual phase

Table 4.2: Summary of the overall algorithm for network optimization.

1. **Initial power allocation**. The source nodes solve the MA sum-rate maximization problem (4.8) and obtain $\mathbf{D}^0$, $\bar{R}_{i\mathrm{r}}(\mathbf{D}_i^0), \forall i$, and $R^{\mathrm{ma}}(\mathbf{D}^0)$. The relay obtains λ^0 and $\hat{R}_{\mathrm{r}i}(\lambda^0), \forall i$.
2. **Determining the cases**. Check if $R^{\mathrm{ma}}(\mathbf{D}^0) \geq \sum_i \hat{R}_{\mathrm{r}i}(\lambda^i)$. If yes, proceed to Step 3. Otherwise, proceed to Step 4.
3. **Case I**. Determine the subcase based on μ_1^0, μ_2^0, μ_{ma}^0, and λ^0. For Subcase I-1, the relay's optimal strategy is $\mathbf{B}_i = \mathbf{V}_{\mathrm{r}i}\mathbf{P}_{\mathrm{r}i}(\lambda^0)\mathbf{V}_{\mathrm{r}i}^{\mathrm{H}}$ while the source nodes solve problem (4.9) for transmission power minimization. For Subcase I-2, use Steps 2 to 6 of the algorithm in Table 4.1 for deriving the optimal strategies for both the source nodes and the relay.
4. **Case II**. Determine the subcase based on μ_1^0, μ_2^0, μ_{ma}^0, and λ^0. For Subcases II-1, II-2, and II-3, the optimal strategy for source i is $\mathbf{D}_i^0$ and the relay minimizes its transmission power via solving the problems (4.18) or (4.19). For Subcase II-4, substitute $R^{\max} = \sum_i \hat{R}_{\mathrm{r}i}(\lambda^0)$ in Step 3 of Table 4.1 by $R^{\max} = R^{\mathrm{ma}}(\mathbf{D}^0)$ and use Steps 2 to 6 of the algorithm in Table 4.1 for finding the optimal strategies for both the source nodes and the relay.

of transmission, which is MA phase for the source nodes and BC phase for the relay. In fact, the power allocations in the two phases have to compromise with each other in order to achieve optimality over two phases. It is so because of the rate balancing constraints (4.5a) and (4.5b). It infers that there is a cost of coordinating the relay and source nodes to achieve optimality over two phases. This cost can be very different depending on the specific subcase. In order to show the difference in this cost, we use the metric *efficiency* defined next. A given power allocation of the relay (source nodes) is considered as *efficient* if it maximizes the BC (MA) phase sum-rate with the actual power consumption of this power allocation. For example, if the power allocation of the relay consumes the power of $P_{\mathrm{r}} \leq P_{\mathrm{r}}^{\max}$ at optimality and achieves sum-rate R^{bc} in the BC phase, then this power allocation is efficient if R^{bc} is the maximum achievable sum-rate in the BC phase with power consumption P_{r}. It is inefficient otherwise. It can be shown that the chance that the optimal power allocation is efficient for both the relay and the source nodes is small (such case happens in Subcase II-1 and possibly Subcases I-1). Therefore, a

joint power allocation of the relay and source nodes is considered to be inefficient if it is inefficient for both the relay and the source nodes, and it is considered to be efficient otherwise. The following conclusions can be drawn for the scenario of network optimization.

First, it can be shown that the optimal power allocation is efficient in Subcase I-1 and generally inefficient in Subcase I-2. Specifically, the optimal power allocation of the relay is always efficient in Subcase I-1 while the optimal power allocation of the source nodes can be either efficient or inefficient. In contrast, the optimal power allocation of the relay is always inefficient in Subcase I-2 while the optimal power allocation of the source nodes is also inefficient in general. For Case II, the optimal power allocation is efficient in Subcases II-1, II-2, and II-3 and generally inefficient in Subcase II-4. Specifically, the optimal power allocation of the source nodes is efficient in Subcases II-1, II-2, and II-3 and generally inefficient in Subcase II-4 while the optimal power allocation of the relay is efficient in Subcase II-1 and inefficient in Subcases II-2, II-3, and II-4.

Second, the optimal power allocation in Subcase I-1 achieves $\hat{R}_{\mathrm{r1}}(\lambda^0)+\hat{R}_{\mathrm{r2}}(\lambda^0)$. In this subcase, the relay uses its full power and achieves the maximum achievable BC phase sum-rate. The source nodes minimize their power consumption while achieving the maximum sum-rate and in general they do not use up all their available power at optimality. Unlike Subcase I-1, both source nodes may use up their available power in Subcase I-2 while the achieved sum-rate is smaller than $\hat{R}_{\mathrm{r1}}(\lambda^0) + \hat{R}_{\mathrm{r2}}(\lambda^0)$. Similarly, the optimal power allocation in Subcases II-1, II-2, and II-3 achieves $R^{\mathrm{ma}}(\mathbf{D}^0)$ with the source nodes using their full power while the relay does not necessarily use up its available power. In contrast, the optimal power allocation in Subcase II-4 consumes all the available power of the relay while the achieved sum-rate is smaller than $R^{\mathrm{ma}}(\mathbf{D}^0)$. Therefore, it can be seen that for Subcase I-1 and Subcases II-1, II-2, and II-3, in which the optimal power allocation is efficient, either the maximum possible sum-rate of the MA phase or that of the BC phase can be achieved at optimality. Moreover, the source nodes and the relay generally do not both use up their available power. In Subcases I-2 and II-4, in which the optimal power allocation is inefficient, the achieved sum-rate is however

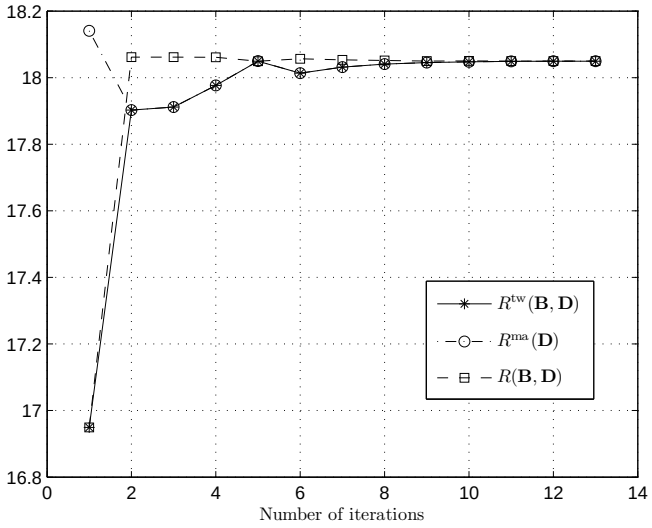
smaller than either the maximum possible sum-rate of the MA phase or that of the BC phase, while it is possible that all nodes use up their available power.

Third, it can be shown for Case I that the difference between $\max_i\{1/\mu_i^0\}$ and $\min_i\{1/\mu_i^0\}$ increases in general as the subcase changes from Subcase I-1 to Subcase I-2. Similar result can be observed in Case II. As the subcase changes from Subcase II-1, via Subcases II-2 and II-3, to Subcase II-4, the difference between $\max_i\{1/\mu_i^0\}$ and $\min_i\{1/\mu_i^0\}$ increases.

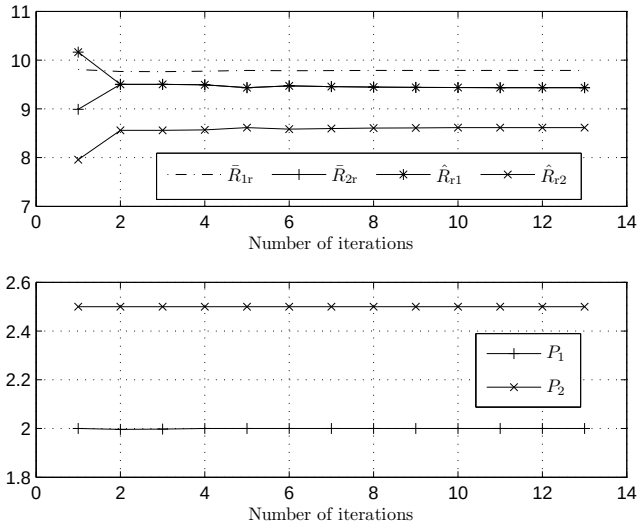
Last, from the definitions of $\mu_i^0, \forall i$, it can be seen that large difference between $\max_i\{1/\mu_i^0\}$ and $\min_i\{1/\mu_i^0\}$ can be, and most likely is, a result of asymmetry in the power limits, number of antennas, and/or channels at the two source nodes. It will also be shown in detail later in the simulations that such asymmetry can increase the occurrence of the two inefficient subcases, i.e., Subcases I-2 and II-4. In contrary, if the two source nodes have the same available power, same number of antennas, and same channel matrices, then $1/\mu_1^0 = 1/\mu_2^0 > 1/\mu_{\rm ma}^0$. As a result, only Subcases I-1 and II-1 are possible, in which the optimal power allocation is efficient. Combining this fact with the observations in the above three paragraphs, it can be seen that the asymmetry in the power limits, number of antennas, and/or channels at the two source nodes can lead to a degradation in the power allocation efficiency for the considered scenario of network optimization. As efficiency reveals the cost of coordination between the relay and source nodes required to achieve optimality over the two phases in the network optimization scenario, it can be seen that such cost is low with source node symmetry and high otherwise.

4.3 Numerical and simulation results

In this section, we provide simulation examples for some results presented earlier and demonstrate the proposed algorithm for network optimization in Table 4.1. The general setup is as follows. The elements of the channels $\mathbf{H}_{ri}$ and $\mathbf{H}_{i\rm r}, \forall i$ are generated from complex Gaussian distribution with zero mean and unit variance. The noise powers $\sigma_i^2, \forall i$ and $\sigma_{\rm r}^2$ are set to 1. The rates $R^{\rm ma}(\mathbf{D})$, $\bar{R}_{i\rm r}(\mathbf{B}_i)$, and $\hat{R}_{{\rm r}i}(\mathbf{D}_i)$ are briefly denoted as $R^{\rm ma}$, $\bar{R}_{i\rm r}$, and $\hat{R}_{{\rm r}i}$, respectively, in all figures.



(a) Convergence of $R^{\mathrm{tw}}(\mathbf{B},\mathbf{D})$, $R^{\mathrm{ma}}(\mathbf{D})$ and $R(\mathbf{B},\mathbf{D})$.



(b) $\bar{R}_{ir}(\mathbf{D}_i)$, $\hat{R}_{ri}(\lambda_i), \forall i$ and the power consumption of the source nodes over the number of iterations.

Figure 4.1: Illustration of the algorithm in Table 4.1 for Subcase I-2.

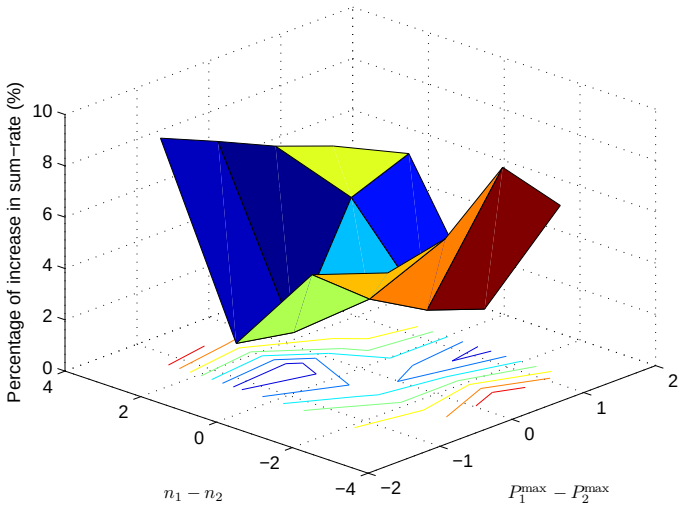
4.3.1 The process of finding the optimal solution for network optimization, Subcase I-2, using the proposed algorithm in Table 4.1

The specific setup for this example is as follows. The number of antennas n_1, n_2, and n_{r} are set to be $6, 4,$ and 8, respectively. Power limits for the source nodes are $P_1^{\max} = 2, P_2^{\max} = 2.5$. The relay's power limit is set to $P_{\mathrm{r}}^{\max} = 3$. Since the

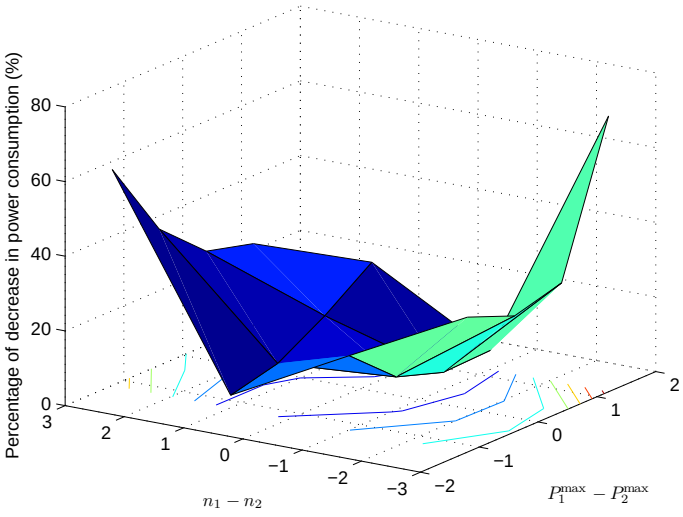
optimality of the solution derived using the algorithm has been proved analytically by the insights from Theorem 4.2, we focus on demonstrating the iterative process and the convergence of the algorithm. Fig. 4.1a shows instantaneous $R^{\mathrm{tw}}(\mathbf{B}, \mathbf{D})$, $R^{\mathrm{ma}}(\mathbf{D})$, and $R(\mathbf{B}, \mathbf{D})$ versus the number of iterations. From the figure, it can be seen that the above three rates converge very fast. Fig. 4.1b shows the instantaneous $\bar{R}_{i\mathrm{r}}(\mathbf{D}_i)$, $\hat{R}_{\mathrm{r}i}(\lambda_i), \forall i$ and the power consumption of the source nodes 1 and 2, denoted as P_1 and P_2, respectively. Two observations can be drawn from Fig. 4.1b. First, $\hat{R}_{\mathrm{r}2}(\lambda_2) < \bar{R}_{1\mathrm{r}}(\mathbf{D}_1)$ and $\hat{R}_{\mathrm{r}1}(\lambda_1) = \bar{R}_{2\mathrm{r}}(\mathbf{D}_2)$ at optimality since the sum-rate is bounded by $R^{\mathrm{ma}}(\mathbf{D}) < \bar{R}_{1\mathrm{r}}(\mathbf{D}_1) + \bar{R}_{2\mathrm{r}}(\mathbf{D}_2)$. Second, both source nodes use all available power at optimality. The latter observation verifies the conclusion that for Case I the optimal power allocation in Subcase I-2 is inefficient for using relatively more power and achieving relatively less sum-rate comparing to that in Subcase I-1.

4.3.2 Comparison with relay optimization in Chapter 3

The specific setup for this example is as follows. The number of antennas at the relay, i.e., n_{r}, is set to be 5. The power limit of the relay, i.e., $P_{\mathrm{r}}^{\max}$, is set to be 3. The total number of antennas at both source nodes is fixed so that $n_1 + n_2 = 5$. The total available power at both source nodes is also fixed so that $P_1^{\max} + P_2^{\max} = 2$. Given the above total number of antennas and total available power at the source nodes, both the relay optimization and the network optimization problems are solved for different combinations of n_1, n_2, $P_1^{\max}$, and $P_2^{\max}$ each with 100 channel realizations. The percentage of the increase in the average sum-rate and the percentage of the decrease in the average power consumption at optimality of the network optimization problem compared to those at optimality of the relay optimization problem are plotted in Figs. 4.2a and 4.2b, respectively. These percentages are shown versus the difference between the number of antennas and the difference between the power limits at the source nodes. From these two figures, it can be seen that although the optimal solution of the network optimization problem on average consumes much less power than that of the relay optimization problem, it still achieves larger sum-rate. Moreover, it can also be seen that the improvements, in either sum-rate or power consumption of the optimal solution of the network optimiza-



(a) Percentage of the increase in sum-rate at optimality.



(b) Percentage of the decrease in power consumption at optimality.

Figure 4.2: Improvements as compared to relay optimization.

tion problem as compared to that of the relay optimization problem, become more obvious when there is more asymmetry in the system. This is because the source nodes and the relay can jointly optimize their power allocations and therefore cope to some extent with the negative effect of the asymmetry in the system in the network optimization scenario. In contrast, the relay optimization scenario does not has such capability to combat the negative effect of asymmetry.

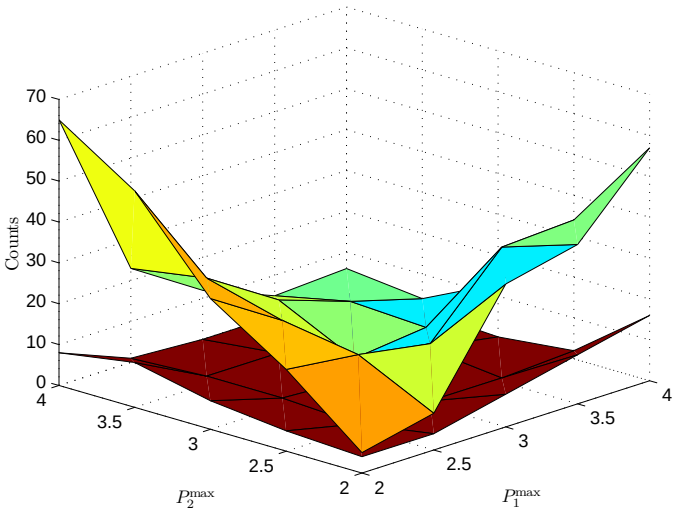
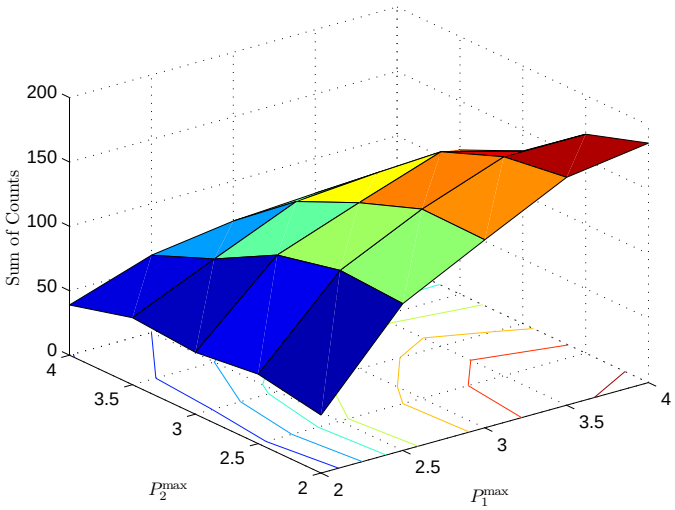


Figure 4.3: Number of channel realizations that Subcases I-2 and II-4 appear depending on the asymmetry in $P_1^{\max}$ and $P_2^{\max}$.

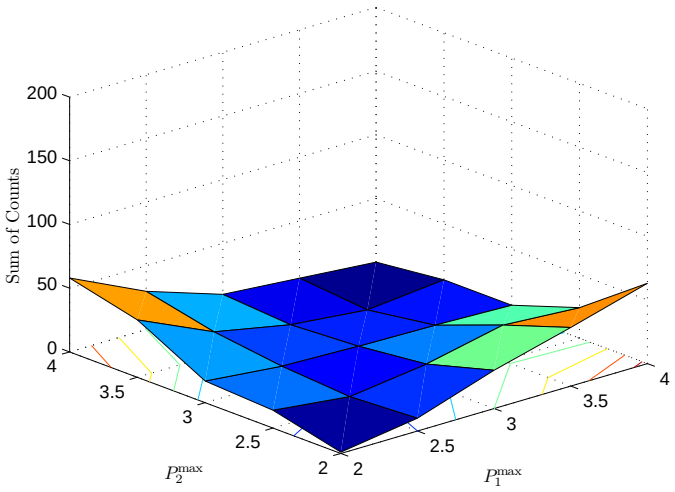
4.3.3 The effect of asymmetry in the scenario of network optimization

First, we solve the network optimization problem for different $P_1^{\max}$ and $P_2^{\max}$ given that $P_{\mathrm{r}}^{\max}$ is fixed. The number of antennas of the relay is set to 8 and the number of antennas of each source node is set to 4. For each combination of $P_1^{\max}$ and $P_2^{\max}$, we use 200 channel realizations and solve the resulting 200 network optimization problems. The number of channel realizations that Subcases I-2 and II-4 appear are plotted in Fig. 4.3. In this figure, the points in the upper surface correspond to the counts of Subcase I-2 while the points in the lower surface correspond to the counts of Subcase II-4. From Fig. 4.3, it can be seen that in general the count of either Subcase I-2 or Subcase II-4 is the smallest when $P_1^{\max} = P_2^{\max}$. Moreover, for any given $P_1^{\max}$ or $P_2^{\max}$, the largest count of either Subcase I-2 or Subcase II-4 mostly appears when the difference between $P_1^{\max}$ and $P_2^{\max}$ is the largest.[7] The above two observations are accurate for most of the times in Fig. 4.3, which shows that the asymmetry of $P_i^{\max}$ leads to the rise of the occurrence of Subcases I-2 and II-4.

[7]Note, however, that subcases are also determined by the number of antennas at the relay and the source nodes, the power limit $P_{\mathrm{r}}^{\max}$, the channel realizations, and other factors instead of only by $P_i^{\max}, \forall i$.



(a) Sum of the counts of Subcases I-2 and II-4 versus $P_1^{\max}$ and $P_2^{\max}$, $n_1 = 4, n_2 = 6$.



(b) Sum of the counts of Subcases I-2 and II-4 versus $P_1^{\max}$ and $P_2^{\max}$, $n_1 = n_2 = 6$.

Figure 4.4: Illustration of the effect of asymmetry in the number of antennas at the source nodes.

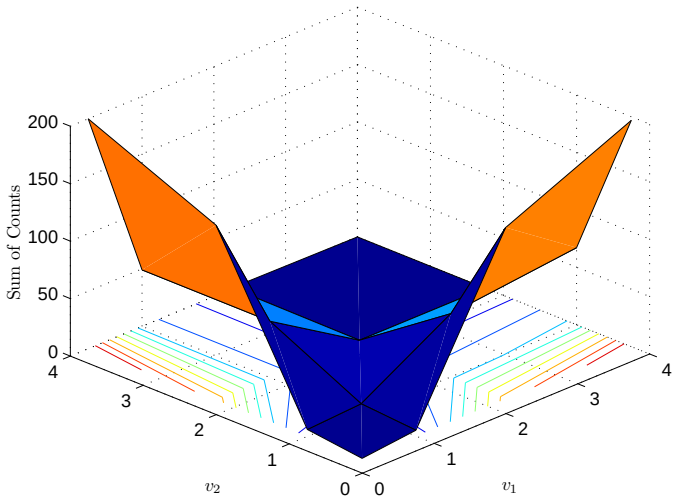
Next we demonstrate the effect of asymmetry in the number of antennas at the source nodes. The number of antennas at the relay is still 8 and $P_{\mathrm{r}}^{\max}$ is still 4. However, the number of antennas at the sources nodes 1 and 2 are first set to 4 and 6, respectively, and then set to both 6. The network optimization problem is solved for different $P_1^{\max}$ and $P_2^{\max}$ and the sum of the counts of Subcases I-2 and II-4 in 200 channel realizations is plotted in Fig. 4.4 for each combination of $P_1^{\max}$ and

$P_2^{\max}$. From Fig. 4.4a, it can be seen that the sum of the counts of Subcases I-2 and II-4 substantially increases when $n_1 = 4$ and $n_2 = 6$ as compared to the sum of the counts in Fig. 4.3 on most of the points. However, as shown in Fig. 4.4b, when $n_1 = n_2 = 6$, the sum of the counts of Subcases I-2 and II-4 drops to the same level as the sum of the counts in Fig. 4.3. Therefore, it can be seen that asymmetry in the number of antennas at the source nodes leads to larger chance of Subcases I-2 and II-4.

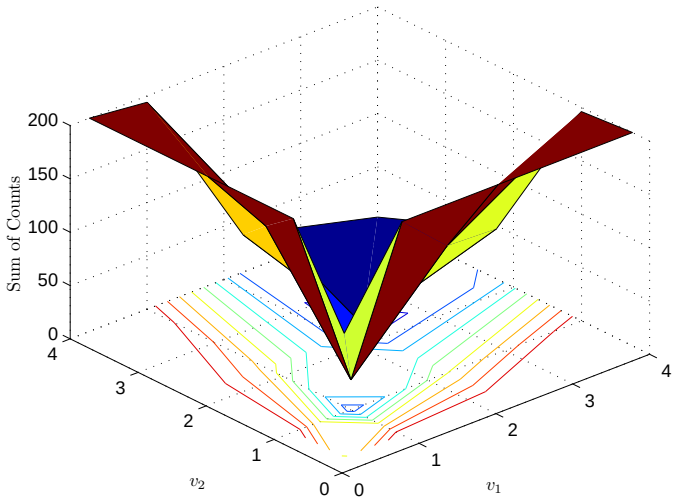
Lastly, we show the effect of asymmetry in channels. Instead of generating the real and imaginary parts of each element of $\mathbf{H}_{i\mathrm{r}}, \forall i$ from Gaussian distributions with zero mean and unit variance, we use here Gaussian distribution with zero mean and variance v_i to generate the real and imaginary parts of each element of $\mathbf{H}_{i\mathrm{r}}, \forall i$. For each combination of v_1 and v_2, we use 200 channel realizations and solve the resulting 200 network optimization problems. The number of antennas at the relay is set to 6 and the number of antennas at each source node is set to 4. The power limits are $P_{\mathrm{r}}^{\max} = 5$ and $P_i^{\max} = 3, \forall i$. The sum of the counts of Subcases I-2 and II-4 is plotted in Fig. 4.5 versus v_1 and v_2. In Fig. 4.5a, channel reciprocity is not assumed and the real and imaginary parts of each element of $\mathbf{H}_{\mathrm{r}i}, \forall i$ are generated from Gaussian distribution with zero mean and unit variance. In Fig. 4.5b, channel reciprocity is assumed i.e., $\mathbf{H}_{\mathrm{r}i} = \mathbf{H}_{i\mathrm{r}}^{\mathrm{T}}, \forall i$ where $(\cdot)^{\mathrm{T}}$ represents transpose. It can be seen from both Figs. 4.5a and 4.5b that the sum of the counts of Subcases I-2 and II-4 tends to increase when the difference between v_1 and v_2 becomes larger. Therefore, Fig. 4.5 clearly shows that the asymmetry in the channels also leads to larger chance of the inefficient Subcases I-2 and II-4.

4.4 Conclusions

In this chapter, we have solved the problem of sum-rate maximization using minimum total transmission power for MIMO DF TWR in the scenario of network optimization. For finding the optimal solution, we study the original problem in two cases each of which has several subcases. It has been shown that for all except two subcases, the originally nonconvex problem can be simplified into correspond-



(a) Sum of the counts of Subcases I-2 and II-4 versus v_1 and v_2 (without assuming channel reciprocity).



(b) Sum of the counts of Subcases I-2 and II-4 versus v_1 and v_2 (assuming channel reciprocity).

Figure 4.5: Illustration of the effect of asymmetry in channel statistics.

ing convex optimization problems. For the remaining two subcases, we have found the properties that the optimal solution must satisfy and have proposed the algorithm to find the optimal solution based on these properties. We have shown that the optimal power allocation in these two subcases are inefficient in the sense that it always consumes all the available power of the relay (and sometimes all the available power of the source nodes as well) yet cannot achieve the maximum sum-rate

of either the MA or BC phase. We have also shown that the asymmetry in the power limits, number of antennas, and channels leads to a higher probability of the above-mentioned two inefficient subcases. Together with Chapter 3, we have provided a complete and detailed study of the problem of sum-rate maximization using minimum power consumption for MIMO DF TWR.

$\sim$

Chapter 5

Jamming in Multi-User Wireless Communications

In this chapter, the worst-case noise jamming in MIMO wireless communications and the worst-case multi-target correlated jamming are investigated. For characterizing the worst-case jamming threat, it is assumed that the jammer has the channel information of all channels. Two scenarios are considered. In the first scenario, the jammer knows the covariance of the target signal and optimizes its jamming signal to perform worst-case noise jamming. In the second scenario, the jammer has the knowledge of multiple legitimate signals and performs multi-target correlated jamming.

5.1 Optimal non-correlated jamming

In this section, the problem of finding the non-correlated jamming[1] strategy to minimize the rate of a legitimate communication will be investigated. Since the jammer aims at causing maximum damage to the legitimate communication, the worst-case jamming is achieved when the jammer's strategy is optimized. Therefore, the term "optimal jamming" will be used as an alternative for the term "worst-case jamming" throughout the chapter.

[1]The term "non-correlated jamming" is used as an alternative for the term "noise jamming" in this chapter.

5.1.1 System model

A legitimate transmitter with n_{t} antennas sends a signal $\mathbf{s}$ to a receiver with n_{r} antennas. The elements of $\mathbf{s}$ are independent and identically distributed Gaussian with zero mean and covariance $\mathbf{Q}_{\mathrm{s}}$. A jammer with n_{z} antennas attempts to jam the legitimate communication by transmitting a jamming signal $\mathbf{z}$ to the receiver. Denote the legitimate channel from the legitimate transmitter to the receiver as $\mathbf{H}_{\mathrm{r}}$ (of size $n_{\mathrm{r}} \times n_{\mathrm{t}}$) and the jamming channel from the jammer to the receiver as $\mathbf{H}_{\mathrm{z}}$ (of size $n_{\mathrm{r}} \times n_{\mathrm{z}}$). In the presence of the jamming signal, the received signal at the legitimate receiver is expressed as

$$\mathbf{y} = \mathbf{H}_{\mathrm{r}}\mathbf{s} + \mathbf{H}_{\mathrm{z}}\mathbf{z} + \mathbf{n} \tag{5.1}$$

where $\mathbf{n}$ is the noise at the legitimate receiver with zero mean and covariance $\sigma^2\mathbf{I}$. Note that given the Gaussian channel and Gaussian target signal, the worst-case form of jamming signal is also Gaussian [83]. Denote the covariance of $\mathbf{z}$ as $\mathbf{Q}_{\mathrm{z}}$. Then the information rate of the legitimate communication in presence of the jamming is

$$R^{\mathrm{J}} = \log\left|\mathbf{I} + \mathbf{H}_{\mathrm{r}}\mathbf{Q}_{\mathrm{s}}\mathbf{H}_{\mathrm{r}}^{\mathrm{H}}(\mathbf{H}_{\mathrm{z}}\mathbf{Q}_{\mathrm{z}}\mathbf{H}_{\mathrm{z}}^{\mathrm{H}} + \sigma^2\mathbf{I})^{-1}\right|. \tag{5.2}$$

The jammer aims at decreasing the above rate as much as possible given its power limit P_{z}. Assuming that the jammer has the knowledge of $\mathbf{H}_{\mathrm{r}}$, $\mathbf{H}_{\mathrm{z}}$, and $\mathbf{Q}_{\mathrm{s}}$ but does not know the exact $\mathbf{s}$, the jammer can use the available knowledge to find the optimal $\mathbf{Q}_{\mathrm{z}}$ such that the rate (5.2) is minimized. This problem is studied in details in the following section.

5.1.2 Optimal jamming in closed-form under PSD condition

Given the system model, the optimal non-correlated jamming strategy can be found by solving the following problem[2]

$$\min_{\mathbf{Q}_{\mathrm{z}}} \quad R^{\mathrm{J}} \tag{5.3a}$$

$$\mathbf{s.t.} \quad \mathrm{Tr}\{\mathbf{Q}_{\mathrm{z}}\} \leq P_{\mathrm{z}}. \tag{5.3b}$$

[2]The PSD constraint $\mathbf{Q}_{\mathrm{z}} \succeq 0$ is assumed as default and omitted for brevity throughout this chapter.

With only one pair of transceiver, the above problem is a basic jamming problem on a MIMO channel.

Denote the SVD of $\mathbf{H}_{\mathrm{z}}$ as $\mathbf{H}_{\mathrm{z}} = \mathbf{U}_{\mathrm{z}}\mathbf{\Omega}_{\mathrm{z}}\mathbf{V}_{\mathrm{z}}^{\mathrm{H}}$. The matrices $\mathbf{U}_{\mathrm{z}}$, $\mathbf{\Omega}_{\mathrm{z}}$, and $\mathbf{V}_{\mathrm{z}}$ are of sizes $n_{\mathrm{r}} \times n_{\mathrm{r}}$, $n_{\mathrm{r}} \times n_{\mathrm{z}}$, and $n_{\mathrm{z}} \times n_{\mathrm{z}}$, respectively. Define $\mathbf{B} \triangleq \mathbf{U}_{\mathrm{z}}^{\mathrm{H}}\mathbf{H}_{\mathrm{r}}\mathbf{Q}_{\mathrm{s}}\mathbf{H}_{\mathrm{r}}^{\mathrm{H}}\mathbf{U}_{\mathrm{z}}$. Note that $\mathbf{B}$ has the same rank as $\mathbf{H}_{\mathrm{r}}\mathbf{Q}_{\mathrm{s}}\mathbf{H}_{\mathrm{r}}^{\mathrm{H}}$. Using the definition of $\mathbf{B}$ and the SVD of $\mathbf{H}_{\mathrm{z}}$, the objective function in (5.2) can be rewritten as

$$R^{\mathrm{J}} = \log|\mathbf{I} + \mathbf{B}(\mathbf{\Omega}_{\mathrm{z}}\hat{\mathbf{Q}}_{\mathrm{z}}\mathbf{\Omega}_{\mathrm{z}}^{\mathrm{H}} + \sigma^2\mathbf{I})^{-1}| \tag{5.4}$$

where

$$\hat{\mathbf{Q}}_{\mathrm{z}} \triangleq \mathbf{V}_{\mathrm{z}}^{\mathrm{H}}\mathbf{Q}_{\mathrm{z}}\mathbf{V}_{\mathrm{z}}. \tag{5.5}$$

In order to solve the problem (5.3), we start from introducing the following two lemmas.

Lemma 5.1: Given a constant Hermitian matrix $\mathbf{A}$ with $\mathbf{A} \succ 0$, the optimization problem over positive definite matrix $\mathbf{X}$

$$\begin{aligned} \min_{\mathbf{X}} \quad & \log\left|\mathbf{I} + \mathbf{A}\mathbf{X}^{-1}\right| && \text{(5.6a)} \\ \mathbf{s.t.} \quad & \mathrm{Tr}\{\mathbf{X}\} \leq 1 && \text{(5.6b)} \\ & \mathbf{X} \succeq 0 && \text{(5.6c)} \end{aligned}$$

has the following closed-form solution

$$\mathbf{X} = \mathbf{U}_{\mathrm{A}}\sqrt{\frac{\mathbf{\Lambda}_{\mathrm{A}}}{\lambda} + \frac{\mathbf{\Lambda}_{\mathrm{A}}^2}{4}}\mathbf{U}_{\mathrm{A}}^{\mathrm{H}} - \frac{\mathbf{A}}{2} \tag{5.7}$$

where $\mathbf{U}_{\mathrm{A}}$ and $\mathbf{\Lambda}_{\mathrm{A}}$ are the eigenvector and eigenvalue matrices, respectively, obtained from the EVD $\mathbf{A} = \mathbf{U}_{\mathrm{A}}\mathbf{\Lambda}_{\mathrm{A}}\mathbf{U}_{\mathrm{A}}^{\mathrm{H}}$, and λ is chosen so that the power constraint (5.6b) is satisfied with equality.

Proof: See Section C.1 in Appendix C.

Denote the rank of $\mathbf{H}_{\mathrm{z}}$ as r_{z} and assume without loss of generality that the first r_{z} elements on the main diagonal of $\mathbf{\Omega}_{\mathrm{z}}$ are non-zero. Whether or not $\mathbf{B}$ is positive definite, i.e., has the rank of n_{z}, has an impact on the optimal form of $\hat{\mathbf{Q}}_{\mathrm{z}}$ in (5.4). Therefore, the following lemma regarding $\mathbf{B}$ is in order.

Lemma 5.2: If we denote $\mathbf{B}$ using blocks such that

$$\mathbf{B} = \begin{array}{c} \\ r_{\mathrm{z}} \\ n_{\mathrm{z}}-r_{\mathrm{z}} \end{array} \begin{array}{c} \begin{array}{cc} r_{\mathrm{z}} & n_{\mathrm{z}}-r_{\mathrm{z}} \end{array} \\ \begin{bmatrix} \mathbf{B}_{11} & \mathbf{B}_{12} \\ \mathbf{B}_{21} & \mathbf{B}_{22} \end{bmatrix} \end{array} \tag{5.8}$$

and define

$$\tilde{\mathbf{B}} \triangleq \mathbf{B}_{11} - \mathbf{B}_{12}(\sigma^2\mathbf{I} + \mathbf{B}_{22})^{-1}\mathbf{B}_{21}, \tag{5.9}$$

then $\tilde{\mathbf{B}}$ is positive definite if $\mathbf{B}$ is positive definite.

Proof: See Section C.2 in Appendix C.

Let us define a new eigen channel $\tilde{\mathbf{\Omega}}_{\mathrm{z}}$ as

$$\tilde{\mathbf{\Omega}}_{\mathrm{z}} \triangleq \begin{array}{c} \\ r_{\mathrm{z}} \\ n_{\mathrm{r}}-r_{\mathrm{z}} \end{array} \begin{array}{c} \begin{array}{cc} r_{\mathrm{z}} & n_{\mathrm{r}}-r_{\mathrm{z}} \end{array} \\ \begin{bmatrix} \mathbf{\Omega}_{\mathrm{z}}^{+} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} \end{bmatrix} \end{array} \tag{5.10}$$

where $\mathbf{\Omega}_{\mathrm{z}}^{+}$ is an $r_{\mathrm{z}} \times r_{\mathrm{z}}$ diagonal matrix made of the positive diagonal elements of $\mathbf{\Omega}_{\mathrm{z}}$. Also define a new jamming covariance matrix $\tilde{\mathbf{Q}}_{\mathrm{z}}$ as

$$\tilde{\mathbf{Q}}_{\mathrm{z}} \triangleq \begin{array}{c} \\ r_{\mathrm{z}} \\ n_{\mathrm{r}}-r_{\mathrm{z}} \end{array} \begin{array}{c} \begin{array}{cc} r_{\mathrm{z}} & n_{\mathrm{r}}-r_{\mathrm{z}} \end{array} \\ \begin{bmatrix} \mathbf{Q}'_{\mathrm{z}} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \end{array} \tag{5.11}$$

where $\mathbf{Q}'_{\mathrm{z}}$ is the part of the matrix to be determined. With the above $\tilde{\mathbf{\Omega}}_{\mathrm{z}}$ and $\tilde{\mathbf{Q}}_{\mathrm{z}}$, the rate in (5.4) can be equivalently rewritten as

$$R^{\mathrm{J}} = \log\left|\mathbf{I} + \mathbf{B}(\tilde{\mathbf{\Omega}}_{\mathrm{z}}\tilde{\mathbf{Q}}_{\mathrm{z}}\tilde{\mathbf{\Omega}}_{\mathrm{z}}^{\mathrm{H}} + \sigma^2\mathbf{I})^{-1}\right|. \tag{5.12}$$

Therefore, we consider $\tilde{\mathbf{\Omega}}_{\mathrm{z}}$ and $\tilde{\mathbf{Q}}_{\mathrm{z}}$ as the equivalent channel matrix and the equivalent jamming covariance matrix to $\mathbf{\Omega}_{\mathrm{z}}$ and $\hat{\mathbf{Q}}_{\mathrm{z}}$, respectively.

The equivalent channel $\tilde{\mathbf{\Omega}}_{\mathrm{z}}$ has the size $n_{\mathrm{r}} \times n_{\mathrm{r}}$, which is larger than the size of $\mathbf{\Omega}_{\mathrm{z}}$ if $n_{\mathrm{r}} > n_{\mathrm{z}}$ and smaller than the size of $\mathbf{\Omega}_{\mathrm{z}}$ if $n_{\mathrm{r}} < n_{\mathrm{z}}$. Correspondingly, the allocation of jamming power in (5.11) represented by $\mathbf{Q}'_{\mathrm{z}}$ is limited to at most r_{z} dimensions corresponding to the r_{z} non-zero eigenvalues of $\mathbf{\Omega}_{\mathrm{z}}^{+}$. It can be seen that allocating jamming power anywhere else has no effect on the received signal and only leads to jamming power waste. Therefore, the optimal structure of $\hat{\mathbf{Q}}_{\mathrm{z}}$ has to be in the form

$$\hat{\mathbf{Q}}_{\mathrm{z}} = \begin{array}{c} \\ n_{\mathrm{r}} \\ n_{\mathrm{z}}-n_{\mathrm{r}} \end{array} \begin{array}{c} \begin{array}{cc} n_{\mathrm{r}} & n_{\mathrm{z}}-n_{\mathrm{r}} \end{array} \\ \begin{bmatrix} \tilde{\mathbf{Q}}_{\mathrm{z}} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \end{array} = \begin{array}{c} \\ r_{\mathrm{z}} \\ n_{\mathrm{z}}-r_{\mathrm{z}} \end{array} \begin{array}{c} \begin{array}{cc} r_{\mathrm{z}} & n_{\mathrm{z}}-r_{\mathrm{z}} \end{array} \\ \begin{bmatrix} \mathbf{Q}'_{\mathrm{z}} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \end{array}. \tag{5.13}$$

Using (5.5) and (5.13) it can be seen that the optimal form of $\mathbf{Q}_{\mathrm{z}}$ is

$$\mathbf{Q}_{\mathrm{z}} = \mathbf{V}_{\mathrm{z}} \begin{bmatrix} \mathbf{Q}'_{\mathrm{z}} & \mathbf{0} \\ \mathbf{0} & \mathbf{0}_{n_{\mathrm{z}}-r_{\mathrm{z}}} \end{bmatrix} \mathbf{V}_{\mathrm{z}}^{\mathrm{H}}. \tag{5.14}$$

Given the above definitions and lemmas, we next solve the problem (5.3) by finding the optimal $\mathbf{Q}'_{\mathrm{z}}$ in (5.14). First, we consider a specific case that $\mathbf{H}_{\mathrm{r}}\mathbf{Q}_{\mathrm{s}}\mathbf{H}_{\mathrm{r}}^{\mathrm{H}}$ is positive definite. Then, we will extend the solution to the general case that $\mathbf{H}_{\mathrm{r}}\mathbf{Q}_{\mathrm{s}}\mathbf{H}_{\mathrm{r}}^{\mathrm{H}}$ is PSD but not necessarily positive definite.

Theorem 5.1: When $\mathbf{H}_{\mathrm{r}}\mathbf{Q}_{\mathrm{s}}\mathbf{H}_{\mathrm{r}}^{\mathrm{H}}$ is positive definite, the problem (5.3) has the following closed-form optimal solution

$$\mathbf{Q}'_{\mathrm{z}} = \mathbf{U}_{\tilde{\mathbf{A}}} \sqrt{\frac{1}{\lambda}\mathbf{\Lambda}_{\tilde{\mathbf{A}}} + \frac{1}{4}\mathbf{\Lambda}_{\tilde{\mathbf{A}}}^2}\, \mathbf{U}_{\tilde{\mathbf{A}}}^{\mathrm{H}} - \mathbf{\Omega}_{\mathrm{z}}^{+-1}\big(\frac{1}{2}\tilde{\mathbf{B}} + \sigma^2\mathbf{I}\big)\mathbf{\Omega}_{\mathrm{z}}^{+-\mathrm{H}} \tag{5.15}$$

under the condition that the above $\mathbf{Q}'_{\mathrm{z}}$ is PSD, where $\tilde{\mathbf{B}}$ is given by (5.9), $\mathbf{U}_{\tilde{\mathbf{A}}}$ and $\mathbf{\Lambda}_{\tilde{\mathbf{A}}}$ are obtained from the EVD $\tilde{\mathbf{A}} = \mathbf{U}_{\tilde{\mathbf{A}}}\mathbf{\Lambda}_{\tilde{\mathbf{A}}}\mathbf{U}_{\tilde{\mathbf{A}}}^{\mathrm{H}}$ with

$$\tilde{\mathbf{A}} \triangleq \mathbf{\Omega}_{\mathrm{z}}^{+-1}\tilde{\mathbf{B}}\mathbf{\Omega}_{\mathrm{z}}^{+-\mathrm{H}}, \tag{5.16}$$

and λ is chosen such that the jammer's power constraint (5.3b) is satisfied with equality.

Proof: Please see Section C.3 in Appendix C.

As mentioned in the introduction in Chapter 1, a special case of the problem (5.3) that assumes the jamming channel $\mathbf{H}_{\mathrm{z}}$ is the identity matrix $\mathbf{I}$ is investigated in [47]. Consequently, $\mathbf{U}_{\mathrm{z}}$, $\mathbf{\Omega}_{\mathrm{z}}$, and $\mathbf{V}_{\mathrm{z}}^{\mathrm{H}}$ are all equal to $\mathbf{I}$. Therefore, $\tilde{\mathbf{A}}$ and $\mathbf{\Omega}_{\mathrm{z}}^{+}$ simplify to $\tilde{\mathbf{B}}$ and $\mathbf{I}$, respectively. Moreover, the above simplification in [47] leads to the result that $r_{\mathrm{z}} = n_{\mathrm{z}}$, which further simplifies the case so that $\tilde{\mathbf{B}} = \mathbf{B}$ and $\mathbf{Q}_{\mathrm{z}} = \mathbf{Q}'_{\mathrm{z}}$. In such case, the solution in (5.15) simplifies to

$$\mathbf{Q}'_{\mathrm{z}} = \mathbf{U}_{\mathbf{B}}\big(\sqrt{\frac{1}{\lambda}\mathbf{\Lambda}_{\mathbf{B}} + \frac{1}{4}\mathbf{\Lambda}_{\mathbf{B}}^2} - \frac{1}{2}\mathbf{\Lambda}_{\mathbf{B}} - \sigma^2\mathbf{I}\big)\mathbf{U}_{\mathbf{B}}^{\mathrm{H}} \tag{5.17}$$

where $\mathbf{U}_{\mathbf{B}}$ and $\mathbf{\Lambda}_{\mathbf{B}}$ are obtained from the EVD $\mathbf{B} = \mathbf{U}_{\mathbf{B}}\mathbf{\Lambda}_{\mathbf{B}}\mathbf{U}_{\mathbf{B}}^{\mathrm{H}}$. An equivalent scalar form of the above solution is given in [47] for the above simplified case of the problem. By forcing the negative elements (if any) of $\sqrt{\mathbf{\Lambda}_{\mathbf{B}}/\lambda + \mathbf{\Lambda}_{\mathbf{B}}^2/4} - \frac{1}{2}\mathbf{\Lambda}_{\mathbf{B}} - \sigma^2\mathbf{I}$ to be zero and adjusting λ to satisfy the power constraint, the solution given in (5.17) can always be made PSD.

The solution of $\mathbf{Q}_{\rm z}'$ given by (5.15) is not necessarily PSD for the case considered in Theorem 5.1. It can be indefinite when the jammer's power limit $P_{\rm z}$ is sufficiently small. It can be seen that $1/\lambda$ decreases when the jammer's power limit becomes smaller. As a result, $\mathbf{Q}_{\rm z}'$ has a larger chance to be indefinite and thereby invalid as a solution of a covariance matrix. For a given power limit $P_{\rm z}$, whether or not $\mathbf{Q}_{\rm z}'$ in (5.15) is PSD depends on the channel $\mathbf{H}_{\rm z}$, or essentially, the elements of $\mathbf{\Omega}_{\rm z}^{+}$. It can be shown that, for a small $P_{\rm z}$ and a given $\mathbf{\Omega}_{\rm z}^{+}$ such that $\mathbf{Q}_{\rm z}'$ given by (5.15) is indefinite, there always exists $\tilde{\mathbf{\Omega}}_{\rm z}^{+}$ with the same trace as $\mathbf{\Omega}_{\rm z}^{+}$ (i.e., ${\rm Tr}\{\tilde{\mathbf{\Omega}}_{\rm z}^{+}\} = {\rm Tr}\{\mathbf{\Omega}_{\rm z}^{+}\}$) but different elements, such that $\mathbf{Q}_{\rm z}'$ is PSD if $\mathbf{\Omega}_{\rm z}^{+}$ in (5.15) is substituted by $\tilde{\mathbf{\Omega}}_{\rm z}^{+}$. Therefore, the power limit of the jammer as well as the gains of the eigen-channels determine whether or not $\mathbf{Q}_{\rm z}'$ is PSD. The above fact, which reveals the effect of the jamming power limit and the jamming channel on the jammer's strategy, has not been observed before as the jamming channel has been neglected.

Unlike the case of [47] in which the solution can always be made PSD by forcing the negative elements to be zero and adjusting the λ to satisfy the power constraint, such method does not work for the case considered here. The problem of finding the solution when $\mathbf{Q}_{\rm z}'$ in (5.15) is indefinite will be studied in Section 5.1.3.

Now consider the general case that $\mathbf{H}_{\rm r}\mathbf{Q}_{\rm s}\mathbf{H}_{\rm r}^{\rm H}$ is PSD but not necessarily positive definite. Since $\mathbf{H}_{\rm r}\mathbf{Q}_{\rm s}\mathbf{H}_{\rm r}^{\rm H}$, or equivalently $\mathbf{B}$, is PSD but not necessarily positive definite in this case, $\tilde{\mathbf{B}}$ in (5.9) and consequently $\tilde{\mathbf{A}}$ in (5.16) can be rank deficient. In this situation, assume that the rank of $\tilde{\mathbf{A}}$ is $r_{\tilde{\mathbf{A}}}$ and denote the diagonal matrix made of the $r_{\tilde{\mathbf{A}}}$ positive eigenvalues of $\tilde{\mathbf{A}}$ as $\mathbf{\Lambda}_{\tilde{\mathbf{A}}}^{+}$. Denote the EVD of $\tilde{\mathbf{A}}$ as

$$\tilde{\mathbf{A}} = \mathbf{U}_{\tilde{\mathbf{A}}}\mathbf{\Lambda}_{\tilde{\mathbf{A}}}\mathbf{U}_{\tilde{\mathbf{A}}}^{\rm H} = \begin{bmatrix} \overset{r_{\tilde{\mathbf{A}}}}{\mathbf{U}_{\tilde{\mathbf{A}}1}} & \overset{r_{\rm z}-r_{\tilde{\mathbf{A}}}}{\mathbf{U}_{\tilde{\mathbf{A}}2}} \end{bmatrix} \begin{bmatrix} \mathbf{\Lambda}_{\tilde{\mathbf{A}}}^{+} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{U}_{\tilde{\mathbf{A}}1}^{\rm H} \\ \mathbf{U}_{\tilde{\mathbf{A}}2}^{\rm H} \end{bmatrix}. \tag{5.18}$$

The following theorem regarding the solution in this general case is in order.

Theorem 5.2: When $\mathbf{H}_{\rm r}\mathbf{Q}_{\rm s}\mathbf{H}_{\rm r}^{\rm H}$ is PSD but not necessarily positive definite, the problem (5.3) has the following closed-form optimal solution

$$\mathbf{Q}_{\rm z}' = \mathbf{U}_{\tilde{\mathbf{A}}1}\sqrt{\frac{1}{\lambda}\mathbf{\Lambda}_{\tilde{\mathbf{A}}}^{+}+\frac{1}{4}\mathbf{\Lambda}_{\tilde{\mathbf{A}}}^{+\,2}}\mathbf{U}_{\tilde{\mathbf{A}}1}^{\rm H} - \frac{1}{2}\mathbf{U}_{\tilde{\mathbf{A}}1}\mathbf{\Lambda}_{\tilde{\mathbf{A}}}^{+}\mathbf{U}_{\tilde{\mathbf{A}}1}^{\rm H} - \sigma^2\mathbf{\Omega}_{\rm z}^{+-1}\mathbf{\Omega}_{\rm z}^{+-{\rm H}} \tag{5.19}$$

under the condition that the above $\mathbf{Q}_{\mathrm{z}}'$ is PSD, where λ is chosen such that the jammer's power constraint (5.3b) is satisfied with equality.

Proof: See Section C.4 in Appendix C.

It can be seen that if $\tilde{\mathbf{A}}$ has full rank, then (5.19) is equivalent to (5.15). Similarly, $\mathbf{Q}_{\mathrm{z}}'$ given by (5.19) can be indefinite depending on the jammer's power limit P_{z} and the jamming channel $\mathbf{\Omega}_{\mathrm{z}}^{+}$. To tackle this problem, we next find solutions of the problem (5.3) when $\mathbf{Q}_{\mathrm{z}}'$ given in (5.15) or (5.19) is indefinite.

5.1.3 Optimal numeric solution and closed-form approximation

As mentioned earlier, the closed-form expressions of $\mathbf{Q}_{\mathrm{z}}'$ given by (5.15) and (5.19) when $\mathbf{H}_{\mathrm{r}}\mathbf{Q}_{\mathrm{s}}\mathbf{H}_{\mathrm{r}}^{\mathrm{H}}$ is positive definite and PSD, respectively, may not be valid when the power constraint P_{z} is small. In such case, the optimal solution may not be found in closed-form. To solve this problem, we propose two different approaches in this section. The first one is to find the optimal solution numerically. The second one is to find a sub-optimal solution in closed-form. The two approaches provide a choice between accuracy and complexity. We start from describing an algorithm for finding the optimal solution of (5.3) numerically.

Substituting (5.10) and (5.11) into (5.12) and using the definitions (5.9) and (5.16), it can be shown[3] that the original problem of minimizing (5.4) is equivalent to the minimization of

$$\bar{R}^{\mathrm{J}} = \log\Big|\mathbf{I} + \tilde{\mathbf{A}}(\mathbf{Q}_{\mathrm{z}}' + \sigma^2\mathbf{\Omega}_{\mathrm{z}}^{+^{-1}}\mathbf{\Omega}_{\mathrm{z}}^{+^{-\mathrm{H}}})^{-1}\Big|. \tag{5.20}$$

Although the minimization of (5.20) subject to a power constraint is a convex problem, it is not a disciplined convex problem [84]. Therefore, the optimal solution cannot be obtained using classic convex optimization methods. In order to find the optimal solution, we first rewrite the problem into the following equivalent form

$$\begin{aligned}
\underset{\alpha,\mathbf{Q}_{\mathrm{z}}'}{\mathbf{min}} \quad & \alpha - \log|\mathbf{Q}_{\mathrm{z}}' + \mathbf{D}_0| && \text{(5.21a)}\\
\mathbf{s.t.} \quad & \alpha \geq \log\Big|\mathbf{Q}_{\mathrm{z}}' + \mathbf{D}_0 + \tilde{\mathbf{A}}\Big| && \text{(5.21b)}\\
& \mathrm{Tr}\{\mathbf{Q}_{\mathrm{z}}'\} \leq P_{\mathrm{z}} && \text{(5.21c)}
\end{aligned}$$

[3]The details can be found in the proof of Theorem 5.1, from (C.14) to (C.18), Section C.3 in Appendix C.

Table 5.1: Steps for finding the optimal solution of the problem (5.21).

1. Select a starting $\mathbf{Q}_{\mathrm{z}}'^{\dagger}$ subject to $\mathrm{Tr}\{\mathbf{Q}_{\mathrm{z}}'^{\dagger}\} \leq P_{\mathrm{z}}$.
2. Solve the problem (5.22) given $\mathbf{Q}_{\mathrm{z}}'^{\dagger}$. Denote the corresponding optimal solution of $\mathbf{Q}_{\mathrm{z}}'$ as $\mathbf{Q}_{\mathrm{z}}'^{*}$.
3. Set $\mathbf{Q}_{\mathrm{z}}'^{\dagger} = \mathbf{Q}_{\mathrm{z}}'^{*}$.
4. Repeat the Steps 2 and 3 until the solution converges.

in which $\mathbf{D}_0 \triangleq \sigma^2 \mathbf{\Omega}_{\mathrm{z}}^{+-1} \mathbf{\Omega}_{\mathrm{z}}^{+-\mathrm{H}}$. In the above problem, the objective function is convex while the first constraint is not. In order to solve the problem (5.21), we first consider the following problem in a similar form

$$\begin{aligned}
\underset{\alpha, \mathbf{Q}_{\mathrm{z}}'}{\mathbf{min}} \quad & \alpha - \log|\mathbf{Q}_{\mathrm{z}}' + \mathbf{D}_0| && \text{(5.22a)} \\
\mathbf{s.t.} \quad & \alpha \geq \log\left|\mathbf{Q}_{\mathrm{z}}'^{\dagger} + \mathbf{D}_0 + \tilde{\mathbf{A}}\right| + \mathrm{Tr}\{\left(\mathbf{Q}_{\mathrm{z}}'^{\dagger} + \mathbf{D}_0 \right. \\
& \qquad \left. + \tilde{\mathbf{A}}\right)^{-1} \mathbf{Q}_{\mathrm{z}}'\} - \mathrm{Tr}\{\left(\mathbf{Q}_{\mathrm{z}}'^{\dagger} + \mathbf{D}_0 + \tilde{\mathbf{A}}\right)^{-1} \mathbf{Q}_{\mathrm{z}}'^{\dagger}\} && \text{(5.22b)} \\
& \mathrm{Tr}\{\mathbf{Q}_{\mathrm{z}}'\} \leq P_{\mathrm{z}}. && \text{(5.22c)}
\end{aligned}$$

Here $\mathbf{Q}_{\mathrm{z}}'^{\dagger}$ stands for a given $\mathbf{Q}_{\mathrm{z}}'$ subject to (5.21c). The optimal solution of the problem (5.21) can be found from solving the problem (5.22) iteratively. Specifically, the corresponding algorithm is summarized in Table 5.1.

Lemma 5.3: The $\mathbf{Q}_{\mathrm{z}}'^{*}$ in the procedure described in Table 5.1 converges to the optimal solution of the problem (5.21).

Proof: See Section C.5 in Appendix C.

After obtaining the optimal $\mathbf{Q}_{\mathrm{z}}'^{*}$ using the algorithm in Table 5.1, the optimal $\mathbf{Q}_{\mathrm{z}}$ can be obtained using (5.14).

Using the algorithm for finding the optimal $\mathbf{Q}_{\mathrm{z}}'$ can be computationally complex as compared to obtaining a closed-form solution. Therefore, we next give an approximation of the optimal solution in closed-form when the $\mathbf{Q}_{\mathrm{z}}'$ given by (5.15) (when $\mathbf{H}_{\mathrm{r}}\mathbf{Q}_{\mathrm{s}}\mathbf{H}_{\mathrm{r}}^{\mathrm{H}}$ is positive definite) or (5.19) (when $\mathbf{H}_{\mathrm{r}}\mathbf{Q}_{\mathrm{s}}\mathbf{H}_{\mathrm{r}}^{\mathrm{H}}$ is PSD) is indefinite.

When $\mathbf{H}_{\mathrm{r}}\mathbf{Q}_{\mathrm{s}}\mathbf{H}_{\mathrm{r}}^{\mathrm{H}}$ is positive definite, a suboptimal closed-form solution to the considered problem when the $\mathbf{Q}_{\mathrm{z}}'$ in (5.15) is indefinite can be given as

$$\mathbf{Q}_{\mathrm{z}}' = \mathbf{U}_{\tilde{\mathbf{A}}} \sqrt{\frac{1}{\tilde{\lambda}} \mathbf{\Lambda}_{\tilde{\mathbf{A}}} + \frac{1}{4} \mathbf{\Lambda}_{\tilde{\mathbf{A}}}^2} \mathbf{U}_{\tilde{\mathbf{A}}}^{\mathrm{H}} - \frac{1}{2} \tilde{\mathbf{A}} + (\tilde{\epsilon} - 1) \mathbf{D}_0 \qquad \text{(5.23)}$$

in which $\tilde{\epsilon}$ and $\tilde{\lambda}$ are the optimal solution to the problem

$$\min_{\epsilon,\lambda} \quad \epsilon \tag{5.24a}$$

$$\text{s.t.} \quad \mathbf{U}_{\tilde{\mathbf{A}}}\sqrt{\frac{1}{\lambda}\mathbf{\Lambda}_{\tilde{\mathbf{A}}}+\frac{1}{4}\mathbf{\Lambda}_{\tilde{\mathbf{A}}}^2}\mathbf{U}_{\tilde{\mathbf{A}}}^{\mathrm{H}}-\frac{1}{2}\tilde{\mathbf{A}}+(\epsilon-1)\mathbf{D}_0\succeq 0 \tag{5.24b}$$

$$\mathrm{Tr}\left\{\sqrt{\frac{1}{\lambda}\mathbf{\Lambda}_{\tilde{\mathbf{A}}}+\frac{1}{4}\mathbf{\Lambda}_{\tilde{\mathbf{A}}}^2}-\frac{1}{2}\tilde{\mathbf{A}}+(\epsilon-1)\mathbf{D}_0\right\}=P_{\mathrm{z}} \tag{5.24c}$$

$$0\leq\epsilon\leq 1 \tag{5.24d}$$

$$\lambda>0. \tag{5.24e}$$

It is worth mentioning that the constraints (5.24b)-(5.24e) specify a non-empty feasible set. It can be found that the suboptimal solution (5.23) is equal to the expression in (5.15) plus $\tilde{\epsilon}\mathbf{D}_0$ (using the definitions (5.16) and $\mathbf{D}_0=\sigma^2\mathbf{\Omega}_{\mathrm{z}}^{+-1}\mathbf{\Omega}_{\mathrm{z}}^{+-\mathrm{H}}$). The logic behind the suboptimal solution (5.23) is that the remaining part of the expression (5.15) without $-\mathbf{D}_0$ is always PSD. Therefore, there exists a non-negative factor $\epsilon<1$ such that the summation is PSD if $-\mathbf{D}_0$ is scaled by $1-\epsilon$ and added back to the remaining part of (5.15). In order to remain as close as possible to the form of (5.15) in the above modification, the minimum ϵ that results in a PSD $\mathbf{Q}_{\mathrm{z}}'$ is used.

The above suboptimal solution given by (5.23) is proposed based on the following reasons. First and most important, it can be shown that $\mathbf{Q}_{\mathrm{z}}'$ given by the above suboptimal solution is the same as the $\mathbf{Q}_{\mathrm{z}}'$ given by (5.15) when the latter one is PSD (and consequently $\tilde{\epsilon}=0$). Therefore, the use of (5.23) is sufficient for calculating the jamming strategy in all cases because (5.23) gives the optimal solution when it exists in closed-form and gives the suboptimal solution otherwise. Second, when it is not optimal, the suboptimal solution given by (5.23) is in fact very close to the optimal one found numerically (as will be shown in simulations). Third, compared to the numerical solution, the suboptimal solution given by (5.23) can be obtained with negligible complexity since the parameters $\tilde{\epsilon}$ and $\tilde{\lambda}$ can be obtained by a simple bisectional search. Last, the above suboptimal solution is always PSD as can be seen from the constraint (5.24b).

The closed-form suboptimal solution for the general case when $\mathbf{H}_{\mathrm{r}}\mathbf{Q}_{\mathrm{s}}\mathbf{H}_{\mathrm{r}}^{\mathrm{H}}$ is PSD but not necessarily positive definite can be obtained similarly. In this case, the

Table 5.2: Summarizing the procedure for finding the solution to the problem (5.3).

1. Check whether or not $\mathbf{H}_{\mathrm{r}}\mathbf{Q}_{\mathrm{s}}\mathbf{H}_{\mathrm{r}}^{\mathrm{H}}$ is positive definite. If yes, obtain $\mathbf{Q}_{\mathrm{z}}'$ using (5.15). Otherwise, obtain $\mathbf{Q}_{\mathrm{z}}'$ using (5.19).
2. Check whether or not the above obtained $\mathbf{Q}_{\mathrm{z}}'$ is PSD. If yes, substitute the obtained $\mathbf{Q}_{\mathrm{z}}'$ into (5.14) to find the optimal $\mathbf{Q}_{\mathrm{z}}$. Otherwise, select from two options: a) optimal numerical solution; b) sub-optimal closed-form solution. For a), proceed to step 3. For b), proceed to step 4.
3. Use the algorithm in Table 5.1 to obtain the optimal numerical solution. Exit.
4. Obtain $\tilde{\epsilon}$ and $\tilde{\lambda}$ from solving the problem (5.24) (if $\mathbf{H}_{\mathrm{r}}\mathbf{Q}_{\mathrm{s}}\mathbf{H}_{\mathrm{r}}^{\mathrm{H}}$ is positive definite) or problem (5.26) (if $\mathbf{H}_{\mathrm{r}}\mathbf{Q}_{\mathrm{s}}\mathbf{H}_{\mathrm{r}}^{\mathrm{H}}$ is PSD but not positive definite). Then obtain the suboptimal closed-form solution accordingly using (5.14) with (5.23) (if $\mathbf{H}_{\mathrm{r}}\mathbf{Q}_{\mathrm{s}}\mathbf{H}_{\mathrm{r}}^{\mathrm{H}}$ is positive definite) or (5.25) (if $\mathbf{H}_{\mathrm{r}}\mathbf{Q}_{\mathrm{s}}\mathbf{H}_{\mathrm{r}}^{\mathrm{H}}$ is PSD but not positive definite). Exit.

suboptimal solution in closed-form is expressed as

$$\mathbf{Q}_{\mathrm{z}}' = \mathbf{U}_{\tilde{\mathbf{A}}1}\sqrt{\frac{1}{\tilde{\lambda}}\mathbf{\Lambda}_{\tilde{\mathbf{A}}}^{+}+\frac{1}{4}\mathbf{\Lambda}_{\tilde{\mathbf{A}}}^{+2}}\mathbf{U}_{\tilde{\mathbf{A}}1}^{\mathrm{H}} - \frac{1}{2}\mathbf{U}_{\tilde{\mathbf{A}}1}\mathbf{\Lambda}_{\tilde{\mathbf{A}}}^{+}\mathbf{U}_{\tilde{\mathbf{A}}1}^{\mathrm{H}} + (\tilde{\epsilon}-1)\mathbf{D}_0 \tag{5.25}$$

in which $\tilde{\epsilon}$ and $\tilde{\lambda}$ are the optimal solution to the problem

$$\min_{\epsilon,\lambda} \quad \epsilon \tag{5.26a}$$

$$\mathbf{s.t.} \quad \mathbf{U}_{\tilde{\mathbf{A}}1}\sqrt{\frac{1}{\lambda}\mathbf{\Lambda}_{\tilde{\mathbf{A}}}^{+}+\frac{1}{4}\mathbf{\Lambda}_{\tilde{\mathbf{A}}}^{+2}}\mathbf{U}_{\tilde{\mathbf{A}}1}^{\mathrm{H}} - \frac{1}{2}\mathbf{U}_{\tilde{\mathbf{A}}1}\mathbf{\Lambda}_{\tilde{\mathbf{A}}}^{+}\mathbf{U}_{\tilde{\mathbf{A}}1}^{\mathrm{H}} + (\epsilon-1)\mathbf{D}_0 \succeq 0 \tag{5.26b}$$

$$\mathrm{Tr}\left\{\sqrt{\frac{1}{\lambda}\mathbf{\Lambda}_{\tilde{\mathbf{A}}}^{+}+\frac{1}{4}\mathbf{\Lambda}_{\tilde{\mathbf{A}}}^{+2}} - \frac{1}{2}\mathbf{\Lambda}_{\tilde{\mathbf{A}}}^{+} + (\epsilon-1)\mathbf{D}_0\right\} = P_{\mathrm{z}} \tag{5.26c}$$

$$0 \leq \epsilon \leq 1 \tag{5.26d}$$

$$\lambda > 0. \tag{5.26e}$$

With the proposed closed-form optimal and sub-optimal solutions and the algorithm for finding the optimal numerical solution, the complete procedure of calculating the non-correlated jamming strategy $\mathbf{Q}_{\mathrm{z}}$ is summarized in Table 5.2.

5.2 Multi-target correlated jamming

If the jammer has the information of the legitimate signal (as considered in [16] and [85]), it can perform another form of jamming, i.e., correlated jamming. Instead of causing interference to the legitimate receiver, the objective of the jammer in

correlated jamming is to weaken or even completely cancel out the legitimate signal depending on its power limit. Therefore, correlated jamming can be very efficient for a jammer. Unlike the cases in [16] and [85], in which the authors investigate correlated jamming without considering the jamming channel as if the jamming is applied directly at the target receiver, here the jamming channel is taken into account in the investigation of jamming strategy. Moreover, the jammer needs to perform correlated jamming with more than one target. This significantly increases the complexity of the problem.

5.2.1 System model and problem formulation

To be consistent with Section 5.1 that considers MIMO, the following describes a general system model for multi-target correlated jamming in which each node has multiple antennas. There are m legitimate transceiver pairs and one jammer. The transmitter and receiver in the ith ($i = 1, \ldots, m$) transceiver pair have $n_{\mathrm{s}i}$ and $n_{\mathrm{r}i}$ antennas, respectively. The channel between the ith transceiver pair is denoted as $\mathbf{H}_i$ and the transmitted signal over channel $\mathbf{H}_i$ is $\mathbf{x}_i$. It is assumed that the elements of $\mathbf{x}_i$ are independent and identically distributed Gaussian with zero mean and unit variance, i.e., $\mathbb{E}\{\mathbf{x}_i\} = \mathbf{0}$ and $\mathbb{E}\{\mathbf{x}_i\mathbf{x}_i^{\mathrm{H}}\} = \mathbf{I}, \forall i$. The signals from different transmitters are uncorrelated, i.e., $\mathbb{E}\{\mathbf{x}_i\mathbf{x}_j^{\mathrm{H}}\} = \mathbf{0}, \forall j \neq i, \forall i$. The jammer has n_{z} antennas and the channel from the jammer to the receiver in the ith transceiver pair is denoted as $\mathbf{H}_{\mathrm{z}i}$. The maximum jamming power of the jammer is limited by P_{z}. The jammer has the knowledge of $\mathbf{H}_i$ and $\mathbf{H}_{\mathrm{z}i}$ and $\mathbf{x}_i, \forall i$ and therefore is able to perform correlated-jamming. It is assumed that the legitimate communications of all the transceiver pairs are interference-free (e.g., with frequency/time division multiplexing) or the interference is negligible (as compared to the effect of jamming).

To completely cancel out the signal from the ith transmitter using correlated jamming, the jammer should transmit a signal $-\mathbf{v}_i$ such that

$$-\mathbf{H}_{\mathrm{z}i}\mathbf{v}_i = -\mathbf{H}_i\mathbf{x}_i. \tag{5.27}$$

There exists a $\mathbf{v}_i$ that satisfies (5.27) only if the following two conditions are satis-

fied. First, the jammer must have at least the same number of antennas as that of the target receiver, i.e., $n_{\rm z} \geq n_{{\rm r}i}$. Second, $\mathbf{H}_{{\rm z}i}$ must have full rank, i.e., the rank of $\mathbf{H}_{{\rm z}i}$ should be $n_{{\rm r}i}$. Under the above two conditions, $\mathbf{v}_i$ can be given in the following form

$$\mathbf{v}_i = \mathbf{H}_{{\rm z}i}^{\rm H}(\mathbf{H}_{{\rm z}i}\mathbf{H}_{{\rm z}i}^{\rm H})^{-1}\mathbf{H}_i\mathbf{x}_i. \tag{5.28}$$

Due to the power limit of the jammer, it may not have sufficient power to transmit $-\sum_i \mathbf{v}_i$ to cancel all target signals. Therefore, the overall jammer's signal targeting at all the legitimate signals in correlated jamming can be generally expressed as

$$\mathbf{x}_{\rm z} = \sum_i \xi_i \mathbf{v}_i + \mathbf{n}_{\rm z} \tag{5.29}$$

where the weight $\xi_i \in [-1, 0]$ is determined by the power that the jammer uses targeting at the ith signal, and $\mathbf{n}_{\rm z}$ is the non-correlated noise jamming part of the jamming signal with zero mean and covariance $\sigma_{\rm z}^2\mathbf{I}$. The non-correlated part $\mathbf{n}_{\rm z}$ is uncorrelated with the legitimate signals $\mathbf{x}_i, \forall i$.

The received signal at the ith receiver can be written as

$$\begin{aligned}\mathbf{y}_i &= \mathbf{H}_i\mathbf{x}_i + \mathbf{H}_{{\rm z}i}\mathbf{x}_{\rm z} + \mathbf{n}_i \\ &= (1+\xi_i)\mathbf{H}_i\mathbf{x}_i + \mathbf{H}_{{\rm z}i}\sum_{j\neq i}\xi_j\mathbf{v}_j + \mathbf{H}_{{\rm z}i}\mathbf{n}_{\rm z} + \mathbf{n}_i\end{aligned} \tag{5.30}$$

where $\mathbf{n}_i$ is the noise at the ith receiver with zero mean and covariance $\sigma_i^2\mathbf{I}$.

Define $\mathbf{Q}_i^{\rm v} \triangleq \mathbb{E}\{\mathbf{v}_i\mathbf{v}_i^{\rm H}\}$ and $\mathbf{G}_i \triangleq \mathbf{H}_i\mathbf{H}_i^{\rm H}$. Then the information rate at receiver i in the presence of correlated jamming is given as

$$R_i^{\rm C} = \log\left|\mathbf{I} + (1+\xi_i)^2\mathbf{G}_i\Big(\mathbf{H}_{{\rm z}i}\Big(\sum_{j\neq i}\xi_j^2\mathbf{Q}_j^{\rm v} + \sigma_{\rm z}^2\mathbf{I}\Big)\mathbf{H}_{{\rm z}i}^{\rm H} + \sigma_i^2\mathbf{I}\Big)^{-1}\right|, \forall i. \tag{5.31}$$

Due to the power limit of the jammer, the following constraint must be satisfied

$$\sum_i \xi_i^2 {\rm Tr}\{\mathbf{Q}_i^{\rm v}\} + n_{\rm z}\sigma_{\rm z}^2 \leq P_{\rm z}. \tag{5.32}$$

The jammer aims at minimizing the weighted sum-rate $\sum_i w_i R_i^{\rm C}$ by optimizing $\xi_i, \forall i$ and $\sigma_{\rm z}^2$ subject to the power constraint in (5.32), where w_i's are positive weights satisfying $\sum_i w_i = 1$. If the jammer's power is sufficiently large, i.e.,

$P_{\rm z} \geq \sum_i {\rm Tr}\{\mathbf{Q}_i^{\rm v}\}$, it can completely cancel out the target signals at the corresponding receivers by setting $\xi_i = -1, \forall i$ and $\sigma_{\rm z}^2 = 0$. As a result, the rate $R_i^{\rm C}$ in (5.31) becomes zero for each i, which suggests that no information is received at any receiver. This is the ideal case for the jammer. However, a more likely situation is that $P_{\rm z}$ is not large enough to cancel all target signals, i.e., $P_{\rm z} < \sum_i {\rm Tr}\{\mathbf{Q}_i^{\rm v}\}$. In this situation, the jammer needs to jam the receivers with different priorities and optimize the weights $\xi_i, \forall i$ and $\sigma_{\rm z}^2$ in order to minimize $\sum_i w_i R_i^{\rm C}$. The focus of our investigation in multi-target correlated jamming is on finding the jamming strategy in the power-limited case, i.e., $P_{\rm z} < \sum_i {\rm Tr}\{\mathbf{Q}_i^{\rm v}\}$.

In the power-limited case, the problem of finding the optimal correlated jamming strategy can be formulated as the following optimization problem

$$\begin{aligned} \underset{\{\xi_i\},\sigma_{\rm z}^2}{\mathbf{min}} \quad & \sum_i w_i R_i^{\rm C} && \text{(5.33a)} \\ \mathbf{s.t.} \quad & \sum_i \xi_i^2 {\rm Tr}\{\mathbf{Q}_i^{\rm v}\} + n_{\rm z}\sigma_{\rm z}^2 \leq P_{\rm z} && \text{(5.33b)} \\ & -\min\{1, \sqrt{\gamma_i}\} \leq \xi_i \leq 0, \forall i. && \text{(5.33c)} \end{aligned}$$

where $\gamma_i \triangleq P_{\rm z}/{\rm Tr}\{\mathbf{Q}_i^{\rm v}\}$ with $\sqrt{\gamma_i}$ represents the maximum absolute value of ξ_i when the jammer uses all power for correlated jamming target i. The above problem is nonconvex in general.

Given the above general system model, we investigate the above problem in the case that all nodes have only a single antenna and therefore all channels are scalars, as considered in [16]. It should be noted that the difference of our work from [16] is that we consider multiple targets.

5.2.2 Multi-target correlated jamming: The SISO case

In the case that all nodes have only one antenna, $\mathbf{x}_i$, $\mathbf{x}_{\rm z}$, $\mathbf{H}_i$, $\mathbf{H}_{{\rm z}i}$, $\mathbf{n}_{\rm z}$, and $\mathbf{n}_i$ simplify to x_i, $x_{\rm z}$, h_i, $h_{{\rm z}i}$, $\sigma_{\rm z}^2$, and σ_i^2, respectively. Accordingly, $\mathbf{v}_i$, $\mathbf{Q}_i^{\rm v}$, and $\mathbf{G}_i$ simplify to

$$v_i = h_{{\rm z}i}^{-1} h_i x_i, \; q_i^{\rm v} = |h_{{\rm z}i}^{-1} h_i|^2, \; g_i = |h_i|^2 \tag{5.34}$$

respectively. Then the objective function in (5.33a) can be rewritten as

$$\sum_i w_i R_i^{\rm C} = \sum_i w_i \log\left(1 + (1+\xi_i)^2 g_i \left(|h_{{\rm z}i}|^2 (\sum_{j\neq i} \xi_j^2 q_j^{\rm v} + \sigma_{\rm z}^2) + \sigma_i^2\right)^{-1}\right). \tag{5.35}$$

With the above simplification, the following theorem is in order.

Theorem 5.3: Under the condition that the jammer uses full jamming power, i.e., $\sum_i \xi_i^2 q_i^{\mathrm{v}} + \sigma_{\mathrm{z}}^2 = P_{\mathrm{z}}$, the summation $\sum_i w_i R_i^{\mathrm{C}}$ becomes a convex function of $\xi_i, \forall i$ in the interval that $\xi_i \in [-\min\{1, \sqrt{\gamma_i}\}, 0], \forall i$, where $\gamma_i = P_{\mathrm{z}}/q_i^{\mathrm{v}}$ in the SISO case.

Proof: See Section C.6 in Appendix C.

Since the power-limited case, i.e., $P_{\mathrm{z}} < \sum_i q_i^{\mathrm{v}}$, is considered, it can be seen that using full jamming power is a necessary condition of the optimal jamming strategy. With the objective function proved to be convex under the full-power jamming condition, the solution can be found using an algorithm similar to the one in Table 5.1 used for numerically finding the solution of the optimal jamming problem in Section 5.1.3. Specifically, the problem can be first rewritten into the following equivalent disciplined form[4]

$$\begin{aligned}
\underset{\{\alpha_i\},\{\xi_i\}}{\mathbf{min}} \quad & \sum_i w_i \Big(\alpha_i - \log\big(\gamma_i - \xi_i^2 + \rho_i\big) \Big) && \text{(5.36a)} \\
\mathbf{s.t.} \quad & \alpha_i \geq \log\big(\gamma_i - \xi_i^2 + \rho_i + (1+\xi_i)^2\big), \forall i && \text{(5.36b)} \\
& \sum_i \xi_i^2 q_i^{\mathrm{v}} \leq P_{\mathrm{z}} && \text{(5.36c)} \\
& -\min\{1, \sqrt{\gamma_i}\} \leq \xi_i \leq 0, \forall i && \text{(5.36d)}
\end{aligned}$$

where $\rho_i \triangleq \sigma_i^2/g_i$. The objective function of the problem (5.36) is convex while the first constraint is not. In order to solve the problem (5.36), we further rewrite the above problem into the following form

$$\begin{aligned}
\underset{\{\alpha_i\},\{\xi_i\}}{\mathbf{min}} \quad & \sum_i w_i \Big(\alpha_i - \log(\gamma_i - \xi_i^2 + \rho_i) \Big) && \text{(5.37a)} \\
\mathbf{s.t.} \quad & \alpha_i \geq \log\Big(\gamma_i + \rho_i + 2\xi_i^{\dagger} + 1\Big) + \frac{2(\xi_i - \xi_i^{\dagger})}{\gamma_i + \rho_i + 2\xi_i^{\dagger} + 1}, \forall i && \text{(5.37b)} \\
& \sum_i \xi_i^2 q_i^{\mathrm{v}} \leq P_{\mathrm{z}} && \text{(5.37c)} \\
& -\min\{1, \sqrt{\gamma_i}\} \leq \xi_i \leq 0, \forall i && \text{(5.37d)}
\end{aligned}$$

where ${\xi_i}^{\dagger}$ stands for a given ξ_i subject to (5.36c) and (5.36d). Starting from an initial value of ${\xi_i}^{\dagger}, \forall i$, we solve the problem (5.37) and update ${\xi_i}^{\dagger}, \forall i$ using the resulting optimal solution of (5.37). The above process is repeated until convergence. The

[4]Details can be found in the proof of Theorem 5.3, Section C.6, Appendix C, and are omitted here.

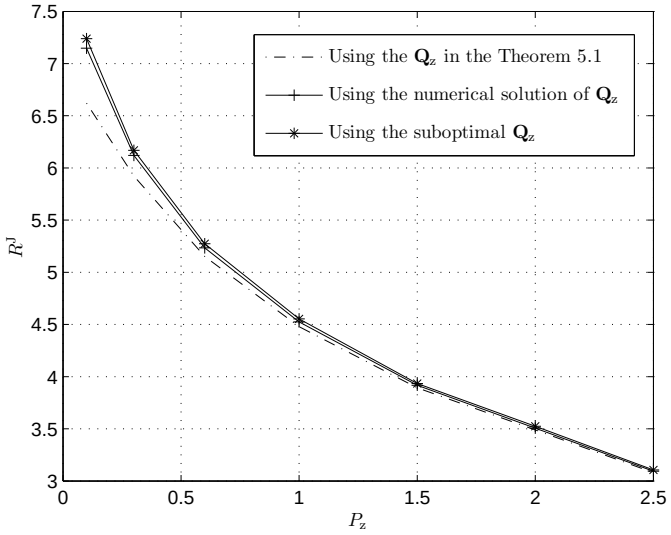


Figure 5.1: Comparison of R^{J} versus P_{z} with $\mathbf{Q}'_{\mathrm{z}}$ given by (5.15), the algorithm in Table 5.1, and (5.23), respectively.

proof of convergence to the optimal solution is similar to the proof for Lemma 5.3 and is omitted here.

5.3 Numerical and simulation results

In this section, we provide simulation examples for some results presented earlier for both non-correlated and correlated jamming.

5.3.1 The optimal and suboptimal solution for non-correlated jamming

In this simulation, we compare the rates of the legitimate communication under jamming when the jammer's strategy $\mathbf{Q}'_{\mathrm{z}}$ is given by (i) the expression in (5.15), (ii) the optimal solution obtained numerically using the algorithm in Table 5.1, and (iii) the approximation in (5.23), respectively.

The specific setup of this simulation is as follows. The number of antennas at the legitimate transmitter and receiver are set to be 4 and 3, respectively, while the number of antennas at the jammer is 5. The power limit for the legitimate transmitter is 3 and the power allocation at the legitimate transmitter is based on

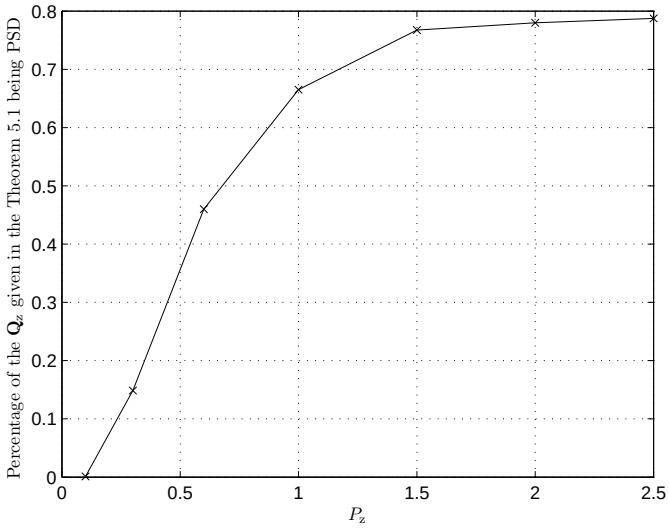


Figure 5.2: The percentage of times when the $\mathbf{Q}_{\mathrm{z}}'$ given by (5.15) is PSD versus P_{z}

waterfilling. The noise variance σ^2 is set to be 1. The elements of the target signal $\mathbf{s}$ and the channels $\mathbf{H}_{\mathrm{r}}$ and $\mathbf{H}_{\mathrm{z}}$ are generated from complex Gaussian distribution with zero mean and unit variance. As a result $\mathbf{H}_{\mathrm{r}}\mathbf{Q}_{\mathrm{s}}\mathbf{H}_{\mathrm{r}}^{\mathrm{H}}$ is always positive definite. We use 800 channel realizations and calculate the average R^{J} versus the power limit of the jammer P_{z}.

Fig. 5.1 shows the average R^{J} with $\mathbf{Q}_{\mathrm{z}}'$ obtained using the three aforementioned methods. Three observations can be made from this figure. First, there is a gap between the average R^{J} with $\mathbf{Q}_{\mathrm{z}}'$ given by (5.15) and the average R^{J} with the optimal $\mathbf{Q}_{\mathrm{z}}'$ found numerically when P_{z} is small. The gap exists because $\mathbf{Q}_{\mathrm{z}}'$ given by (5.15) is not always PSD and when it is not PSD, it no longer gives the optimal solution of the problem. Second, the gap between the average R^{J} with $\mathbf{Q}_{\mathrm{z}}'$ obtained numerically and the average R^{J} given by the suboptimal $\mathbf{Q}_{\mathrm{z}}'$ in (5.23) is very small. It verifies that the proposed suboptimal solution is in fact very close to the optimal solution of the considered problem. Third, the three curves of average R^{J} converge when P_{z} increases.

Fig. 5.2 shows the percentage that $\mathbf{Q}_{\mathrm{z}}'$ given by (5.15) is PSD in all 800 channel realizations. It verifies the aforementioned fact that $\mathbf{Q}_{\mathrm{z}}'$ given by (5.15) can be indefinite when the jammer's power limit P_{z} is small. Even when P_{z} is larger (above

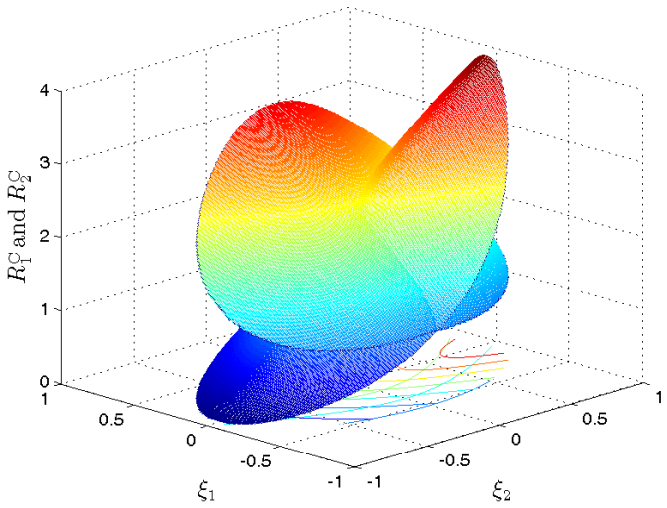


Figure 5.3: The rates R_1^{C} and R_2^{C} versus coefficients ξ_1 and ξ_2.

2), there remains a 20% chance that $\mathbf{Q}_{\mathrm{z}}'$ given by (5.15) is indefinite. It verifies the other fact that whether $\mathbf{Q}_{\mathrm{z}}'$ given by (5.15) is PSD also depends on the jamming channel.

Using the observations from the two figures, it can be seen that the suboptimal solution given by (5.23) is a very good approximation of the optimal jamming strategy since it is very close to the optimal one when $\mathbf{Q}_{\mathrm{z}}'$ given by (5.15) is indefinite while it becomes optimal when $\mathbf{Q}_{\mathrm{z}}'$ given by (5.15) is PSD.

5.3.2 The SISO correlated jamming

First, we demonstrate the rates and sum-rate of two legitimate communications under correlated jamming from one jammer when the jamming power is not necessarily fully used. The specific setup of this simulation is as follows. The number of antennas at the jammer and all legitimate transceivers is 1. The legitimate signals x_1 and x_2, the legitimate channels h_1 and h_2, and the jamming channels $h_{\mathrm{z}1}$ and $h_{\mathrm{z}2}$ are generated from complex Gaussian distribution with zero mean and unit variance. The noise covariance σ_i^2 is set to be 0.1 for both i. The power limit for the jammer is 0.5. The non-correlated jamming part n_{z} is set to be 0. Fig. 5.3 shows the rates of R_1^{C} and R_2^{C}, respectively, calculated using (5.31) versus the correlated jamming coefficients ξ_1 and ξ_2. Note that the rates are shown in the ellipse repre-

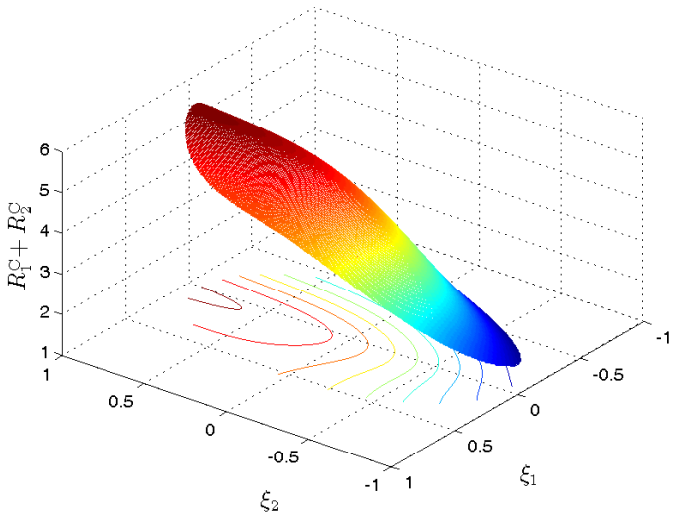


Figure 5.4: The sum-rate $R_1^{\mathrm{C}} + R_2^{\mathrm{C}}$ versus coefficients ξ_1 and ξ_2.

sented by (5.32) for better effect of displaying while only the part with $\xi_i \leq 0, \forall i$ corresponds to correlated jamming. It can be seen from the figure that the rate of legitimate communications decreases steeply with the magnitude of the corresponding coefficients, which infers that correlated jamming is effective. Fig. 5.4 shows the sum-rate of the two legitimate communications versus the correlated jamming coefficients. From the figure, it can be seen that the weighted sum-rate is a nonconvex function of ξ_1 and ξ_2.

Then we demonstrate the sum-rate of the legitimate communications under the condition that the jammer spends all of its jamming power. There are two legitimate communications and one jammer. The legitimate signals x_1 and x_2, the legitimate channels h_1 and h_2, and the jamming channels $h_{\mathrm{z}1}$ and $h_{\mathrm{z}2}$ are generated from complex Gaussian distribution with zero mean and unit variance. The power limit for the jammer P_{z} is 0.5. However, unlike the case in Fig. 5.4, in which the non-correlated jamming part n_{z} is set to be 0, the n_{z} in this simulation is calculated according to the values of x_1 and x_2 so that full jamming power is used at any point. Fig. 5.5 shows the resulting weighted sum-rate versus the correlated jamming coefficients ξ_1 and ξ_2. The diamond in the figure corresponds to the optimal solution found using the method in Section 5.2.2. The convexity of the sum-rate in the figure versus ξ_1 and ξ_2 verifies Theorem 5.3.

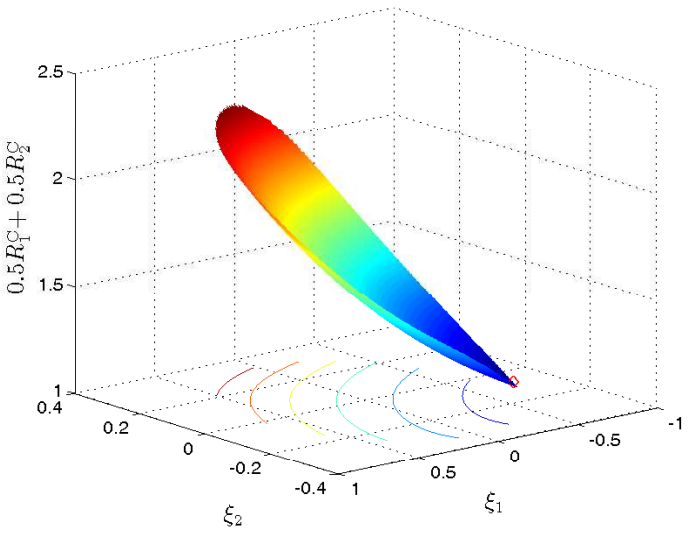


Figure 5.5: Illustration of the convexity for SISO case under the full-power jamming condition and the optimal solution found by the proposed method.

5.4 Conclusions

In this chapter, we have found the optimal noise jamming and correlated jamming for the worst-jamming multi-user systems. For the optimal noise jamming in MIMO communications, we have found the optimal jamming strategy in closed-form under PSD conditions. A numerical solution and a sub-optimal solution in closed-form have also been obtained for the case when the PSD conditions are not satisfied and the optimal solution may not be found in closed-form. For the multi-target correlated jamming, we have proved that the problem of finding the optimal jamming strategy in the SISO case is convex if the jammer always uses its full power. Simulation results demonstrate the effectiveness of the proposed solutions for both non-correlated and correlated jamming.

~

Chapter 6

Mixed Strategy Nash Equilibria in Two-User Resource Allocation Games

The objective of this chapter is to investigate mixed strategies and mixed-strategy Nash equilibrium (MSNE)[1] in non-cooperative resource allocation games in which the users' strategies are represented by continuous probability distributions with discrete distributions as special cases. The necessary and sufficient conditions for the existence and uniqueness of MSNE are derived for both two-channel and N-channel games. In the two-channel game, the MSNE which maximizes the utilities for both users is obtained, while for the N-channel game, an algorithm is provided to perform channel selection for users in order to achieve MSNE.[2]

6.1 A two-user two-channel system model

Consider first a system of two users, i.e., two transmitter-receiver pairs, sharing two channels. The total transmission power for user i is limited by $P_i^{\max}$ while each user is assumed to be able to communicate on either one or both of the channels. The users interfere with each other when they transmit on the same channel. The channel state information is known at both transmitter and receiver sides for both users and both users use Gaussian codebooks. Each user here is a player which

[1] MSNE includes pure strategy Nash equilibrium as a special case.

[2] *A version of this chapter has been published in Proc. IEEE Inter. Symp. Info. Theory,, 2011, pp. 2632-2636.*

seeks to maximize its own utility defined as its sum information rate on the two channels.

It is important to identify the condition under which the noncooperative approach that we use in this chapter is appropriate. In general, a cooperative approach tends to be more efficient, in terms of providing a larger rate region, than the noncooperative one if interference is the major impairment to the data transmission of the users. However, the cooperative approach becomes inefficient as interference power decreases given that the noise power is fixed. Thus, the noncooperative approach is preferable if noise is larger than interference. Moreover, the noncooperative approach does not require any cooperation between the users and, thus, causes no information overhead. Therefore, the study in this chapter focuses on the case of large noise power. Then the utility of user i can be approximated by[3]

$$u_i(t_i, t_j) = \frac{b_{ii}^1 t_i}{\sigma_1^2 + b_{ji}^1 t_j} + \frac{b_{ii}^2 (1 - t_i)}{\sigma_2^2 + b_{ji}^2 (1 - t_j)}, \ \forall i \tag{6.1}$$

where the approximation $\log(1+x) \approx \log(e) \cdot x$ is used and the constant multiplier $\log(e)$ is neglected, $t_i \in [0, 1]$ denotes the portion of $P_i^{\max}$ that user i allocates on channel 1, σ_k^2 is the noise power on channel k, $b_{ii}^k = P_i^{\max} |h_{ii}^k|^2$ and $b_{ij}^k = P_i^{\max} |h_{ij}^k|^2$, and h_{ii}^k and h_{ij}^k are the channel gain of the kth channel from the transmitter of user i to the receiver of user i and from the transmitter of user i to the receiver of user j, respectively.

A mixed strategy of user i is represented by the probability distribution of t_i. Denote the mixed strategy of user i as $f_i(t_i)$, which is assumed to be continuous in general, but a discrete distribution is considered as a special case. A combination of strategies $\{f_1(t_1), f_2(t_2)\}$ is called a *strategy profile*. An MSNE is a strategy profile $\{f_1^\star(t_1), f_2^\star(t_2)\}$ that satisfies [89]

$$\begin{aligned} &\mathbb{E}_{t_i, t_j} \{ u_i(t_i, t_j) |_{f_i(t_i) = f_i^\star(t_i), f_j(t_j) = f_j^\star(t_j)} \} \\ &= \max_{f_i(t_i)} \mathbb{E}_{t_i, t_j} \{ u_i(t_i, t_j) |_{f_i(t_i), f_j(t_j) = f_j^\star(t_j)} \}, \forall i. \end{aligned} \tag{6.2}$$

[3]Similar cases are considered in [86] and [87], while received signal-to-interference-plus-noise ratios are directly chosen as users' utilities in [88] instead of the information rates. Note also that we neglect the condition $i \neq j$ for brevity and assume it as default when applies.

It can be proved that $\{f_1^\star(t_1), f_2^\star(t_2)\}$ satisfies (6.2) iff the conditions

$$\mathbb{E}_{t_j}\{u_i(t_i,t_j)\} = c_i,\ \forall t_i \in \mathcal{S}_i^\star, \tag{6.3}$$

$$\mathbb{E}_{t_j}\{u_i(t_i,t_j)\} \geq E_{t_j}\{u_i(t_i^{'},t_j)\}, \forall t_i \in \mathcal{S}_i^\star,\ \forall t_i^{'} \notin \mathcal{S}_i^\star \tag{6.4}$$

are satisfied for all i given that $f_j(t_j) = f_j^\star(t_j)$, where $\mathcal{S}_i^\star = \{t_i \,|\, t_i \in [0,1], f_i^\star(t_i) \neq 0\}$ is defined as the *support* of $f_i^\star(t_i)$ and c_i is a constant. The following is a brief illustration of the necessity. If (6.3) is not satisfied or, equivalently, if there exist $t_i^1, t_i^2 \in \mathcal{S}_i^\star$ such that $\mathbb{E}_{t_j}\{u_i(t_i^2,t_j)\} > \mathbb{E}_{t_j}\{u_i(t_i^1,t_j)\}$, then $u_i(t_i,t_j)$ can be increased by transferring the probability density assigned on t_i^1 to t_i^2. If (6.4) is not satisfied or, equivalently, if there exist $t_i^3 \in \mathcal{S}_i^\star$ and $t_i^4 \notin \mathcal{S}_i^\star$ such that $\mathbb{E}_{t_j}\{u_i(t_i^4,t_j)\} > \mathbb{E}_{t_j}\{u_i(t_i^3,t_j)\}$, then $u_i(t_i,t_j)$ can be increased by transferring the probability density assigned on t_i^3 to t_i^4. The illustration for the sufficiency is straightforward and neglected here due to the space limit.

6.2 MSNE in a two-user two-channel game

The following theorem provides a result on the existence and uniqueness of MSNE in the considered game.

Theorem 6.1: *The considered two-user game has either a unique or infinitely many MSNEs. The necessary and sufficient condition for the latter is*

$$\frac{\sigma_1^2}{\sigma_2^2 + b_{ji}^2} \leq \frac{b_{ii}^1}{b_{ii}^2} \leq \frac{\sigma_1^2 + b_{ji}^1}{\sigma_2^2},\ \forall i. \tag{6.5}$$

According to Theorem 6.1, there could be infinitely many MSNEs in the considered two-user game, which can lead to different utilities for the users. Therefore, it is also of interest to investigate the most efficient MSNE.

Theorem 6.2: *In the case when there exist infinitely many MSNEs in the considered game, the one MSNE among all which maximizes the utilities for both users is*

$$\tilde{f}_j(t_j) = \xi_j \delta(t_j) + (1-\xi_j)\delta(t_j - 1),\ \forall j \tag{6.6}$$

where $\delta(\cdot)$ is the Dirac delta function and

$$\xi_j = \frac{\frac{b_{ii}^2}{\sigma_2^2} - \frac{b_{ii}^1}{\sigma_1^2+b_{ji}^1}}{\frac{b_{ij}^2}{\sigma_2^2} - \frac{b_{ii}^1}{\sigma_1^2+b_{ji}^1} + \frac{b_{ij}^1}{\sigma_1^2} - \frac{b_{ii}^2}{\sigma_2^2+b_{ji}^2}}, \ \forall j. \tag{6.7}$$

6.3 Extension to a two-user N-channel game

The two-user N-channel case is more general yet complicated. In the N-channel case, the utility of user i extends as

$$u_i(\mathbf{t}_i, \mathbf{t}_j) = \sum_{k=1}^{N-1} \frac{b_{ii}^k t_i^k}{\sigma_k^2 + b_{ji}^k t_j^k} + \frac{b_{ii}^N (1 - \sum_{k=1}^{N-1} t_i^k)}{\sigma_N^2 + b_{ji}^N (1 - \sum_{k=1}^{N-1} t_j^k)}, \forall i \tag{6.8}$$

where $\mathbf{t}_i = [t_i^1, \ldots, t_i^{N-1}]$ and $t_i^k \in [0,1]$ is the portion of $P_i^{\max}$ that user i allocates on channel k subject to $\sum_{k=1}^{N-1} t_i^k \in [0,1]$. Conditions (6.3)-(6.4) extend accordingly as

$$\mathbb{E}_{\mathbf{t}_j}\{u_i(\mathbf{t}_i, \mathbf{t}_j)\} = c_i, \ \forall \mathbf{t}_i \in \mathcal{S}_i^\star, \tag{6.9}$$

$$\mathbb{E}_{\mathbf{t}_j}\{u_i(\mathbf{t}_i, \mathbf{t}_j)\} \geq \mathbb{E}_{\mathbf{t}_j}\{u_i(\mathbf{t}_i', \mathbf{t}_j)\}, \forall \mathbf{t}_i \in \mathcal{S}_i^\star, \ \forall \mathbf{t}_i' \notin \mathcal{S}_i^\star \tag{6.10}$$

with the mixed strategy of user i now represented by the joint distribution of $t_i^k, \forall k \in \{1, \ldots, N-1\}$ and denoted as $f_i(\mathbf{t}_i)$.

The existence and uniqueness of MSNE in the N-channel game can be derived based on the outputs of the algorithm in Table 6.1.

Theorem 6.3. *The following properties hold for the proposed algorithm in Table I.*

i) The algorithm converges to the same result regardless of the ordering of users or channels.

ii) Denote $\Gamma_i = \{k \in \Delta_i, k \notin \Delta_j\}$, then $L(\Gamma_i) \leq 1, \forall i$ at the output of the algorithm.

iii) MSNE is unique in the game iff $L(\Delta_{i=1}) = 1$ or $L(\Delta_{i=2}) = 1$. Otherwise, infinitely many MSNEs exist.[4]

[4]There is a trivial exception. If $\Delta_i = \{k_1\}$, $\Delta_j = \{k_1, k_2\}$ and $\nu_j^1(k_2) = \nu_j^2(k_1)$ where $k_1, k_2 \in \{1, \ldots, N\}$, infinitely many MSNEs exist. However, in this case all other MSNEs generate smaller utilities for user i (and the same utility for user j) than the MSNE which is achievable and unique in the case when $\Delta_i = \{k_1\}$, $\Delta_j = \{k_2\}$. Therefore, we consider the former case as equivalent to the latter one.

Table 6.1: Algorithm for channel selection in two-user N channel game

1. Let $\nu_i^1 = [b_{ii}^1/\sigma_1^2, \ldots, b_{ii}^N/\sigma_N^2]$ and $\nu_i^2 = [b_{ii}^1/(\sigma_1^2 + b_{ji}^1), \ldots, b_{ii}^N/(\sigma_N^2 + b_{ji}^N)]$ for each user i. Let $\Delta_{i=1} = \Delta_{i=2} = \{1, \ldots, N\}$. Initialize $d = 1$.
2. For user $i = 1$, let $k = \Delta_{i=1}(d)$, where $\Delta_{i=1}(d)$ is the dth element of the set $\Delta_{i=1}$, and check if the inequalities $\nu_1^1(k) > \nu_1^2(l), \forall l \in \Delta_{i=1} \neq k$ are all satisfied. If not, let $t_1^k = 0$, $\nu_2^2(k) = b_{22}^k/\sigma_k^2$, remove $\Delta_{i=1}(d)$ from $\Delta_{i=1}$ and set $d = d - 1$. Check if $d < L(\Delta_{i=1})$ where $L(\cdot)$ denotes the cardinality of a set. If yes, set $d = d + 1$ and repeat the above procedure in Step 2. If no, set $d = 1$ and proceed to Step 3.
3. For user $i = 2$, let $k = \Delta_{i=2}(d)$ and check if the inequalities $\nu_2^1(k) > \nu_2^2(l), \forall l \in \Delta_{i=2} \neq k$ are all satisfied. If not, let $t_2^k = 0$, $\nu_1^2(k) = b_{11}^k/\sigma_k^2$, remove $\Delta_{i=2}(d)$ from $\Delta_{i=2}$ and set $d = d - 1$. Check if $d < L(\Delta_{i=2})$, If yes, set $d = d + 1$ and repeat the above procedure in Step 3. If no, and no element was deleted from $\Delta_{i=2}$ in this step, proceed to Step 4; otherwise set $d = 1$ and return to the beginning of Step 2.
4. Output $\Delta_{i=1}$ and $\Delta_{i=2}$.

6.4 Numerical and simulation results

Our simulation example illustrates the iterative process of channel selection described in Table 6.1. Here $N = 8$, $P_i^{\max} = 1, \forall i$, and $\sigma_k^2, \forall k$ are uniformly generated from the interval $[1, 2]$. The real and imaginary parts of h_{ii}^k and $h_{ij}^k \forall i, \forall k$ are generated from zero-mean normal distributions with variances 1 and 0.25, respectively. The results are shown in Fig. 6.1, where the diamonds and squares are generated at coordinates $(Re(h_{11}^k), Im(h_{11}^k), (|h_{21}|^k)^2), \forall k$ and $(Re(h_{22}^k), Im(h_{22}^k), (|h_{12}|^k)^2), \forall k$, respectively. The diamond and square corresponding to the same k are connected by dash-dot lines for all k. A diamond/square closer to the corners implies a channel with higher channel gain for the corresponding user, while a diamond/square closer to the top implies a channel with higher gain of interference from the transmitter of the other user to the receiver of the corresponding user.

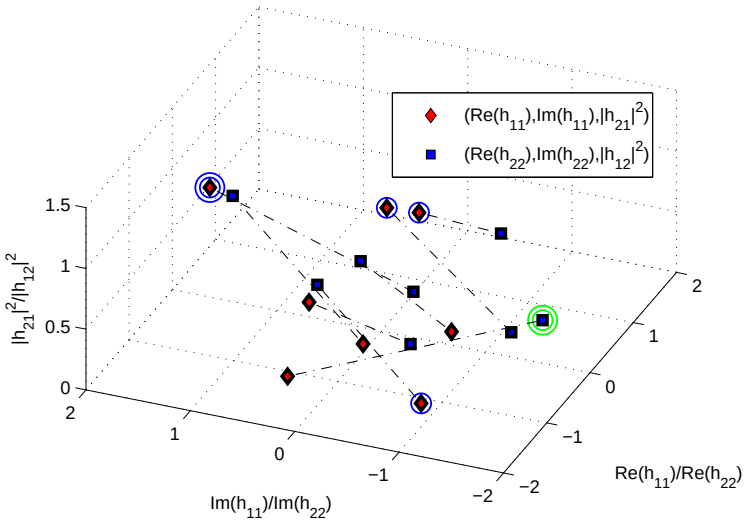
At the end of Step 2/Step 3, the diamonds/squares corresponding to the channel indexes in the updated Δ_1/Δ_2 are circled in Fig. 6.1. If the algorithm iterates, the diamonds/squares corresponding to the channel indexes in the most updated Δ_1/Δ_2 are circled by circles with a larger radius at the end of each iteration of Step 2/Step 3. The diamonds/squares with the maximum number of circles correspond to the chan-

nel indexes in Δ_1 and Δ_2 at the output of the algorithm. The upper plot of Fig. 6.1 shows the case of $\Delta_1 = 1, \Delta_2 = 1$, in which a unique MSNE exists according to Theorem 6.3. It can be seen that the first run of Step 2 selects four channels for user 1 while the second run further selects one out of the four. The lower plot of Fig. 6.1 shows the case of $\Delta_1 = 2, \Delta_2 = 3$, in which two of the eight channels are shared and infinitely many MSNEs exist according to Theorem 6.3. Note that the users interfere with each other only on the channels corresponding to the dash-dot line with the maximum number of circles at both ends in the plots. From the figure, it can be seen that the channels selected by the users achieve high channel gains and low interference.

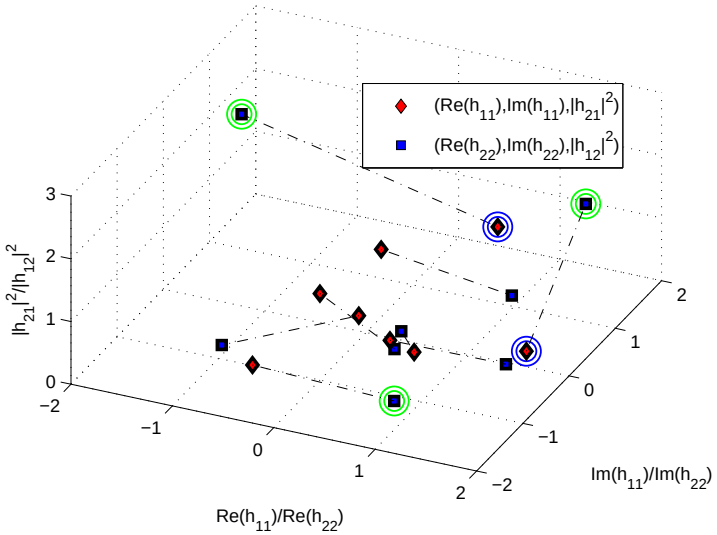
6.5 Conclusion

Noncooperative resource allocation games are studied in mixed strategies. It is shown that applying mixed strategies can potentially lead to MSNE which is more efficient than NE in pure strategies. For two-channel games, the sufficient and necessary condition for the uniqueness of MSNE is derived. The game becomes significantly more complicated in the case of N channels. A channel selection algorithm which simplifies the game is proposed. Based on the outputs of the algorithm, the sufficient and necessary condition for the uniqueness of MSNE in this game is also derived. Our simulation results demonstrate how the proposed algorithm selects channels in the N-channel game.

$\sim$



(a) At the output of the algorithm $\Delta_1 = 1, \Delta_2 = 1$



(b) At the output of the algorithm $\Delta_1 = 2, \Delta_2 = 3$

Figure 6.1: Illustration of the cannel selection algorithm in two examples

Chapter 7

Conclusion and Future Work

The problems of spectral and power efficiency, worst-case jamming threat, and resource allocation in multi-user wireless communications are studied in this thesis. The following main results are derived.

- Chapter 3 obtains the optimal solution to the problem of maximizing spectral efficiency, in terms of sum-rate, with minimum relay power consumption in MIMO DF TWR when there is limited coordination in the system. It is shown that the relay optimization scenario may not be energy-efficient to the source nodes as they can possibly waste part of their transmission power. The asymmetry is shown to have negative effect on the spectrum and power efficiency of the DF TWR.

- Chapter 4 obtains the optimal solution to the problem of maximizing spectral efficiency with minimum total power consumption in MIMO DF TWR with full coordination. The optimal solution is found in closed-form or using proposed algorithms. It is shown that the cooperation among the participating nodes can dramatically improve energy-efficiency in the system while at the same time achieving the same or better spectrum efficiency. The negative effect of asymmetry is also demonstrated.

- Chapter 5 obtains the optimal form of noise jamming for the worst-case jamming in multi-user wireless communications. The optimal noise jamming is shown to be in closed-form under certain conditions. A sub-optimal noise jamming is proposed in closed form and shown to be a good approximation

of the optimal jamming solution when the latter cannot found in closed-form. The problem of worst-case multi-target correlated jamming is proved to be convex in the SISO case under the condition that the jammer uses its full power.

- Chapter 6 obtains the conditions on the existence and uniqueness of the MSNE in a continuous resource allocation game with mixed-strategies. For the two-user two-channel case, the most efficient MSNE is found. It is shown that mixed-strategy Nash equilibria (MSNEs) are more efficient than the Nash equilibria (NEs) in pure strategies in the considered game. For the two-user multiple channel case, an algorithm is proposed for achieving the MSNE.

There are some open problems related to the topics of this thesis, which will be considered in future work.

- With respect to spectral and energy efficiency, the study of sum-rate maximization with minimum power consumption in TWR can be extended by considering the optimal time division between the MA and BC phases. In Chapter 3 and 4, it is assumed that each of the MA and BC phases takes a half of the entire time of message exchange. However, it is not necessarily the optimal division of time between the two phases. The spectral efficiency can be further improved if the length of the MA and BC phases are optimally divided. It is also possible to extend the study of sum-rate maximization with minimum power consumption to multiuser TWR with multiple pairs of source nodes and one or multiple relays. Relay selection for source nodes can be taken into account. The relays can help multiple pairs of source nodes exchange information by adopting time/frequency division multiple access.

- With respect to the worst-case jamming threat, the study of optimal noise jamming strategy can be extended to the case of multiple legitimate communications. Optimal solution or sub-optimal in closed-form could be obtained. For the optimal correlated jamming, it could be extended to the general MIMO case with the objective to find an efficient suboptimal jamming strategy. It

could also be extended by adopting game theory to study the interactions between the legitimate transceivers and the jammer(s).

- With respect to the multi-user resource allocation game, the study of the existence and uniqueness of MSNE could be extended by considering the multi-user multi-channel game with mixed strategy. It is also of interest to consider the case that the utilities of the users are of uncertainty, i.e., the utilities of the users are subject to small fluctuations.

~

Bibliography

[1] D. Raychaudhuri and N. B. Mandayam, "Frontiers of wireless and mobile communications," *Proc. IEEE*, vol. 100, no. 4, pp. 824–840, 2012.

[2] A. Sheikh, *Wireless Communications: Theory and Techniques*. Kluwer Academics Publishers, 2004.

[3] R. Fan and H. Jiang, "Optimal multi-channel cooperative sensing in cognitive radio networks," vol. 9, no. 3, pp. 1128–1138, 2010.

[4] R. Baldemair, E. Dahlman, G. Fodor, G. Mildh, S. Parkvall, Y. Selen, H. Tullberg, and K. Balachandran, "Evolving wireless communications: Addressing the challenges and expectations of the future," *IEEE Veh. Technol. Mag.*, vol. 8, no. 1, pp. 24–30, 2013.

[5] E. Telatar, "Capacity of multi-antenna gaussian channels," *European Trans. Telecommun.*, vol. 10, no. 6, pp. 585–595, Nov. 1999.

[6] L. Zheng and D. Tse, "Diversity and multiplexing: a fundamental tradeoff in multiple-antenna channels," *IEEE Trans. Inf. Theory*, vol. 49, no. 5, pp. 1073 – 1096, May 2003.

[7] B. Rankov and A. Wittneben, "Spectral efficient protocols for half-duplex fading relay channels," *IEEE J. Select. Areas Commun.*, vol. 25, no. 2, pp. 379–389, Feb. 2007.

[8] D. Feng, C. Jiang, G. Lim, L. J. Cimini, G. Feng, and G. Y. Li, "A survey of energy-efficient wireless communications," *IEEE Commun. Surveys Tutorials*, vol. 15, no. 1, pp. 167–178, 2013.

[9] G. Li, Z. Xu, C. Xiong, C. Yang, S. Zhang, Y. Chen, and S. Xu, "Energy-efficient wireless communications: tutorial, survey, and open issues," *IEEE Wireless Commun.*, vol. 18, no. 6, pp. 28–35, 2011.

[10] X. Lu, E. Erkip, Y. Wang, and D. Goodman, "Power efficient multimedia communication over wireless channels," *IEEE J. Select. Areas Commun.*, vol. 21, no. 10, pp. 1738–1751, 2003.

[11] X. Wang, D. Wang, H. Zhuang, and S. Morgera, "Fair energy-efficient resource allocation in wireless sensor networks over fading tdma channels," *IEEE J. Select. Areas Commun.*, vol. 28, no. 7, pp. 1063–1072, 2010.

[12] W. Xu, W. Trappe, Y. Zhang, and T. Wood, "The feasibility of launching and detecting jamming attacks in wireless networks," in *Proc. ACM Int. Symp. Mobile ad hoc Netw. and Comput.*, New York, NY, USA, 2005.

[13] R. Poisel, *Modern Communications Jamming: Principles and Techniques*, 2nd ed. Artech House, 2011.

[14] M. Li, I. Koutsopoulos, and R. Poovendran, "Optimal jamming attack strategies and network defense policies in wireless sensor networks," vol. 9, no. 8, pp. 1119–1133, Aug. 2010.

[15] H.-M. Sun, S.-P. Hsu, and C.-M. Chen, "Mobile jamming attack and its countermeasure in wireless sensor networks," in *Int. Conf. Advanced Info. Netw. and Applicat. Workshops*, 2007, pp. 457–462.

[16] A. Kashyap, T. Basar, and R. Srikant, "Correlated jamming on MIMO Gaussian fading channels," *IEEE Trans. Inf. Theory*, vol. 50, no. 9, pp. 2119–2123, Sept. 2004.

[17] M. Brady, M. Mohseni, and J. Cioffi, "Spatially-correlated jamming in gaussian multiple access and broadcast channels," in *40th Annu. Conf. Info. Sci. and Syst.*, 2006, pp. 1635–1639.

[18] A. Ali, “Worst-case partial-band noise jamming of Rician fading channels,” *IEEE Trans. Commun.*, vol. 44, no. 6, pp. 660–662, 1996.

[19] R. Wang and M. Tao, “Joint source and relay precoding designs for MIMO two-way relaying based on MSE criterion,” *IEEE Trans. Signal Process.*, vol. 60, no. 3, pp. 1352–1365, May 2012.

[20] A. Khabbazibasmenj and S. A. Vorobyov, “Power allocation based on SEP minimization in two-hop decode-and-forward relay networks,” *IEEE Trans. Signal Process.*, vol. 59, no. 8, pp. 3954–3963, Aug. 2011.

[21] M. Pischella and D. L. Ruyet, “Optimal power allocation for the two-way relay channel with data rate fairness,” *IEEE Commun. Lett.*, vol. 15, no. 9, pp. 959–961, Sept. 2011.

[22] I. Hammerstrom, M. Kuhn, C. Esli, J. Zhao, A. Wittneben, and G. Bauch, “MIMO two-way relaying with transmit CSI at the relay,” in *IEEE Workshop on Signal Processing advances in Wireless Commun.*, 2007.

[23] K. Jitvanichphaibool, R. Zhang, and Y.-C. Liang, “Optimal resource allocation for two-way relay-assisted OFDMA,” *IEEE Trans. Veh. Technol.*, vol. 58, no. 7, pp. 3311–3321, Sept. 2009.

[24] Q. F. Zhou, Y. Li, F. C. M. Lau, and B. Vucetic, “Decode-and-forward two-way relaying with network coding and opportunistic relay selection,” *IEEE Trans. Commun.*, vol. 58, no. 11, pp. 3070–3076, Nov. 2010.

[25] I. Krikidis, “Relay selection for two-way relay channels with MABC DF: a diversity perspective,” *IEEE Trans. Veh. Technol.*, vol. 59, no. 9, pp. 4620–4628, Nov. 2010.

[26] P. Liu and I.-M. Kim, “Performance analysis of bidirectional communication protocols based on decode-and-forward relaying,” *IEEE Trans. Commun.*, vol. 58, no. 9, pp. 2683–2696, Sept. 2010.

[27] T. J. Oechtering, R. F. Wyrembelski, and H. Boche, "Multiantenna bidirectional broadcast channels – optimal transmit strategies," *IEEE Trans. Signal Process.*, vol. 57, no. 5, pp. 1948–1958, May 2009.

[28] A. Khabbazibasmenj, F. Roemer, S. A. Vorobyov, and M. Haardt, "Sum-rate maximization in two-way AF MIMO relaying: Polynomial time solutions to a class of dc programming problems," *IEEE Trans. Signal Process.*, vol. 60, no. 10, pp. 5478–5493, Oct. 2012.

[29] S. Xu and Y. Hua, "Optimal design of spatial source-and-relay matrices for a non-regenerative two-way MIMO relay system," *IEEE Trans. Wireless Commun.*, vol. 10, no. 5, pp. 1645–1655, May 2011.

[30] C. Y. Leow, Z. Ding, and K. K. Leung, "Joint beamforming and power management for nonregenerative MIMO two-way relaying channels," *IEEE Trans. Veh. Technol.*, vol. 60, no. 9, pp. 4374–4383, Nov. 2011.

[31] J. Zou, H. Luo, M. Tao, and R. Wang, "Joint source and relay optimization for non-regenerative MIMO two-way relay systems with imperfect CSI," *IEEE Trans. Wireless Commun.*, vol. 11, no. 9, pp. 3305–3315, Sept. 2012.

[32] S. J. Kim, N. Devroye, P. Mitran, and V. Tarokh, "Achievable rate regions and performance comparison of half duplex bi-directional relaying protocols," *IEEE Trans. Inf. Theory*, Oct. 2011.

[33] T. J. Oechtering and H. Boche, "Stability region of an optimized bidirectional regenerative half-duplex relaying protocol," *IEEE Trans. Commun.*, vol. 56, no. 9, pp. 1519–1529, Sept. 2008.

[34] ——, "Optimal transmit strategies in multi-antenna bidirectional relaying," in *Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Process. (ICASSP)*, Honolulu, USA, Apr. 2007, pp. 145–148.

[35] T. J. Oechtering, E. A. Jorswieck, R. F. Wyrembelski, and H. Boche, "On the optimal transmit strategy for the mimo bidirectional broadcast channel," *IEEE Trans. Commun.*, vol. 57, no. 12, pp. 3817–3826, Dec. 2009.

[36] V. Havary-Nassab, S. Shahbazpanahi, and A. Grami, "Optimal distributed beamforming for two-way relay networks," *IEEE Trans. Signal Process.*, vol. 58, no. 3, pp. 1238–1250, 2010.

[37] M. Zaeri-Amirani, S. Shahbazpanahi, T. Mirfakhraie, and K. Ozdemir, "Performance tradeoffs in amplify-and-forward bidirectional network beamforming," *IEEE Trans. Signal Process.*, vol. 60, no. 8, pp. 4196–4209, 2012.

[38] T. Basar, "The gaussian test channel with an intelligent jammer," *IEEE Trans. Inf. Theory*, vol. 29, no. 1, pp. 152–157, 1983.

[39] M. Wilhelm, I. Martinovic, J. B. Schmitt, and V. Lenders, "Short paper: reactive jamming in wireless networks: how realistic is the threat?" in *Proc. Fourth ACM conf. Wireless Network Security*, 2011, pp. 47–52.

[40] H. Kwon and P. Lee, "Combined tone and noise jamming against coded FH/MFSK ECCM radios," *IEEE J. Select. Areas Commun.*, vol. 8, no. 5, pp. 871–883, June 1990.

[41] S. Prasad and D. Thuente, "Jamming attacks in 802.11g - a cognitive radio based approach," in *Military Commun. Conf.*, Baltimore, USA, Nov. 2011, pp. 1219–1224.

[42] X. Zhou, B. Maham, and A. Hjorungnes, "Pilot contamination for active eavesdropping," *IEEE Trans. Wireless Commun.*, vol. 11, no. 3, pp. 903–907, 2012.

[43] M. M*é*dard, "Capacity of correlated jamming channels," in *Proc. Allerton Conf. Commun. Comput. and Control*, 1997, pp. 1043–1052.

[44] S. Shafiee and S. Ulukus, "Mutual information games in multiuser channels with correlated jamming," *IEEE Trans. Inf. Theory*, vol. 55, no. 10, pp. 4598–4607, 2009.

[45] A. Bayesteh, M. Ansari, and A. K. Khandani, "Effect of jamming on the capacity of MIMO channels," in *Proc. Allerton Conf. Commun. Comput. and Control*, Monticello, USA, September 2004, pp. 401–410.

[46] H. B. E. A. Jorswieck and M. Weckerle, "Optimal transmitter and jamming strategies in gaussian mimo channels,," in *Proc. IEEE Veh. Tech. Conf.*, Stockholm, Sweden, May 2005, pp. 978–982.

[47] M. R. D. Rodrigues and G. Ramos, "On multiple-input multiple-output gaussian channels with arbitrary inputs subject to jamming," in *Proc. IEEE Int. Symp. Info. Theory*, Seoul, Korea, June 2009, pp. 2512–2516.

[48] R. Myerson, *Game Theory: Analysis of Conflict*. Harvard University Press, 1997.

[49] M. J. Osborne, *An introduction to Game Theory*, 1st ed. Oxford University Press, 2003.

[50] Y. Zhang and M. Guizani, Eds., *Game Theory for Wireless Communications and Networking*, 1st ed. CRC Press, 2011.

[51] J. Chen and A. Swindlehurst, "Applying bargaining solutions to resource allocation in multiuser MIMO-OFDMA broadcast systems," *IEEE J. Select. Areas Commun.*, vol. 6, no. 2, pp. 127–139, 2012.

[52] E. Larsson, E. Jorswieck, J. Lindblom, and R. Mochaourab, "Game theory and the flat-fading gaussian interference channel," *IEEE Trans. Signal Process.*, vol. 26, no. 5, pp. 18–27, 2009.

[53] G. Scutari, D. Palomar, F. Facchinei, and J.-S. Pang, "Convex optimization, game theory, and variational inequality theory," *IEEE Signal Processing Magazine*, vol. 27, no. 3, pp. 35–49, 2010.

[54] G. Scutari, D. Palomar, and S. Barbarossa, "Competitive design of multiuser MIMO systems based on game theory: A unified view," *IEEE J. Select. Areas Commun.*, vol. 26, no. 7, pp. 1089–1103, 2008.

[55] Z. Han and K. Liu, "Noncooperative power-control game and throughput game over wireless networks," *IEEE Trans. Commun.*, vol. 53, no. 10, pp. 1625–1629, 2005.

[56] S.-L. Hew and L. White, "Cooperative resource allocation games in shared networks: symmetric and asymmetric fair bargaining models," *IEEE Trans. Wireless Commun.*, vol. 7, no. 11, pp. 4166–4175, 2008.

[57] H. Boche and M. Schubert, "A generalization of nash bargaining and proportional fairness to log-convex utility sets with power constraints," *IEEE Trans. Inf. Theory*, vol. 57, no. 6, pp. 3390–3404, 2011.

[58] E.-V. Belmega, S. Lasaulce, and M. Debbah, "Power allocation games for MIMO multiple access channels with coordination," *IEEE Trans. Wireless Commun.*, vol. 8, no. 6, pp. 3182–3192, 2009.

[59] C. St.Jean and B. Jabbari, "On game-theoretic power control under successive interference cancellation," *IEEE Trans. Wireless Commun.*, vol. 8, no. 4, pp. 1655–1657, 2009.

[60] J. Gao, S. Vorobyov, and H. Jiang, "Cooperative resource allocation games under spectral mask and total power constraints," *IEEE Trans. Signal Process.*, vol. 58, no. 8, pp. 4379–4395, 2010.

[61] Y. Wu and D. H. K. Tsang, "Distributed power allocation algorithm for spectrum sharing cognitive radio networks with QoS guarantee," in *IEEE Int. Conf. Computer Commu. (INFOCOM)*, Rio de Janerio, Brazil, Apr. 2009, pp. 981–989.

[62] N. B. Chang and M. Liu, "Optimal competitive algorithms for opportunistic spectrum access," *IEEE J. Select. Areas Commun.*, vol. 26, no. 7, pp. 1183–1192, Sept. 2008.

[63] A. Rubinstein and M. J. Osborne, *A Course in Game Theory*, 1st ed. Cambridge, MA, USA: MIT Press, 1994.

[64] S. Brahma and M. Chatterjee, "Mitigating self-interference among IEEE 802.22 networks: A game theoretic perspective," in *Proc. IEEE Global Telecommn. Conf. (GLOBECOM)*, 2009.

[65] O. Awwad, A. Al-Fuqaha, B. Khan, D. Benhaddou, M. Guizani, and A. Rayes, "Bayesian-based game theoretic model to guarantee cooperativeness in hybrid rf/fso mesh networks," in *Proc. IEEE Global Telecommn. Conf. (GLOBECOM)*, 2009.

[66] H. N. Pham, J. Xiang, Y. Zhang, and T. Skeie, "QoS-aware channel selection in cognitive radio networks: A game-theoretic approach," in *Proc. IEEE Global Telecommn. Conf. (GLOBECOM)*, Dec. 2008.

[67] A. Goldsmith, S. A. Jafar, N. Jindal, and S. Vishwanath, "Capacity limits of MIMO channels," *IEEE J. Select. Areas Commun.*, vol. 21, no. 5, pp. 684–702, June 2003.

[68] R. Bhatia, *Matrix Analysis*. Heidelberg, Germany: Springer, 1997.

[69] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*, 1st ed. New York, USA: Cambridge University Press, 2005.

[70] W. Yu, R. Wonjong, S. Boyd, and J. M. Cioffi, "Iterative water-filling for Gaussian vector multiple-access channels," *IEEE Trans. Inf. Theory*, vol. 50, no. 1, pp. 145–152, Jan. 2004.

[71] M. Chen and A. Yener, "Power allocation for F/TDMA multiuser two-way relay networks," *IEEE Trans. Wireless Commun.*, vol. 9, no. 2, pp. 546–551, Feb. 2010.

[72] P. Popovski and H. Yomo, "Physical network coding in two-way wireless relay channels," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Glasgow, Scotland, 2007, pp. 707–712.

[73] C.-H. Liu and F. Xue, “Network coding for two-way relaying: Rate region, sum rate and opportunistic scheduling,” in *Proc. IEEE Int. Conf. Commun. (ICC)*, Beijing, China, 2008, pp. 1044–1049.

[74] A. Agustin, J. Vidal, and O. Munoz, “Protocols and resource allocation for the two-way relay channel with half-duplex terminals,” in *Proc. IEEE Int. Conf. Commun. (ICC)*, 2009.

[75] Y. Shim, H. Park, and H. M. Kwon, “Optimal power allocation for two-way decode-and-forward relay networks with equal transmit power at source nodes,” in *IEEE Wireless Commun. and Networking Conf. (WCNC)*, Shang Hai, China, 2013, pp. 3335–3340.

[76] Y. Jia and A. Vosoughi, “Two-way relaying for energy constrained systems: Joint transmit power optimization,” in *Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Process. (ICASSP)*, Dallas, USA, 2010, pp. 3026–3029.

[77] E. Rasmusen, *Games and Information: An Introduction to Game Theory*. Wiley, 2006.

[78] M. Nakayama, “The dawn of modern theory of games,” in *Advances in Mathematical Economics*, ser. Advances in Mathematical Economics, S. Kusuoka and A. Yamazaki, Eds. Springer Japan, 2006, vol. 9, pp. 73–97.

[79] C. H. Liu and F. Xue, “Network coding for two-way relaying: rate region, sum rate and opportunistic scheduling,” in *Proc. IEEE Int. Conf. Commun. (ICC)*, Beijing, China, May 2008, pp. 1044–1049.

[80] J. Liu, M. Tao, Y. Xu, and X. Wang, “Superimposed XOR: a new physical layer network coding scheme for two-way relay channels,” in *Proc. IEEE Global Telecommn. Conf. (GLOBECOM)*, Dec. 2009.

[81] *CVX: Matlab software for disciplined convex programming*, Aug. 2013. [Online]. Available: http://cvxr.com/cvx/

[82] I. Hammerstrom, M. Kuhn, C. Esli, J. Zhao, A. Wittneben, and G. Bauch, "MIMO two-way relaying with transmit CSI at the relay," in *IEEE Workshop on Signal Processing advances in Wireless Commun.*, June 2007.

[83] S. N. Diggavi and T. M. Cover, "The worst additive noise under a covariance constraint," *IEEE Trans. Inf. Theory*, vol. 47, no. 7, pp. 3072–3081, Nov. 2001.

[84] S. B. Michael Grant and Y. Ye, *Global Optimization: From Theory to Implementation*, L. Liberti and N. Maculan, Eds. Springer, 2006.

[85] S. Shafiee and S. Ulukus, "Mutual information games in multiuser channels with correlated jamming," *IEEE Trans. Inf. Theory*, vol. 55, no. 10, pp. 4598–4607, Oct. 2009.

[86] M. C. Gursoy, "Secure communication in the low-SNR regime: A characterization of the energy-secrecy tradeoff," in *IEEE Int. Symposium on Inform. Theory*, Seoul, Korea, July 2009, pp. 2291–2295.

[87] J. N. Laneman, D. N. C. Tse, and G. W. Wornell, "Cooperative diversity in wireless networks: Efficient protocols and outage behavior," *IEEE Trans. Inf. Theory*, vol. 50, no. 12, pp. 3062–3080, Dec. 2004.

[88] E. Altman, A. Kumar, C. Singh, and R. Sundaresan, "Spatial SINR games combining base station placement and mobile association," in *IEEE Int. Conf. Computer Commu. (INFOCOM)*, Rio de Janerio, Brazil, Apr. 2009, pp. 1629–1637.

[89] G. Owen, "Existence of equilibrium pairs in continuous games," *Int. J. Game Theory*, vol. 5, no. 2-3, pp. 97–105, June 1976.

[90] S. Boyd and L. Vandenberghe, *Convex Optimization*, 1st ed. New York, USA: Cambridge University Press, 2004.

[91] F. A. Dietrich, *Robust Signal Processing for Wireless Communications*. Heidelberg, Germany: Springer, 2007.

[92] H. Ltkepohl, *Handbook of Matrices*, 1st ed. Chichester, England: John Wiley & Sons, 1997.

[93] A. Beck, A. Ben-Tal, and L. Tetruashvili, "A sequential parametric convex approximation method with applications to nonconvex truss topology design problems," *J. Global Optimization*, vol. 47, no. 1, pp. 29–51, May 2010.

[94] J. Dattorro, *Convex Optimization & Euclidean Distance Geometry*, 3rd ed. Meboo Publishing, 2008.

~

Appendix A

Proofs for Chapter 3

A.1 Proof of Lemma 3.2

Lemma 3.2 is proved in two steps, i.e., Steps A and B. In Step A, we prove that $\sum_l \hat{R}_{\mathrm{r}l}(\lambda_l')$ can be increased by modifying the current power allocation on two specific subchannels. In Step B, we show that $\sum_l \hat{R}_{\mathrm{r}l}(\lambda_l')$ may be further increased.

Step A: $\sum_l \hat{R}_{\mathrm{r}l}(\lambda_l')$ can be increased. Given the fact that $\sum_l \mathrm{Tr}\{\mathbf{P}_{\mathrm{r}l}(\lambda_l)\} = \sum_l \mathrm{Tr}\{\mathbf{P}_{\mathrm{r}l}(\lambda_l')\}$, it can be shown that $1/\lambda_i' > \min_k\{1/\alpha_i(k)\}$ as long as $1/\lambda_j > \min_k\{1/\alpha_j(k)\}$. As a result, there exist k_1 and k_2 such that $1/\lambda_i' > 1/\alpha_i(k_1)$ and $1/\lambda_j > 1/\alpha_j(k_2)$. Define $f(p_{\mathrm{r}i}(k_1)) \triangleq \log 1 + \alpha_i(k_1)p_{\mathrm{r}i}(k_1) + \log 1 + \alpha_j(k_2)p_{\mathrm{r}j}(k_2)$ where $p_{\mathrm{r}j}(k_2) = p - p_{\mathrm{r}i}(k_1)$ and p is a positive constant. It can be seen that $f(p_{\mathrm{r}i}(k_1))$ is strictly concave in $p_{\mathrm{r}i}(k_1) \in [0, p], \forall p > 0$. Set $p = \left(1/\lambda_j^{'} - 1/\alpha_j(k_2)\right)^+ + 1/\lambda_i' - 1/\alpha_i(k_1)$. The optimal allocation of the power p on $\alpha_i(k_1)$ and $\alpha_j(k_2)$ that maximizes $f(p_{\mathrm{r}i}(k_1))$ is $p_{\mathrm{r}i}(k_1) = \left(1/\lambda^{\mathrm{opt}}(p) - 1/\alpha_i(k_1)\right)^+$ and $p_{\mathrm{r}j}(k_2) = \left(1/\lambda^{\mathrm{opt}}(p) - 1/\alpha_j(k_2)\right)^+$ where $\lambda^{\mathrm{opt}}(p)$ is a function of p and $1/\lambda^{\mathrm{opt}}(p)$ is the optimal water level. It can be shown that $1/\lambda^{\mathrm{opt}}(p) < 1/\lambda_i'$. There exist two cases, i.e., $1/\lambda^{\mathrm{opt}}(p) \leq 1/\lambda_i$ and $1/\lambda^{\mathrm{opt}}(p) > 1/\lambda_i$. In the case when $1/\lambda^{\mathrm{opt}}(p) \leq 1/\lambda_i$, it follows that $\left(1/\lambda^{\mathrm{opt}}(p) - 1/\alpha_i(k_1)\right)^+ \leq \left(1/\lambda_i - 1/\alpha_i(k_1)\right)^+ < 1/\lambda_i' - 1/\alpha_i(k_1)$. The power allocation on k_1 and k_2 using λ_i' and λ_j' is

$$p_{\mathrm{r}i}(k_1) = \left(\frac{1}{\lambda_i'} - \frac{1}{\alpha_i(k_1)}\right)^+ \tag{A.1a}$$

$$p_{\mathrm{r}j}(k_2) = \left(\frac{1}{\lambda_j'} - \frac{1}{\alpha_j(k_2)}\right)^+. \tag{A.1b}$$

Since $f(p_{\mathrm{r}i}(k_1))$ is strictly concave as mentioned above, it can be seen that the power allocation

$$p_{\mathrm{r}i}(k_1) = \left(\frac{1}{\lambda_i} - \frac{1}{\alpha_i(k_1)}\right)^+ \tag{A.2a}$$

$$p_{\mathrm{r}j}(k_2) = \left(\frac{1}{\lambda'_j} - \frac{1}{\alpha_j(k_2)}\right)^+ + \frac{1}{\lambda'_i} - \frac{1}{\alpha_i(k_1)} - \left(\frac{1}{\lambda_i} - \frac{1}{\alpha_i(k_1)}\right)^+ \tag{A.2b}$$

which reduces $p_{\mathrm{r}i}(k_1)$ and increases $p_{\mathrm{r}j}(k_2)$, both by $1/\lambda'_i - 1/\alpha_i(k_1) - \left(1/\lambda_i - 1/\alpha_i(k_1)\right)^+$, yields higher $f(p_{\mathrm{r}i}(k_1))$ than the power allocation in (A.1).

Therefore, the sum-rate $\sum_l \sum_k \log 1 + \alpha_l(k) p_{\mathrm{r}l}(k)$ achieved using (A.2) and

$$p_{\mathrm{r}i}(k) = \left(\frac{1}{\lambda'_i} - \frac{1}{\alpha_i(k)}\right)^+, \forall k \in \mathcal{I}_i \setminus \{k_1\} \tag{A.3a}$$

$$p_{\mathrm{r}j}(k) = \left(\frac{1}{\lambda'_j} - \frac{1}{\alpha_j(k)}\right)^+, \forall k \in \mathcal{I}_j \setminus \{k_2\} \tag{A.3b}$$

is larger than $\sum_l \hat{R}_{\mathrm{r}l}(\lambda'_l)$. This is the first step of increasing sum-rate. Moreover, it can be seen that there exists $\tilde{\lambda}_j$ such that

$$\frac{1}{\lambda'_j} < \frac{1}{\tilde{\lambda}_j} < \frac{1}{\lambda_j} \tag{A.4a}$$

$$\mathrm{Tr}\{\mathbf{P}_{\mathrm{r}i}(\lambda'_i)\} - \left(\frac{1}{\lambda'_i} - \frac{1}{\alpha_i(k_1)}\right)^+ + \left(\frac{1}{\lambda_i} - \frac{1}{\alpha_i(k_1)}\right)^+ + \mathrm{Tr}\{\mathbf{P}_{\mathrm{r}j}(\tilde{\lambda}_j)\} = \sum_l \mathrm{Tr}\{\mathbf{P}_{\mathrm{r}l}(\lambda'_l)\} \tag{A.4b}$$

and the power allocation

$$p_{\mathrm{r}i}(k_1) = \left(\frac{1}{\lambda_i} - \frac{1}{\alpha_i(k)}\right)^+ \tag{A.5a}$$

$$p_{\mathrm{r}i}(k) = \left(\frac{1}{\lambda'_i} - \frac{1}{\alpha_i(k)}\right)^+, \forall k \in \mathcal{I}_i \setminus \{k_1\} \tag{A.5b}$$

$$p_{\mathrm{r}j}(k) = \left(\frac{1}{\tilde{\lambda}_j} - \frac{1}{\alpha_j(k)}\right)^+, \forall k \in \mathcal{I}_j \tag{A.5c}$$

which spreads the power $1/\lambda'_i - 1/\alpha_i(k_1) - \left(1/\lambda_i - 1/\alpha_i(k_1)\right)^+$ over $\alpha_j(k)$'s, $\forall k \in \mathcal{I}_j$, achieves even higher sum-rate than that achieved by the power allocation specified by (A.2) and (A.3). This is the second step of increasing the sum-rate.

For the second case in which $1/\lambda_i < 1/\lambda^{\rm opt}(p) < 1/\lambda_i'$, the following process is adopted. Similar to the two steps of increasing the sum-rate in the first case, the sum-rate $\sum_l \sum_k \log 1 + \alpha_l(k) p_{{\rm r}l}(k)$ increases after each of the following two adjustments of power allocation. First, reduce $p_{{\rm r}i}(k_1)$ from $1/\lambda_i' - 1/\alpha_i(k_1)$ to $\left(1/\lambda^{\rm opt}(p) - 1/\alpha_i(k_1)\right)^+$. Then, spread the reduced power $1/\lambda_i' - 1/\alpha_i(k_1) - \left(1/\lambda^{\rm opt}(p) - 1/\alpha_i(k_1)\right)^+$ over $\alpha_j(k)$'s , $k \in \mathcal{I}_j$ by finding and using $1/\tilde{\lambda}_j'$ which satisfies

$$\begin{aligned} {\rm Tr}\{\mathbf{P}_{{\rm r}i}(\lambda_i')\} - \left(\tfrac{1}{\lambda_i'} - \tfrac{1}{\alpha_i(k_1)}\right)^+ + \left(\tfrac{1}{\lambda^{\rm opt}(p)} - \tfrac{1}{\alpha_i(k_1)}\right)^+ \\ + {\rm Tr}\{\mathbf{P}_{{\rm r}j}(\tilde{\lambda}_j')\} = \sum_l {\rm Tr}\{\mathbf{P}_{{\rm r}l}(\lambda_l')\}. \end{aligned} \tag{A.6}$$

After the adjustments, it is straightforward to see that the total power allocated on k_1 and k_2 is reduced from $p = \left(1/\lambda_j' - 1/\alpha_j(k_2)\right)^+ + 1/\lambda_i' - 1/\alpha_i(k_1)$ to $\bar{p} = \left(1/\tilde{\lambda}_j' - 1/\alpha_j(k_2)\right)^+ + \left(1/\lambda^{\rm opt}(p) - 1/\alpha_i(k_1)\right)^+$. In consequence, there exists a new optimal water level $1/\lambda^{\rm opt}(\bar{p})$ based on which the optimal allocation of the power $\bar{p}$, i.e., $p_{{\rm r}i}(k_1) = \left(1/\lambda^{\rm opt}(\bar{p}) - 1/\alpha_i(k_1)\right)^+$ and $p_{{\rm r}j}(k_2) = 1/\lambda^{\rm opt}(\bar{p}) - 1/\alpha_j(k_2)$, maximizes $f(p_{{\rm r}i}(k_1))$ when p in $f(p_{{\rm r}i}(k_1))$ is substituted by $\bar{p}$. Since $\bar{p} < p$, it can be seen that $1/\lambda^{\rm opt}(\bar{p}) < 1/\lambda^{\rm opt}(p)$. Update p and $1/\lambda^{\rm opt}(p)$ so that $p = \bar{p}$ and $1/\lambda^{\rm opt}(p) = 1/\lambda^{\rm opt}(\bar{p})$. Then the above process of reducing $p_{{\rm r}i}(k_1)$ to $\left(1/\lambda^{\rm opt}(p) - 1/\alpha_i(k_1)\right)^+$, finding the new $1/\tilde{\lambda}_j'$ and the new $1/\lambda^{\rm opt}(p)$ can be repeated until (a) $1/\lambda^{\rm opt}(p) \leq 1/\lambda_i$ or until (b) $1/\lambda^{\rm opt}(p) \leq 1/\alpha_i(k_1)$. The former matches the condition for the first case discussed in the previous paragraph and therefore can be dealt with in the same way as in the first case, which leads to (A.5). The latter implies that $1/\lambda_i < 1/\lambda^{\rm opt}(p) \leq 1/\alpha_i(k_1)$, in which case the power allocation can also be equivalently written as (A.5). Note that during this process the sum-rate $\sum_l \sum_k \log 1 + \alpha_l(k) p_{{\rm r}l}(k)$ increases. Therefore, summarizing the above two cases of $1/\lambda^{\rm opt}(p) \leq 1/\lambda_i$ and $1/\lambda^{\rm opt}(p) > 1/\lambda_i$, it is proved that the sum-rate can be increased by reducing $p_{{\rm r}i}(k_1)$ from $1/\lambda_i' - 1/\alpha_i(k_1)$ to $\left(1/\lambda_i - 1/\alpha_i(k_1)\right)^+$ and using the power allocation in (A.5).

Step B: $\sum_l \hat{R}_{{\rm r}l}(\lambda_l')$ may be further increased. Keep the above selected k_2 unchanged. As long as there exists k such that $p_{{\rm r}i}(k) = \left(1/\lambda_i' - 1/\alpha_i(k_1)\right)^+$ and $p_{{\rm r}i}(k) > 0$, this k can be selected as k_1 and the procedure of reducing $p_{{\rm r}i}(k_1)$

from $1/\lambda_i' - 1/\alpha_i(k_1)$ to $\big(1/\lambda_i - 1/\alpha_i(k_1)\big)^+$ and spreading the reduced power over $\alpha_j(k)$'s, $\forall k \in \mathcal{I}_j$ as specified in (A.5) can be performed. This process can be repeated until $p_{\mathrm{r}i}(k) = \big(1/\lambda_i - 1/\alpha_i(k)\big)^+, \forall k \in \{q \in \mathcal{I}_i | \big(1/\lambda_i' - 1/\alpha_i(q)\big)^+ > 0\}$ and $p_{\mathrm{r}i}(k) = 0, \forall k \in \{q \in \mathcal{I}_i | \big(1/\lambda_i' - 1/\alpha_i(q)\big)^+ = 0\}$. Note that the sum-rate $\sum_l \sum_k \log 1 + \alpha_l(k) p_{\mathrm{r}l}(k)$ increases in the above process for every qualifying k_1. The resulting power allocation on $\alpha_i(k)$'s, $\forall k \in \mathcal{I}_i$ is equivalent to $p_{\mathrm{r}i}(k) = \big(1/\lambda_i - 1/\alpha_i(k)\big)^+, \forall k \in \mathcal{I}_i$ since $\big(1/\lambda_i - 1/\alpha_i(k)\big)^+ = 0$ if $\big(1/\lambda_i' - 1/\alpha_i(k)\big)^+ = 0$. From the procedure described in the previous paragraphes, the resulting power allocation on $\alpha_j(k)$'s, $\forall k \in \mathcal{I}_j$ is $p_{\mathrm{r}j}(k) = \big(1/\tilde{\lambda}_j - 1/\alpha_j(k)\big)^+, \forall k$. According to the power constraint $\sum_l \mathrm{Tr}\{\mathbf{P}_{\mathrm{r}l}(\lambda_l)\} = \sum_l \mathrm{Tr}\{\mathbf{P}_{\mathrm{r}l}(\lambda_l')\}$ and the fact that the total power consumption is fixed at all time, it can be seen that $1/\tilde{\lambda}_j = 1/\lambda_j$.

Summarizing the above two steps, Lemma 3.2 is proved.

A.2 Proof of Lemma 3.3

Given that $\lambda_i' \leq \lambda_j$, we have $\lambda_i < \lambda_i' \leq \lambda_j < \lambda_j'$. According to Lemma 3.2, there exists $\tilde{\lambda}_i < \lambda_i'$ such that

$$\begin{aligned} &\mathrm{Tr}\{\mathbf{P}_{\mathrm{r}i}(\lambda_i')\} + \mathrm{Tr}\{\mathbf{P}_{\mathrm{r}j}(\lambda_j)\} \\ &\qquad = \mathrm{Tr}\{\mathbf{P}_{\mathrm{r}i}(\tilde{\lambda}_i)\} + \mathrm{Tr}\{\mathbf{P}_{\mathrm{r}j}(\lambda_j')\} \end{aligned} \tag{A.7}$$

and

$$\hat{R}_{\mathrm{r}i}(\lambda_i') + \hat{R}_{\mathrm{r}j}(\lambda_j) > \hat{R}_{\mathrm{r}i}(\tilde{\lambda}_i) + \hat{R}_{\mathrm{r}j}(\lambda_j'). \tag{A.8}$$

Therefore, given that

$$\hat{R}_{\mathrm{r}i}(\lambda_i') + \hat{R}_{\mathrm{r}j}(\lambda_j) = \hat{R}_{\mathrm{r}i}(\lambda_i) + \hat{R}_{\mathrm{r}j}(\lambda_j') \tag{A.9}$$

it is necessary that $\tilde{\lambda}_i > \lambda_i$. As a result, it leads to

$$\begin{aligned} &\mathrm{Tr}\{\mathbf{P}_{\mathrm{r}i}(\lambda_i')\} + \mathrm{Tr}\{\mathbf{P}_{\mathrm{r}j}(\lambda_j)\} \\ &\qquad < \mathrm{Tr}\{\mathbf{P}_{\mathrm{r}i}(\lambda_i)\} + \mathrm{Tr}\{\mathbf{P}_{\mathrm{r}j}(\lambda_j')\}. \end{aligned} \tag{A.10}$$

Lemma 3.3 is thereby proved.

A.3 Proof of Theorem 3.1

First we prove that the optimal water-levels must satisfy the condition (3.18a). It can be seen that the maximum $R^{\text{tw}}(\mathbf{B}, \mathbf{D})$ is achieved with minimum power consumption using $\lambda_1 = \lambda_2 = \max\{\lambda^0, \mu^0_{\text{ma}}\}$ when $\min\{1/\mu^0_1, 1/\mu^0_2\} \geq 1/\mu_{\text{ma}}$ at the optimality. Therefore, it is necessary that $\min\{1/\mu^0_1, 1/\mu^0_2\} < 1/\mu^0_{\text{ma}}$ given that $\lambda_1 \neq \lambda_2$ at optimality. Let us consider the case when $\min\{1/\lambda_1, 1/\lambda_2\} = 1/\lambda_1 < 1/\lambda_2$ at optimality. According to the constraint (3.15a), we have that $1/\lambda_1 \leq 1/\mu^0_2$ at optimality. Similarly, it can be seen that $1/\lambda_2 \leq 1/\mu^0_1$ at optimality. Since $1/\lambda_1 < 1/\lambda_2$, it leads to the result that $1/\lambda_1 \leq 1/\mu^0_2 < 1/\mu^0_1$ at optimality. Assuming that $\min\{1/\mu^0_1, 1/\mu^0_2\} \neq 1/\lambda_1$ at optimality when $\lambda_1 \neq \lambda_2$, it infers that $1/\lambda_1 < 1/\mu^0_2 < 1/\lambda_2$. However, it can be seen that the power allocation using $1/\lambda_1 < 1/\mu^0_2 < 1/\lambda_2$ does not provide the maximum achievable $R^{\text{tw}}(\mathbf{B}, \mathbf{D})$ according to Lemma 3.2. Consequently, the resulting power allocation is not optimal. It contradicts the assumption that $\min\{1/\mu^0_1, 1/\mu^0_2\} \neq 1/\lambda_1$ at optimality. Thus, the above assumption is invalid and it is necessary that $\min\{1/\mu^0_1, 1/\mu^0_2\} = 1/\lambda_1$ at optimality when $\lambda_1 \neq \lambda_2$. Similarly, it can be proved that $\min\{1/\mu^0_1, 1/\mu^0_2\} = 1/\lambda_2$ at optimality when $\lambda_1 \neq \lambda_2$ for the case when $\min\{1/\lambda_1, 1/\lambda_2\} = 1/\lambda_2 < 1/\lambda_1$. Therefore, it always holds true that $\min\{1/\lambda_1, 1/\lambda_2\} = \min\{1/\mu^0_1, 1/\mu^0_2\}$ if $\lambda_1 \neq \lambda_2$.

Next we prove that the optimal water-levels must satisfy condition (3.18b). It is straightforward to see that $1/\lambda_1 = 1/\lambda_2 \leq 1/\lambda^0$. Moreover, according to the constraints (3.15a) and (3.15b), it is not difficult to see that $1/\lambda_1 = 1/\lambda_2 \leq \min\{1/\mu^0_1, 1/\mu^0_2, 1/\mu^0_{\text{ma}}\}$ when $1/\lambda_1 = 1/\lambda_2$ at optimality. Indeed, if $1/\lambda_1 = 1/\lambda_2 > 1/\mu^0_{\text{ma}}$, then (3.15b) cannot be satisfied. If $1/\lambda_1 = 1/\lambda_2 > \min\{1/\mu^0_1, 1/\mu^0_2\}$, then (3.15a) cannot be satisfied. Combining the above two facts, we have $1/\lambda_1 = 1/\lambda_2 \leq \min\{1/\mu^0_1, 1/\mu^0_2, 1/\mu^0_{\text{ma}}, 1/\lambda^0\}$ when $1/\lambda_1 = 1/\lambda_2$ at optimality. For the case that $\min\{1/\mu^0_1, 1/\mu^0_2\} \geq 1/\mu^0_{\text{ma}}$, the above constraint can be written as $1/\lambda_1 = 1/\lambda_2 \leq \min\{1/\mu^0_{\text{ma}}, 1/\lambda^0\}$. For this case, it is straightforward to see that the achieved sum-rate is not maximized if $1/\lambda_1 = 1/\lambda_2 < \min\{1/\mu^0_{\text{ma}}, 1/\lambda^0\}$. Therefore, the optimal water-levels must satisfy the condition (3.18b) when $\min\{1/\mu^0_1,$

$1/\mu_2^0\} \geq 1/\mu_{\rm ma}^0$ given that $1/\lambda_1 = 1/\lambda_2$. For the case when $\min\{1/\mu_1^0, 1/\mu_2^0\} < 1/\mu_{\rm ma}^0$, it can be seen that $1/\lambda^0 \leq \min\{1/\mu_1^0, 1/\mu_2^0\}$ given that $1/\lambda_1 = 1/\lambda_2$ at optimality. Otherwise, it can be shown that either of the following two results must occur. If $1/\lambda^0 > \min\{1/\mu_1^0, 1/\mu_2^0\}$ and $1/\lambda_1 = 1/\lambda_2 \leq \min\{1/\mu_1^0, 1/\mu_2^0\}$, then the sum-rate can be increased. If $1/\lambda^0 > \min\{1/\mu_1^0, 1/\mu_2^0\}$ and $1/\lambda_1 = 1/\lambda_2 \geq \min\{1/\mu_1^0, 1/\mu_2^0\}$, then the constraint (3.15a) cannot be satisfied. Therefore, given that $1/\lambda^0 \leq \min\{1/\mu_1^0, 1/\mu_2^0\}$ for the case when $\min\{1/\mu_1^0, 1/\mu_2^0\} < 1/\mu_{\rm ma}^0$ and $1/\lambda_1 = 1/\lambda_2$ at optimality, we have $1/\lambda^0 \leq \min\{1/\mu_1^0, 1/\mu_2^0\} < 1/\mu_{\rm ma}^0$. Consequently, the constraint $1/\lambda_1 = 1/\lambda_2 \leq \min\{1/\mu_1^0, 1/\mu_2^0, 1/\mu_{\rm ma}^0, 1/\lambda^0\}$ can be rewritten as $1/\lambda_1 = 1/\lambda_2 \leq 1/\lambda^0 = \min\{1/\mu_{\rm ma}^0, 1/\lambda^0\}$. It is straightforward to see for this case that $1/\lambda_1 = 1/\lambda_2 < 1/\lambda^0$ does not maximize the sum-rate. Therefore, it can also be concluded that $1/\lambda_1 = 1/\lambda_2 = 1/\lambda^0 = \min\{1/\mu_{\rm ma}^0, 1/\lambda^0\}$ when $\min\{1/\mu_1^0, 1/\mu_2^0\} < 1/\mu_{\rm ma}^0$. Combining the above two cases of $\min\{1/\mu_1^0, 1/\mu_2^0\} \geq 1/\mu_{\rm ma}^0$ and $\min\{1/\mu_1^0, 1/\mu_2^0\} < 1/\mu_{\rm ma}^0$, it can be seen that the optimal water-levels always satisfy the condition (3.18b) given that $1/\lambda_1 = 1/\lambda_2$.

The above two parts complete the proof of Theorem 3.1.

A.4 Proof of Theorem 3.2

The necessity of the constraints (3.15a) and (3.15b) is straightforward. It can be seen that the power consumption can be reduced without reducing the sum-rate $R^{\rm tw}(\mathbf{B}, \mathbf{D})$ when these constraints are not satisfied. The necessity of the constraints (3.18a) and (3.18b) is proved in Theorem 3.1 in Section A.3. Therefore, we next prove the sufficiency of the constraints (3.15a), (3.15b), (3.18a), and (3.18b).

We use proof by contradiction. Assume that the above constrains are not sufficient to determine the optimal $\{\lambda_1, \lambda_2\}$ with minimum power consumption among all $\{\lambda_1, \lambda_2\}$'s that maximize the sum-rate $R^{\rm tw}(\mathbf{B}, \mathbf{D})$. Then there exists $\{\lambda_1^\dagger, \lambda_2^\dagger\}$ satisfying (3.15) and (3.18a)-(3.18b) that maximizes the sum-rate and does not minimize the power consumption. Consequently, at least one of $1/\lambda_1^\dagger$ and $1/\lambda_2^\dagger$ can be reduced without reducing $R^{\rm tw}(\mathbf{B}, \mathbf{D})$. We consider the following two cases. The first case is when $\lambda_1^\dagger \neq \lambda_2^\dagger$ while the second case is when $\lambda_1^\dagger = \lambda_2^\dagger$. In the

first case, $\{\lambda_1^\dagger, \lambda_2^\dagger\}$ satisfies (3.18a) and it is straightforward to see that reducing $\min\{1/\lambda_1^\dagger, 1/\lambda_2^\dagger\}$ is not optimal according to Lemma 3.3. Reducing $\max\{1/\lambda_1^\dagger, 1/\lambda_2^\dagger\}$, on the other hand, necessarily leads to the decrease of $R^{\text{tw}}(\mathbf{B}, \mathbf{D})$ given that (3.15b) is satisfied. Therefore, reducing either of $1/\lambda_1^\dagger$ and $1/\lambda_2^\dagger$ results in the decrease of the sum-rate, which contradicts the previous assumption. In the second case, $\{\lambda_1^\dagger, \lambda_2^\dagger\}$ satisfies (3.18b). According to Theorem 3.2, it is necessary that $1/\lambda_1^\dagger = 1/\lambda_2^\dagger = \min\{1/\mu_{\text{ma}}^0, 1/\lambda^0\}$. From Lemma 3.2, it can be seen that it is not optimal to reduce only one of $1/\lambda_1^\dagger$ and $1/\lambda_2^\dagger$. Reducing both of $1/\lambda_1^\dagger$ and $1/\lambda_2^\dagger$, on the other hand, necessarily leads to the decrease of $R^{\text{tw}}(\mathbf{B}, \mathbf{D})$ given that (3.15b) is satisfied. Therefore, it is impossible that there exists $\{\lambda_1^\dagger, \lambda_2^\dagger\}$ with $\lambda_1^\dagger = \lambda_2^\dagger$, satisfying (3.15) and (3.18b), that maximizes the sum-rate while the resulting power consumption can be reduced. Combining the above two cases, it can be seen that the power consumption cannot be reduced given that the $\{\lambda_1^\dagger, \lambda_2^\dagger\}$ maximizes the sum-rate subject to the relay power limit and satisfies (3.15) and (3.18a)-(3.18b). This contradicts the assumption that the above constrains are not sufficient to determine the optimal $\{\lambda_1, \lambda_2\}$ with minimum power consumption among all $\{\lambda_1, \lambda_2\}$'s that maximize $R^{\text{tw}}(\mathbf{B}, \mathbf{D})$. This completes the proof for Theorem 3.2.

A.5 Proof of Theorem 3.3

The optimality of the pair $\{\lambda_1, \lambda_2\}$ obtained using the algorithm in Table 3.1 is proved in three steps: (A) Steps 2-5 of the algorithm in Table 3.1 find $\{\lambda_1, \lambda_2\}$ that maximizes $R^{\text{bc}}(\mathbf{B}, \mathbf{D}^0)$ with minimum power consumption subject to the constraint in (3.11) and the constraint (3.15a). (B) The pair $\{\lambda_1, \lambda_2\}$ obtained from Steps 2-5 of the algorithm in Table 3.1 needs to be modified to maximize the objective function in (3.11) with minimum power consumption. Step 6 of the algorithm in Table 3.1 deals with two cases in which $\{\lambda_1, \lambda_2\}$ obtained from the previous steps can be simply modified to obtain the optimal pair $\{\lambda_1, \lambda_2\}$. (C) Step 7 of the algorithm in Table 3.1 deals with the remaining case which is more complicated and finds the corresponding optimal pair $\{\lambda_1, \lambda_2\}$ in this case. It is not difficult to see that the constraint in (3.11) is always satisfied in any step of the proposed algorithm.

It can also be seen that Steps 1, 2, and 6 ensure that (3.18b) is satisfied if $\lambda_1 = \lambda_2$ at the output of the algorithm while Steps 3 to 5 ensure that (3.18a) is satisfied if $\lambda_1 \neq \lambda_2$ at the output. Therefore, in the following we only consider the constraints (3.15a) and (3.15b), which are equivalent to the constraints in (3.12).

(A) Steps 2-5 find the pair $\{\lambda_1, \lambda_2\}$ that maximizes $R(\mathbf{B}, \mathbf{D}^0)$ with minimum power consumption subject to the constraint (3.15a). Note that the maximum $R(\mathbf{B}, \mathbf{D}^0)$ with minimum power consumption is achieved by $\hat{R}_{\mathrm{r}1}(\lambda_1) + \hat{R}_{\mathrm{r}2}(\lambda_2)$ for some specific $\{\lambda_1, \lambda_2\}$ if (3.15a) is satisfied. Therefore, it is equivalent to finding the $\{\lambda_1, \lambda_2\}$ that maximizes $\hat{R}_{\mathrm{r}1}(\lambda_1) + \hat{R}_{\mathrm{r}2}(\lambda_2)$ subject to (3.15a). The initial power allocation in Step 1 of the algorithm in Table 3.1 using $1/\lambda_1 = 1/\lambda_2 = 1/\lambda^0$ maximizes $\hat{R}_{\mathrm{r}1}(\lambda_1) + \hat{R}_{\mathrm{r}2}(\lambda_2)$. Regarding the constraint (3.15a), the following cases are possible.

(A-1) $\lambda_i \geq \mu_j^0, \forall i$. In this case, the constraint (3.15a) is satisfied and $\{\lambda^0, \lambda^0\}$ is the desired $\{\lambda_1, \lambda_2\}$.

(A-2) $\lambda_i < \mu_j^0$ and $\lambda_j \geq \mu_i^0$. In this case, the constraint (3.15a) is not satisfied for i. The relay power consumption can be reduced without decreasing $R(\mathbf{B}, \mathbf{D}^0)$ by increasing λ_i until $\lambda_i = \mu_j^0$. Then, $R(\mathbf{B}, \mathbf{D}^0)$ can be increased by decreasing λ_j until the relay power limit is reached or until $\lambda_j = \mu_i^0$.

(A-3) $\lambda_i < \mu_j^0, \forall i$. In this case, it is straightforward to see that the pair $\{\lambda_1, \lambda_2\}$ that maximizes $R(\mathbf{B}, \mathbf{D}^0)$ with minimum power consumption subject to the constraint (3.15a) satisfies $\lambda_i = \mu_j^0, \forall i$.

The above three cases are determined in Step 2. Case A-1 is dealt with in Step 2 of the algorithm in Table 3.1. Case A-2 is dealt with in Steps 3 and 4. Case A-3 is dealt with in Steps 3 and 5.

(B) Steps 6 and 7 of the algorithm in Table 3.1 find the optimal pair $\{\lambda_1, \lambda_2\}$ that maximizes the objective function in (3.11) with minimum power consumption. Since $R^{\mathrm{ma}}(\mathbf{D}^0) < \bar{R}_{1\mathrm{r}}(\mathbf{D}_1^0) + \bar{R}_{2\mathrm{r}}(\mathbf{D}_2^0)$, it can be seen that $\lambda_i, \forall i$ should either increase or remain the same in order to satisfy the constraint (3.15b) given that the constraint (3.15a) is satisfied. Therefore, the optimal power allocation can be derived by increasing λ_1 and/or λ_2, if necessary, based on the power allocation derived from Steps 1-5. Regarding the constraint (3.15b), the following cases are

possible.

(B-1) $\lambda_i \geq \mu_{\rm ma}^0, \forall i$ or $\Big(\lambda_i \geq \mu_{\rm ma}^0$, $\lambda_j < \mu_{\rm ma}^0$ and $\hat{R}_{\rm r1}(\lambda_1) + \hat{R}_{\rm r2}(\lambda_2) \leq R^{\rm ma}(\mathbf{D}^0)\Big)$. In this case, the constraint (3.15b) is satisfied and the current $\{\lambda_1, \lambda_2\}$ is optimal.

(B-2) $\lambda_i < \mu_{\rm ma}^0, \forall i$ and $\hat{R}_{\rm r1}(\lambda_1) + \hat{R}_{\rm r2}(\lambda_2) > R^{\rm ma}(\mathbf{D}^0)$. In this case, it is not difficult to see that it is optimal to simply set $\lambda_i = \mu_{\rm ma}^0, \forall i$.

(B-3) $\lambda_i > \mu_{\rm ma}^0$, $\lambda_j < \mu_{\rm ma}^0$ and $\hat{R}_{\rm r1}(\lambda_1) + \hat{R}_{\rm r2}(\lambda_2) > R^{\rm ma}(\mathbf{D}^0)$.

Subcases B-1 and B-2 are simple and dealt with in Step 6 of the algorithm in Table 3.1. It can be shown that in these two cases the constraints (3.15a) and (3.15b) are both necessary and sufficient for finding the optimal power allocation in terms of maximizing the sum-rate with minimum power consumption. Subcase B-3 is dealt with in Step 7. The optimal strategy in Subcase B-3, as in Step 7 of the algorithm in Table 3.1, is to increase λ_j while keeping λ_i unchanged until $\hat{R}_{\rm r1}(\lambda_1) + \hat{R}_{\rm r2}(\lambda_2) = R^{\rm ma}(\mathbf{D}^0)$. In order to prove that this strategy is optimal, the following three points are necessary and sufficient.

1. It is optimal to increase $\min_i\{\lambda_i\}$.

2. $\lambda_i = \mu_j^0$ if $\lambda_i > \mu_{\rm ma}^0$ and $\lambda_j < \mu_{\rm ma}^0$.

3. At optimality, the increased λ_j, denoted as λ_j', satisfies $\lambda_j < \lambda_j' < \mu_{\rm ma}^0$.

The first point states that it is optimal to increase λ_j as long as $\lambda_j < \lambda_i$. The second point infers that it is not optimal to decrease λ_i. The third point infers that λ_j' is always larger than λ_i and therefore it is not optimal to increase λ_i at any time. The first point follows from Lemma 3.3. For the second point, assume that $\lambda_i > \mu_j^0$. It follows that $P_{\rm r}^{\max}$ is used up, i.e., $P_{\rm r}^{\max} = \sum_l \sum_k \big(1/\lambda_l - 1/\alpha_l(k)\big)^+$. Otherwise, the equality in the constraint (3.15a) is not achieved for i and the objective function in (3.11) can be increased by decreasing λ_i, which contradicts Steps 1-5 of the algorithm in Table 3.1. Given that $\lambda_i > \mu_j^0$ and $P_{\rm r}^{\max} = \sum_l \sum_k \big(1/\lambda_l - 1/\alpha_l(k)\big)^+$, it can be proved that $1/\lambda_i \geq 1/\lambda_j$. Otherwise, the power allocation can be proved not optimal based on Lemma 3.2 because the objective function in (3.11) is not maximized subject to the constraint (3.15a), which contradicts Steps 1-5 of the algorithm in Table 3.1. However, the conclusion that $1/\lambda_i \geq 1/\lambda_j$ contradicts

Subcase B-3 in which $\lambda_i > \mu^0_{\rm ma}, \lambda_j < \mu^0_{\rm ma}$. Thus, the assumption that $\lambda_i > \mu^0_j$ is invalid. Since $\lambda_i \geq \mu^0_j$ at the output of Steps 1-5 of the algorithm in Table 3.1, we have $\lambda_i = \mu^0_j$. For the third point, assume that $\lambda'_j > \mu^0_{\rm ma}$. Then it follows that $\hat{R}_{\rm r1}(\lambda_1) + \hat{R}_{\rm r2}(\lambda_2) < R^{\rm ma}(\mathbf{D}^0)$, which is not optimal. Therefore, $\lambda'_j < \mu^0_{\rm ma}$ at optimality of Subcase B-3.

(C) Finally, we prove that λ'_j found in Step 7 of the algorithm in Table 3.1 for Subcase B-3 is optimal. The optimal λ'_j for Case B-3 is the solution to the following optimization problem

$$\mathbf{min} \quad \frac{1}{\lambda'_j} \tag{A.11a}$$

$$\mathbf{s.t.} \quad \hat{R}_{\rm ri}(\lambda_i) + \hat{R}_{\rm rj}(\lambda'_j) = R^{\rm ma}(\mathbf{D}^0). \tag{A.11b}$$

Using the definition that $p_{\rm ri}(k) = \big(1/\lambda_i - 1/\alpha_i(k)\big)^+$ and $\mathcal{M}^+_{\rm ri} = \{k | p_{\rm ri}(k) > 0\}$, the constraint in (A.11) is equal to

$$\hat{R}_{\rm ri}(\lambda_i) + \sum_{k \in \mathcal{M}^+_{\rm rj}} \log \frac{\alpha_j(k)}{\lambda'_j} = R^{\rm ma}(\mathbf{D}^0). \tag{A.12}$$

As previously proved, $\lambda_i = \mu^0_j$ in Case B-3, which means that $\hat{R}_{\rm ri}(\lambda_i) = \bar{R}_{j\rm r}(\mathbf{D}^0_j)$. Thus, the above equation can be written as

$$\sum_{k \in \mathcal{M}^+_{\rm rj}} \log \frac{\alpha_j(k)}{\lambda'_j} = R^{\rm ma}(\mathbf{D}^0) - \bar{R}_{j\rm r}(\mathbf{D}^0_j). \tag{A.13}$$

Therefore, the optimal λ'_j satisfies

$$|\mathcal{M}^+_{\rm rj}| \log \lambda'_j = \sum_{k \in \mathcal{M}^+_{\rm rj}} \log \alpha_j(k) - R^{\rm ma}(\mathbf{D}^0) + \bar{R}_{j\rm r}(\mathbf{D}^0_j) \tag{A.14}$$

and the optimality of the water level λ'_j found in Step 7 of the algorithm in Table 3.1 is proved.

The proof of Theorem 3.3 is thereby complete.

$\sim$

Appendix B

Proofs for Chapter 4

B.1 Proof of Lemma 4.1

Proof for claim 1: Given $\widetilde{\mathbf{D}}$ as defined in the lemma, it follows that $R^{\mathrm{ma}}(\widetilde{\mathbf{D}}) = \bar{R}_{j\mathrm{r}}(\widetilde{\mathbf{D}}_j)$. From the definitions (4.2a)-(4.2c), it can be seen that $R^{\mathrm{ma}}(\widetilde{\mathbf{D}}) > \bar{R}_{j\mathrm{r}}(\widetilde{\mathbf{D}}_j) = \bar{R}_{j\mathrm{r}}(\mathbf{D}_j)$ if $1/\mu_{\mathrm{ma}}(\widetilde{\mathbf{D}}) > 1/\mu_j$. Therefore, it is necessary that $1/\mu_{\mathrm{ma}}(\widetilde{\mathbf{D}}) \leq 1/\mu_j$.

Proof for claim 2: First, note that $R^{\mathrm{ma}}(\widehat{\mathbf{D}})$ is a continuous and strictly increasing function of t in $[0, 1]$. Second, based on the definition (4.2c), it follows that $R^{\mathrm{ma}}(\widehat{\mathbf{D}})$ is a strictly increasing function of $1/\mu_{\mathrm{ma}}(\widehat{\mathbf{D}})$ when $1/\mu_{\mathrm{ma}}(\widehat{\mathbf{D}}) > \min\left(\{\right) 1/\alpha_i(k), \forall i, \forall k\}$, or equivalently, $R^{\mathrm{ma}}(\widehat{\mathbf{D}}) > 0$. Since $\mathrm{Tr}\{\mathbf{D}_1\} > 0$ and $\mathrm{Tr}\{\mathbf{D}_2\} > 0$, we have $R^{\mathrm{ma}}(\widehat{\mathbf{D}}) > 0$ for any $t \in [0, 1]$. Thus, given the fact that $1/\mu_{\mathrm{ma}}(\widetilde{\mathbf{D}}) \leq 1/\mu_j$ when $t = 0$ and that $1/\mu_{\mathrm{ma}}(\widetilde{\mathbf{D}}) = 1/\mu_{\mathrm{ma}}(\mathbf{D}) > 1/\mu_j$ when $t = 1$, it can be seen that there exists $\widehat{t} \in [0, 1)$ such that $1/\mu_{\mathrm{ma}}(\widehat{\mathbf{D}}) = 1/\mu_j$ when $t = \widehat{t}$. Using Lemma 3.1 in Chapter 3, i.e., $1/\mu_{\mathrm{ma}} < \max\left(\{\right) 1/\mu_1, 1/\mu_2\}$, it can be seen that $1/\mu_i(\widehat{\mathbf{D}}_i) > 1/\mu_{\mathrm{ma}}(\widehat{\mathbf{D}}) = 1/\mu_j$ when $t = \widehat{t}$.

B.2 Proof of Lemma 4.2

First we prove $\lambda_j = \mu_i > \mu_{\mathrm{ma}}$ if $\lambda_i < \lambda_j$. Using Lemma 3.2 in Chapter 3, it can be seen that $\lambda_i, \forall i$ satisfy $\lambda_1 = \lambda_2$ if $\min\left(\{\right) 1/\mu_i\} \geq 1/\mu_{\mathrm{ma}}$ at optimality. Therefore, we have $\min\left(\{\right) 1/\mu_i\} < 1/\mu_{\mathrm{ma}}$ given that $\lambda_1 \neq \lambda_2$. Using the same lemma and the constraint (4.6a), it can be further concluded that $1/\mu_i < 1/\mu_{\mathrm{ma}}$ at optimality given that $\lambda_i < \lambda_j$. Otherwise, the constraint (4.6b) cannot be satisfied. Therefore, $1/\mu_j > 1/\mu_{\mathrm{ma}}$ according to Lemma 3.1 in Chapter 3. Due to the constraint (4.6a),

we must have $1/\lambda_j \leq 1/\mu_i$ at optimality. Moreover, from Lemma 3.2 in Chapter 3 and the assumption that $\lambda_i < \lambda_j$, it can be seen that $1/\lambda_j < 1/\mu_i$ is not optimal. Therefore, $1/\lambda_j = 1/\mu_i$ if $\lambda_i < \lambda_j$. Following the same approach, we can prove $\lambda_j = \mu_i > \mu_{\rm ma}$ if $\mu_i > \mu_{\rm ma}$.

B.3 Proof of Theorem 4.1

Recall the definitions of μ_1, μ_2, and $\mu_{\rm ma}$ in (4.2a)-(4.2c). Considering the constraints (4.9b)-(4.9d) in the problem (4.9), it can be seen that at optimality we must have $\mu_{\rm ma}^\star \leq \lambda^0$, $\mu_1^\star \leq \lambda^0$, and $\mu_2^\star \leq \lambda^0$. Otherwise, the above mentioned constraints cannot be satisfied. We will prove Theorem 4.1 by contradiction.

Assume that $\mu_{\rm ma}^\star \neq \lambda^0$ at optimality, then $\mu_{\rm ma}^\star < \lambda^0$ according to the above paragraph. Using Lemma 3.1 in Chapter 3, i.e, $1/\mu_{\rm ma} < \max\left(\{\right) 1/\mu_1, 1/\mu_2\}$, and given that $\mu_1^\star \leq \lambda^0$ and $\mu_2^\star \leq \lambda^0$, there are only two possible situations as follows: a) $\max\left(\{\right) 1/\mu_1^\star, 1/\mu_2^\star\} > 1/\mu_{\rm ma}^\star > \min\left(\{\right) 1/\mu_1^\star, 1/\mu_2^\star\} \geq 1/\lambda^0$ and b) $\max\left(\{\right) 1/\mu_1^\star, 1/\mu_2^\star\} \geq \min\left(\{\right) 1/\mu_1^\star, 1/\mu_2^\star\} \geq 1/\mu_{\rm ma}^\star > 1/\lambda^0$. Assume without loss of generality that $\max\left(\{\right) 1/\mu_1^\star, 1/\mu_2^\star\} = 1/\mu_1^\star$ and $\min\left(\{\right) 1/\mu_1^\star, 1/\mu_2^\star\} = 1/\mu_2^\star$. If it is Situation a), then we have $1/\mu_1^\star > 1/\mu_{\rm ma}^\star > 1/\mu_2^\star \geq 1/\lambda^0$. Use Lemma 4.1 with $\widehat{\mathbf{D}}_i = t\mathbf{D}_1^\star$ and $\widehat{\mathbf{D}}_j = \mathbf{D}_2^\star$. As proved in Lemma 4.1, there exists $t \in [0, 1)$ such that $\mu_1(t\mathbf{D}_1^\star) > 1/\mu_{\rm ma}([t\mathbf{D}_1^\star, \mathbf{D}_2^\star]) = 1/\mu_2^\star$. Since $1/\mu_2^\star \geq 1/\lambda^0$, we have $\mu_1(t\mathbf{D}_1^\star) > 1/\mu_{\rm ma}([t\mathbf{D}_1^\star, \mathbf{D}_2^\star]) = 1/\mu_2^\star \geq 1/\lambda^0$, which indicates that $\hat{\mathbf{D}} = [t\mathbf{D}_1^\star, \mathbf{D}_2^\star]$ also satisfies (4.9b)-(4.9e) while $\text{Tr}\{t\mathbf{D}_1^\star\} + \text{Tr}\{\mathbf{D}_2^\star\} < \text{Tr}\{\mathbf{D}_1^\star\} + \text{Tr}\{\mathbf{D}_2^\star\}$. It contradicts the fact that $\mathbf{D}^\star = [\mathbf{D}_1^\star, \mathbf{D}_2^\star]$ is the optimal solution to the problem (4.9). Therefore, Situation a) is impossible. If it is Situation b), there exist two following possible sub-situations: Sub-situation b-1) there exists $i \in \{1, 2\}$ such that $1/\mu_{\rm ma}(\widehat{\mathbf{D}}) = 1/\lambda^0$ and $1/\mu_i(\widehat{\mathbf{D}}_i) \geq 1/\mu_{\rm ma}(\widehat{\mathbf{D}})$ where $\widehat{\mathbf{D}} = [\widehat{\mathbf{D}}_1, \widehat{\mathbf{D}}_2]$ with $\widehat{\mathbf{D}}_i = t_i\mathbf{D}_i^\star$ and $\widehat{\mathbf{D}}_j = \mathbf{D}_j^\star$ for some $t_i \in [0, 1)$ and Sub-situation b-2) there does not exist $t_i \in [0, 1)$ such that $1/\mu_{\rm ma}(\widehat{\mathbf{D}}) = 1/\lambda^0$ and $1/\mu_i(\widehat{\mathbf{D}}_i) \geq 1/\mu_{\rm ma}(\widehat{\mathbf{D}})$ where $\widehat{\mathbf{D}} = [\widehat{\mathbf{D}}_1, \widehat{\mathbf{D}}_2]$ with $\widehat{\mathbf{D}}_i = t_i\mathbf{D}_i^\star$ and $\widehat{\mathbf{D}}_j = \mathbf{D}_j^\star$ for either $i = 1$ or $i = 2$. In Sub-situation b-1), it can be seen that $\widehat{\mathbf{D}}$ satisfies (4.9b)-(4.9e) while $\text{Tr}\{t_i\mathbf{D}_i^\star\} + \text{Tr}\{\mathbf{D}_j^\star\} < \text{Tr}\{\mathbf{D}_1^\star\} + \text{Tr}\{\mathbf{D}_2^\star\}$. It contradicts the fact that $\mathbf{D}^\star = [\mathbf{D}_1^\star, \mathbf{D}_2^\star]$ is the optimal solution to the problem

(4.9). Therefore, Sub-situation b-1) is impossible. If it is Sub-situation b-2), it indicates that with $t_i \in [0,1)$, for either $i=1$ or $i=2$, such that $1/\mu_{\rm ma}(\widehat{\mathbf{D}}) = 1/\lambda^0$, we have $1/\mu_i(\widehat{\mathbf{D}}_i) = 1/\mu_i(t_i\widehat{\mathbf{D}}_i^\star) < 1/\mu_{\rm ma}(\widehat{\mathbf{D}}) = 1/\lambda^0$. As a result, there exists $t_i' \in (t_i, 1)$ such that $1/\mu_i(t_i'\mathbf{D}_i^\star) = 1/\lambda^0$ and $1/\mu_{\rm ma}(\mathbf{D}') > 1/\lambda^0$ where $\mathbf{D}' = [\mathbf{D}_1', \mathbf{D}_2']$ with $\mathbf{D}_i' = t_i'\mathbf{D}_i^\star$ and $\mathbf{D}_j' = \mathbf{D}_j^\star$. Note that $1/\mu_{\rm ma}(\mathbf{D}') > 1/\lambda^0$ because if $1/\mu_i(\mathbf{D}_i') = 1/\lambda^0$ and $1/\mu_{\rm ma}(\mathbf{D}') = 1/\lambda^0$ then we have Sub-situation b-1) instead of Sub-situation b-2). Recalling that $1/\mu_j(\mathbf{D}_j^\star) > 1/\mu_{\rm ma}(\mathbf{D}^\star) > 1/\mu_{\rm ma}(\mathbf{D}')$, we have $1/\mu_j(\mathbf{D}_j^\star) > 1/\mu_{\rm ma}(\mathbf{D}') > 1/\mu_i(\mathbf{D}_i') = 1/\lambda_0$. It indicates that by changing $\mathbf{D}_i^\star$ at optimality to $\mathbf{D}_i' = t_i'\mathbf{D}_i^\star$, and thus, using less power than ${\rm Tr}\{\mathbf{D}_1^\star\} + {\rm Tr}\{\mathbf{D}_2^\star\}$ while satisfying (4.9b)-(4.9e), Sub-situation b-2) changes to Situation a). As it is proved that Situation a) is impossible at optimality, so it is Sub-situation b-2).

Therefore, it is proved that the assumption $\mu_{\rm ma}^\star \neq \lambda^0$ must lead to either of two situations both of which are impossible at optimality. Thus, it is impossible that $\mu_{\rm ma}^\star \neq \lambda^0$. As a result, we must have $\mu_{\rm ma}^\star = \lambda^0$. This completes the proof.

B.4 Proof of Theorem 4.2

Proof of Property 1: First we show that $1/\mu_{\rm ma}^* < 1/\lambda^0$. Since the maximum $\bar{R}_{l\rm r}(\mathbf{D}_l)$, as the objective function of the problem (4.10), cannot achieve $\hat{R}_{\rm r\bar{l}}(\lambda^0)$ in Subcase I-2, it can be seen that $1/\mu_l < 1/\lambda^0$ whenever $1/\mu_{\rm ma} \geq 1/\lambda^0$ and $1/\mu_{\bar{l}} \geq 1/\lambda^0$. As a result, any $\mathbf{D}$ that leads to $1/\mu_{\rm ma} \geq 1/\lambda^0$ is not optimal. The reason is that in such a case the optimal relay power allocation requires $1/\lambda_{\bar{l}} = 1/\mu_l < 1/\lambda^0$ according to Lemma 4.2 and such relay power allocation leads to a BC phase sum-rate $\sum_i \hat{R}_{{\rm r}i}(\lambda_i)$ which is less than $\hat{R}_{\rm r1}(\lambda^0) + \hat{R}_{\rm r2}(\lambda^0)$ according to Lemma 3.2 in Chapter 3. Since $1/\mu_{\rm ma} \geq 1/\lambda^0$ implies that $R^{\rm ma}(\mathbf{D}) \geq \hat{R}_{\rm r1}(\lambda^0) + \hat{R}_{\rm r2}(\lambda^0)$, it can be seen that the constraint (4.5b) is not satisfied and therefore such strategies cannot be optimal. Next we show that $\min_i\{1/\mu_i^*\} < 1/\mu_{\rm ma}^*$. Assuming that $1/\mu_{\rm ma}^* \leq \min_i\{1/\mu_i^*\}$, it leads to $1/\mu_{\rm ma}^* < 1/\lambda^0$ given that the problem (4.9) is infeasible. Moreover, it also leads to the result that $\lambda_i^* = \mu_{\rm ma}^*, \forall i$. However, it is not difficult to see that $R^{\rm ma}(\mathbf{D})$, $\sum_i \hat{R}_{{\rm r}i}(\lambda_i)$ and eventually $R^{\rm tw}(\mathbf{B}, \mathbf{D})$ can be increased in this case through appropriately increasing $1/\mu_{\rm ma}$, which is feasible since

$1/\mu_{\rm ma}^0 > 1/\lambda^0 > 1/\mu_{\rm ma}^*$, and also increasing at least one of $1/\lambda_{\bar{l}}$ and $1/\lambda_l$, which is also feasible since $1/\lambda_{\bar{l}}^* = 1/\mu_{\rm ma}^* < 1/\lambda^0$, given that $1/\mu_{\rm ma}^* \leq \min_i\{1/\mu_i^*\}$ and $1/\mu_{\rm ma}^* < 1/\lambda^0$. It contradicts the assumption that $\mathbf{D}^*$ and $\mathbf{B}^*$ are the optimal solution. Therefore, $1/\mu_{\rm ma}^* > \min_i\{1/\mu_i^*\}$.

Proof of Property 2: Given the fact that $1/\mu_{\rm ma}^* < 1/\lambda^0$, the problem boils down to finding the maximum $1/\mu_{\rm ma}$ such that the corresponding rate $R^{\rm ma}(\mathbf{D})$ can also be achieved by the BC phase sum-rate $\sum_i \hat{R}_{{\rm r}i}(\lambda_i)$ subject to the first constraint in (4.3) and the constraint that $\min_i\{1/\lambda_i\} = \min_i\{1/\mu_i\}$ as stated in Lemma 4.2. Since the maximum $\sum_i \hat{R}_{{\rm r}i}(\lambda_i)$ cannot achieve $R^{\rm ma}(\mathbf{D})$ subject to the above-mentioned constraints as long as $1/\mu_{\rm ma} \geq 1/\lambda^0$, the problem is equivalent to finding the $\lambda_i, \forall i$ to maximize $\sum_i \hat{R}_{{\rm r}i}(\lambda_i)$ such that the resulting $\sum_i \hat{R}_{{\rm r}i}(\lambda_i)$ is achievable by $R^{\rm ma}(\mathbf{D})$ subject to the constraint that $\min_i\{1/\lambda_i\}$ is achievable by $\min_i\{1/\mu_i\}$ (in addition to the power constraints). Consider the problem of maximizing $R^{\rm ma}(\mathbf{D})$ subject to the constraint that $\min_i\{1/\mu_i\} \geq C$ where C is a constant. Note that the maximum of this problem is a non-increasing function of C as long as the problem is feasible and $C < 1/\mu_{\rm ma}$. Recall from Property 1 that $\min_i\{1/\mu_i^*\} < 1/\mu_{\rm ma}^* < 1/\lambda_0$. Assume that the relay does not use full power at optimality, then the maximum achievable $\sum_i \hat{R}_{{\rm r}i}(\lambda_i)$ and the maximum achievable $R^{\rm ma}(\mathbf{D})$ can be both increased subject to all the above constraints by appropriately decreasing $\min_i\{1/\mu_i\}$ (and thereby increasing the maximum achievable $R^{\rm ma}(\mathbf{D})$) while letting the relay decrease $\min_i\{1/\lambda_i\}$ accordingly and at the same time use all the remaining power to increase $\max_i\{1/\lambda_i\}$ (and thereby increasing the maximum achievable $\sum_i \hat{R}_{{\rm r}i}(\lambda_i)$). It contradicts the assumption, which infers that the relay must use full power at optimality.

Proof of Property 3: Define the index $i^- = \arg\min_i\{1/\mu_i\}$. Recall from the proof of Property 1 that $1/\mu_{\rm ma}^* < 1/\lambda^0$. As a result, $R^{\rm ma}(\mathbf{D}^*)$ is not the maximum $R^{\rm ma}(\mathbf{D})$ that can be achieved, which implies that there exists $\mathbf{D}^{\rm s}$ such that $R^{\rm ma}(\mathbf{D}^{\rm s}) > R^{\rm ma}(\mathbf{D}^*)$ and $\bar{R}_{i^-{\rm r}}(\mathbf{D}_{i^-}^{\rm s}) > \bar{R}_{i^-{\rm r}}(\mathbf{D}_{i^-}^*) - \delta$ where δ is a positive number. Define $\mathbf{Z} \triangleq \mathbf{H}_{1{\rm r}}\mathbf{D}_1\mathbf{H}_{1{\rm r}}^{\rm H} + \mathbf{H}_{2{\rm r}}\mathbf{D}_2\mathbf{H}_{2{\rm r}}^{\rm H}$. It can be seen that $R^{\rm ma}(\mathbf{D})$ is a concave function of $\mathbf{Z}$. If $\mathbf{D}^*$ is not the optimal solution to the problem of maximizing $\min_i\{1/\mu_i\}$ subject to the constraints in (4.11), there exists $\mathbf{D}^{\rm q}$ such that

$R^{\mathrm{ma}}(\mathbf{D}^{\mathrm{q}}) \geq R^{\mathrm{ma}}(\mathbf{D}^{*})$ and $\bar{R}_{i^{-}\mathrm{r}}(\mathbf{D}^{\mathrm{q}}_{i^{-}}) > \bar{R}_{i^{-}\mathrm{r}}(\mathbf{D}^{*}_{i^{-}})$. Then, for any $0 < \alpha < 1$, there is a $\mathbf{D}^{\mathrm{c}}$ such that $\mathbf{D}^{\mathrm{c}}_{l} = \alpha\mathbf{D}^{\mathrm{q}}_{l} + (1-\alpha)\mathbf{D}^{\mathrm{s}}_{l}, \forall l$. Moreover, for any α such that

$$\frac{\bar{R}_{i^{-}\mathrm{r}}(\mathbf{D}^{*}_{i^{-}}) - \bar{R}_{i^{-}\mathrm{r}}(\mathbf{D}^{\mathrm{s}}_{i^{-}})}{\bar{R}_{i^{-}\mathrm{r}}(\mathbf{D}^{\mathrm{q}}_{i^{-}}) - \bar{R}_{i^{-}\mathrm{r}}(\mathbf{D}^{\mathrm{s}}_{i^{-}})} < \alpha < 1 \tag{B.1}$$

it can be shown that $\bar{R}_{i^{-}\mathrm{r}}(\mathbf{D}^{\mathrm{c}}_{i^{-}}) > \bar{R}_{i^{-}\mathrm{r}}(\mathbf{D}^{*}_{i^{-}})$ using the fact that $\bar{R}_{i\mathrm{r}}(\mathbf{D}_{i})$ is concave with respect to $\mathbf{D}_{i}, \forall i$. Denoting $\mathbf{Z}^{\mathrm{q}} = \mathbf{H}_{1\mathrm{r}}\mathbf{D}^{\mathrm{q}}_{1}\mathbf{H}^{\mathrm{H}}_{1\mathrm{r}} + \mathbf{H}_{2\mathrm{r}}\mathbf{D}^{\mathrm{q}}_{2}\mathbf{H}^{\mathrm{H}}_{2\mathrm{r}}$ and $\mathbf{Z}^{\mathrm{s}} = \mathbf{H}_{1\mathrm{r}}\mathbf{D}^{\mathrm{s}}_{1}\mathbf{H}^{\mathrm{H}}_{1\mathrm{r}} + \mathbf{H}_{2\mathrm{r}}\mathbf{D}^{\mathrm{s}}_{2}\mathbf{H}^{\mathrm{H}}_{2\mathrm{r}}$, it can be shown that $\mathbf{D}^{\mathrm{c}}_{i}, \forall i$ lead to $\mathbf{Z}^{\mathrm{c}} = \alpha\mathbf{Z}^{\mathrm{q}} + (1-\alpha)\mathbf{Z}^{\mathrm{s}}$ and therefore $R^{\mathrm{ma}}(\mathbf{D}^{\mathrm{c}}) \geq \alpha R^{\mathrm{ma}}(\mathbf{D}^{\mathrm{q}}) + (1-\alpha)R^{\mathrm{ma}}(\mathbf{D}^{\mathrm{s}}) > R^{\mathrm{ma}}(\mathbf{D}^{*})$. Hence, if $\mathbf{D}^{*}$ does not maximize $\bar{R}_{i^{-}\mathrm{r}}(\mathbf{D}_{i^{-}})$ subject to the constraints in (4.11), then $\bar{R}_{i^{-}\mathrm{r}}(\mathbf{D}_{i^{-}})$ and $R^{\mathrm{ma}}(\mathbf{D})$ can be simultaneously increased. The fact that $\bar{R}_{i^{-}\mathrm{r}}(\mathbf{D}_{i^{-}})$ can be increased means that $\min_{i}\{1/\mu_{i}\}$ can be increased, which implies that the BC phase sum-rate $\sum_{i}\hat{R}_{\mathrm{r}i}(\lambda_{i})$ can be increased according to Lemma 3.2 in Chapter 3 subject to the constraint that $\min_{i}\{1/\lambda_{i}\} = \min_{i}\{1/\mu_{i}\}$ as implied by Lemma 4.2. Given this result, the fact that $R^{\mathrm{ma}}(\mathbf{D})$ can be simultaneously increased suggests that $R^{\mathrm{tw}}(\mathbf{B},\mathbf{D})$ can be increased. This contradicts the fact that $\mathbf{D}^{*}$ is the optimal solution that maximizes $R^{\mathrm{tw}}(\mathbf{B},\mathbf{D})$ with $\mathbf{D}^{*}$ subject to the related constraints. Therefore, $\mathbf{D}^{*}$ must maximize $\min_{i}\{1/\mu_{i}\}$ subject to (4.11).

Proof of Property 4: It can be seen that the maximum achievable $1/\mu_{l}$ subject to the constraints

$$R^{\mathrm{ma}}(\mathbf{D}) \geq R^{\mathrm{obj}}, \quad \mathrm{Tr}(\mathbf{D}_{i}) \leq P^{\max}_{i}, \forall i \tag{B.2}$$

is a non-increasing function of R^{obj}. If $1/\mu^{*}_{\bar{l}} \leq 1/\mu^{*}_{l}$, according to property 1 of this theorem and the fact that $1/\mu_{\mathrm{ma}} < \max_{i}\{1/\mu_{i}\}$, it indicates that $1/\mu^{*}_{l} > 1/\mu^{*}_{\mathrm{ma}}$. Since $1/\mu^{0}_{\mathrm{ma}} > 1/\mu^{0}_{l}$ and the maximum achievable $1/\mu_{l}$ is a non-increasing function of R^{obj}, there exists $\tilde{\mathbf{D}}$ such that $1/\mu^{*}_{l} \geq 1/\tilde{\mu}_{l}$ and $1/\tilde{\mu}_{l} = 1/\tilde{\mu}_{\mathrm{ma}} \geq 1/\mu^{*}_{\mathrm{ma}}$. Using $1/\mu_{\mathrm{ma}} < \max_{i}\{1/\mu_{i}\}$ from Lemma 3.1 in Chapter 3, it infers that $1/\tilde{\mu}_{\bar{l}} > 1/\tilde{\mu}_{l} = 1/\tilde{\mu}_{\mathrm{ma}}$ at this point. Since the maximum $\bar{R}_{l\mathrm{r}}(\mathbf{D}_{l})$ cannot achieve $\hat{R}_{\mathrm{r}\bar{l}}(\lambda^{0})$ in the problem (4.10), it can be seen that $1/\tilde{\mu}_{l} = 1/\tilde{\mu}_{\mathrm{ma}} < 1/\lambda^{0}$. In such a case, the optimal strategy of the relay is to use $1/\lambda_{i} = 1/\tilde{\mu}_{\mathrm{ma}} < 1/\lambda^{0}, \forall i$, which does not consume the full power of the relay. Therefore, according to property 2 of this theorem, when $1/\tilde{\mu}_{l} = 1/\tilde{\mu}_{\mathrm{ma}}$, the $R^{\mathrm{tw}}(\mathbf{B},\mathbf{D})$ that can be achieved, specifically $R^{\mathrm{ma}}(\tilde{\mathbf{D}})$, is not the maximum that $R^{\mathrm{tw}}(\mathbf{B},\mathbf{D})$ can achieve. Moreover, since $1/\tilde{\mu}_{\mathrm{ma}} \geq 1/\mu^{*}_{\mathrm{ma}}$, it can be

seen that $R^{\mathrm{ma}}(\mathbf{D}^*) \leq R^{\mathrm{ma}}(\tilde{\mathbf{D}})$. As a result, $R^{\mathrm{tw}}(\mathbf{B}^*, \mathbf{D}^*) = R^{\mathrm{ma}}(\mathbf{D}^*) \leq R^{\mathrm{ma}}(\tilde{\mathbf{D}})$. Using the above-proved fact that $R^{\mathrm{ma}}(\tilde{\mathbf{D}})$ is not the maximum that $R^{\mathrm{tw}}(\mathbf{B}, \mathbf{D})$ can achieve, this result obtained under the assumption $1/\mu_{\bar{l}}^* \leq 1/\mu_l^*$ contradicts the assumption that $\mathbf{B}^*$ and $\mathbf{D}^*$ are optimal. Therefore, the assumption that $1/\mu_{\bar{l}}^* \leq 1/\mu_l^*$ must be invalid.

B.5 Proof of Theorem 4.3

The proof follows the same route as the proof of Theorem 4.2.

Proof of Property 1: As there exists no λ_l which satisfies the constraints in (4.20), it can be seen that $\sum_i \hat{R}_{\mathrm{r}i}(\lambda_i)$ cannot achieve $R^{\mathrm{ma}}(\mathbf{D}^0)$ subject to the constraint $\lambda_{\bar{l}} = \mu_l^0$, which is necessary as stated in Lemma 4.2. Therefore, it is necessary that $1/\mu_{\mathrm{ma}}^* < 1/\mu_{\mathrm{ma}}^0$. Given that $1/\mu_{\mathrm{ma}}^* < 1/\mu_{\mathrm{ma}}^0$, it can be shown that the resulting $R^{\mathrm{tw}}(\mathbf{B}, \mathbf{D})$ is not maximized if $1/\mu_{\mathrm{ma}}^* \leq \min_i\{1/\mu_i^*\}$. Therefore, it is necessary that $1/\mu_{\mathrm{ma}}^* > \min_i\{1/\mu_i^*\}$.

Proof of properties 2-3 from Section B.4 can be applied here after we substitute all λ^0 therein to μ_{ma}^0. Proof of property 4 of Theorem 4.2 can be directly applied here.

$\sim$

Appendix C

Proofs for Chapter 5

C.1 Proof of Lemma 5.1

It is proved that the function $\log|\mathbf{I}+\mathbf{A}\mathbf{X}^{-1}|$ is convex in $\mathbf{X}$ given that $\mathbf{A}$ is PSD [83]. Moreover, strong convexity holds if $\mathbf{A} \succ 0$. Therefore, the optimal solution can be characterized using the Karush-Kuhn-Tucker (KKT) conditions [90].

The Lagrangian of (5.6a) can be written as

$$L(\mathbf{X}, \lambda, \mathbf{Z}) = \log|\mathbf{A}+\mathbf{X}| - \log|\mathbf{X}| + \lambda(\text{Tr}\{\mathbf{X}\}-1) - \text{Tr}\{\mathbf{X}\mathbf{Z}\} \tag{C.1}$$

in which λ and $\mathbf{Z}$ are the Lagrange multipliers associated with (5.6b) and (5.6c), respectively. The KKT optimality conditions for the problem (5.6) are then given as

$$\text{Tr}\{\mathbf{X}\} \leq 1,\ \mathbf{X} \succeq 0,\ \lambda \geq 0, \tag{C.2}$$

$$\mathbf{Z} \succeq 0,\ \lambda\text{Tr}\{\mathbf{X}-1\} = 0,\ \text{Tr}\{\mathbf{X}\mathbf{Z}\} = 0, \tag{C.3}$$

$$(\mathbf{X}+\mathbf{A})^{-\mathrm{T}} - \mathbf{X}^{-\mathrm{T}} + \lambda\mathbf{I} - \mathbf{Z}^{\mathrm{T}} = \mathbf{0}. \tag{C.4}$$

It is not difficult to see that $\mathbf{X} \succ 0$ and $\text{Tr}\{\mathbf{X}\} = 1$ at optimality. Given that $\mathbf{X} \succ 0$ and $\mathbf{Z} \succeq 0$ at optimality, the condition $\text{Tr}\{\mathbf{X}\mathbf{Z}\} = 0$ indicates that $\mathbf{Z} = \mathbf{0}$. Then (C.4) becomes

$$(\mathbf{X}+\mathbf{A})^{-\mathrm{T}} = \mathbf{X}^{-\mathrm{T}} - \lambda\mathbf{I} \tag{C.5}$$

which further indicates that

$$\mathbf{X}+\mathbf{A} = (\mathbf{X}^{-1} - \lambda\mathbf{I})^{-1} \tag{C.6}$$

Using the matrix inversion lemma [91], the right-hand side of (C.6) is equivalent to

$$\mathbf{X} + \mathbf{X}(\mathbf{I} - \lambda\mathbf{X})^{-1}\lambda\mathbf{X}. \tag{C.7}$$

Then (C.6) can be written as

$$\mathbf{A} = \mathbf{X}(\lambda^{-1}\mathbf{I} - \mathbf{X})^{-1}\mathbf{X}. \tag{C.8}$$

Denoting the EVD of $\mathbf{X}$ as $\mathbf{X} = \mathbf{U}_{\mathrm{X}}\mathbf{\Lambda}_{\mathrm{X}}\mathbf{U}_{\mathrm{X}}^{\mathrm{H}}$, the expression (C.8) can be rewritten as

$$\mathbf{U}_{\mathrm{X}}^{\mathrm{H}}\mathbf{A}\mathbf{U}_{\mathrm{X}} = \mathbf{\Lambda}_{\mathrm{X}}(\lambda^{-1}\mathbf{I} - \mathbf{\Lambda}_{\mathrm{X}})^{-1}\mathbf{\Lambda}_{\mathrm{X}}. \tag{C.9}$$

Defining $\mathbf{\Lambda}_1 \triangleq \mathbf{U}_{\mathrm{X}}^{\mathrm{H}}\mathbf{A}\mathbf{U}_{\mathrm{X}}$, and using the fact that $\mathbf{U}_{\mathrm{X}}^{\mathrm{H}}\mathbf{A}\mathbf{U}_{\mathrm{X}}$ and $\mathbf{A}$ share the same eigen values, it can be found that $\mathbf{\Lambda}_1$ contains the eigenvalues of $\mathbf{A}$. Since $\mathbf{U}_{\mathrm{X}}^{\mathrm{H}}\mathbf{A}\mathbf{U}_{\mathrm{X}}$ gives the matrix of eigenvalues of $\mathbf{A}$, it must hold that $\mathbf{U}_{\mathrm{X}} = \mathbf{U}_{\mathrm{A}}$. Therefore, using $\mathbf{U}_{\mathrm{X}} = \mathbf{U}_{\mathrm{A}}$, we obtain that

$$\mathbf{\Lambda}_{\mathrm{A}} = \mathbf{\Lambda}_{\mathrm{X}}(\lambda^{-1}\mathbf{I} - \mathbf{\Lambda}_{\mathrm{X}})^{-1}\mathbf{\Lambda}_{\mathrm{X}} \tag{C.10}$$

which gives (recall that $\mathbf{A} \succ 0$ and $\mathbf{X} \succ 0$ at optimality)

$$\mathbf{\Lambda}_{\mathrm{X}}\mathbf{\Lambda}_{\mathrm{A}}^{-1}\mathbf{\Lambda}_{\mathrm{X}} = \lambda^{-1}\mathbf{I} - \mathbf{\Lambda}_{\mathrm{X}}. \tag{C.11}$$

Finally, the following equation

$$\mathbf{\Lambda}_{\mathrm{X}}^2 + \mathbf{\Lambda}_{\mathbf{A}}\mathbf{\Lambda}_{\mathbf{X}} = \lambda^{-1}\mathbf{\Lambda}_{\mathrm{A}} \tag{C.12}$$

holds, which leads to (5.7).

C.2 Proof of Lemma 5.2

If $\mathbf{B}$ is positive definite, the following matrix

$$\bar{\mathbf{B}} = \mathbf{B} + \begin{matrix} \\ r_{\mathrm{z}} \\ n_{\mathrm{z}}-r_{\mathrm{z}} \end{matrix} \overset{\begin{matrix} r_{\mathrm{z}} & n_{\mathrm{z}}-r_{\mathrm{z}} \end{matrix}}{\begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \sigma^2\mathbf{I} \end{bmatrix}}. \tag{C.13}$$

and its inverse $\bar{\mathbf{B}}^{-1}$ are also positive definite. Given that $\bar{\mathbf{B}}$ is positive definite, it can be seen that the two blocks on the diagonal of $\bar{\mathbf{B}}$ are both positive definite. Then, using block matrix inversion [92], it follows that the first block of $\bar{\mathbf{B}}^{-1}$ is $(\mathbf{B}_{11} - \mathbf{B}_{12}(\sigma^2\mathbf{I} + \mathbf{B}_{22})^{-1}\mathbf{B}_{21})^{-1}$, which is the inverse of $\tilde{\mathbf{B}}$. Given that $\bar{\mathbf{B}}^{-1}$ is positive definite, the first block of $\bar{\mathbf{B}}^{-1}$, i.e., the inverse of $\tilde{\mathbf{B}}$ must also be positive definite. Therefore, $\tilde{\mathbf{B}}$ is also positive definite. This proves Lemma 5.2.

C.3 Proof of Theorem 5.1

Using the definitions (5.9), (5.10), and (5.13), the objective function in (5.4) can be rewritten as

$$
\begin{aligned}
R^{\mathrm{J}} &= \log\left|\mathbf{I}+\mathbf{B}(\tilde{\mathbf{\Omega}}_{\mathrm{z}}\tilde{\mathbf{Q}}_{\mathrm{z}}\tilde{\mathbf{\Omega}}_{\mathrm{z}}^{\mathrm{H}}+\sigma^2\mathbf{I})^{-1}\right| \\
&= \log\left|\mathbf{I}+\tilde{\mathbf{\Omega}}_{\mathrm{z}}^{-1}\mathbf{B}\tilde{\mathbf{\Omega}}_{\mathrm{z}}^{-\mathrm{H}}(\tilde{\mathbf{Q}}_{\mathrm{z}}+\sigma^2\tilde{\mathbf{\Omega}}_{\mathrm{z}}^{-1}\tilde{\mathbf{\Omega}}_{\mathrm{z}}^{-\mathrm{H}})^{-1}\right| \\
&= \log\left|\mathbf{I}+\begin{bmatrix}\mathbf{\Omega}_{\mathrm{z}}^{+} & \mathbf{0}\\ \mathbf{0} & \mathbf{I}\end{bmatrix}^{-1}\begin{bmatrix}\mathbf{B}_{11} & \mathbf{B}_{12}\\ \mathbf{B}_{21} & \mathbf{B}_{22}\end{bmatrix}\begin{bmatrix}\mathbf{\Omega}_{\mathrm{z}}^{+\mathrm{H}} & \mathbf{0}\\ \mathbf{0} & \mathbf{I}\end{bmatrix}^{-1}\left(\begin{bmatrix}\mathbf{Q}_{\mathrm{z}}' & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{bmatrix}+\sigma^2\begin{bmatrix}\mathbf{\Omega}_{\mathrm{z}}^{+-1}\mathbf{\Omega}_{\mathrm{z}}^{+-\mathrm{H}} & \mathbf{0}\\ & \mathbf{0} \quad\quad \mathbf{I}\end{bmatrix}\right)^{-1}\right| \\
&= \log\left|\mathbf{I}+\begin{bmatrix}\mathbf{\Omega}_{\mathrm{z}}^{+-1}\mathbf{B}_{11}\mathbf{\Omega}_{\mathrm{z}}^{+-\mathrm{H}} & \mathbf{\Omega}_{\mathrm{z}}^{+-1}\mathbf{B}_{12}\\ \mathbf{B}_{21}\mathbf{\Omega}_{\mathrm{z}}^{+-\mathrm{H}} & \mathbf{B}_{22}\end{bmatrix}\begin{bmatrix}(\mathbf{Q}_{\mathrm{z}}'+\sigma^2\mathbf{\Omega}_{\mathrm{z}}^{+-1}\mathbf{\Omega}_{\mathrm{z}}^{+-\mathrm{H}})^{-1} & \mathbf{0}\\ \mathbf{0} & \frac{1}{\sigma^2}\mathbf{I}\end{bmatrix}\right| \\
&= \log\left|\begin{bmatrix}\mathbf{I}+\mathbf{\Omega}_{\mathrm{z}}^{+-1}\mathbf{B}_{11}\mathbf{\Omega}_{\mathrm{z}}^{+-\mathrm{H}}\mathbf{J}^{-1} & \frac{1}{\sigma^2}\mathbf{\Omega}_{\mathrm{z}}^{+-1}\mathbf{B}_{12}\\ \mathbf{B}_{21}\mathbf{\Omega}_{z}^{+-\mathrm{H}}\mathbf{J}^{-1} & \mathbf{I}+\frac{1}{\sigma^2}\mathbf{B}_{22}\end{bmatrix}\right|
\end{aligned}
\tag{C.14}
$$

where in the last step $\mathbf{J} \triangleq \mathbf{Q}_{\mathrm{z}}' + \sigma^2\mathbf{\Omega}_{\mathrm{z}}^{+-1}\mathbf{\Omega}_{\mathrm{z}}^{+-\mathrm{H}}$.

Since the matrix $\mathbf{H}_{\mathrm{r}}\mathbf{Q}_{\mathrm{s}}\mathbf{H}_{\mathrm{r}}^{\mathrm{H}}$ is positive definite, $\mathbf{B}$, and consequently $\mathbf{B}_{11}$ and $\mathbf{B}_{22}$, are all positive-definite. The rate R^{J} in (C.14) can be simplified as

$$R^{\mathrm{J}} = R^0 + \bar{R}^{\mathrm{J}} \tag{C.15}$$

where

$$R^0 = \log\left|\mathbf{I}+\frac{1}{\sigma^2}\mathbf{B}_{22}\right| \tag{C.16}$$

is the part of rate that is not affected by jamming which is non-zero if $r_{\mathrm{z}} < n_{\mathrm{r}}$ and

$$\bar{R}^{\mathrm{J}} = \log\left|\mathbf{I}+\mathbf{\Omega}_{\mathrm{z}}^{+-1}\mathbf{B}_{11}\mathbf{\Omega}_{\mathrm{z}}^{+-\mathrm{H}}\mathbf{J}^{-1}-\frac{1}{\sigma^2}\mathbf{\Omega}_{\mathrm{z}}^{+-1}\mathbf{B}_{12}(\mathbf{I}+\frac{1}{\sigma^2}\mathbf{B}_{22})^{-1}\mathbf{B}_{21}\mathbf{\Omega}_{\mathrm{z}}^{+-\mathrm{H}}\mathbf{J}^{-1}\right| \tag{C.17}$$

is the part of the rate that is affected by jamming. Therefore, the minimization of R^{J} in (5.3a) is equivalent to minimizing $\bar{R}^{\mathrm{J}}$. Using the definition of $\tilde{\mathbf{B}}$ in (5.9), $\bar{R}^{\mathrm{J}}$ can be rewritten as

$$\bar{R}^{\mathrm{J}} = \log\left|\mathbf{I}+\mathbf{\Omega}_{\mathrm{z}}^{+-1}\tilde{\mathbf{B}}\mathbf{\Omega}_{\mathrm{z}}^{+-\mathrm{H}}(\mathbf{Q}_{\mathrm{z}}'+\sigma^2\mathbf{\Omega}_{\mathrm{z}}^{+-1}\mathbf{\Omega}_{\mathrm{z}}^{+-\mathrm{H}})^{-1}\right|. \tag{C.18}$$

Using Lemma 5.2, it can be seen that $\tilde{\mathbf{B}}$ is positive definite when $\mathbf{B}$ is positive definite. Then, Lemma 5.1 can be used to find such $\mathbf{Q}'$ that minimizes (C.18)

subject to the trace constraint $\mathrm{Tr}\{\mathbf{Q}_{\mathrm{z}}'\} \leq P_{\mathrm{z}}$. Using (5.7), the definition $\tilde{\mathbf{A}} \triangleq \mathbf{\Omega}_{\mathrm{z}}^{+-1}\tilde{\mathbf{B}}\mathbf{\Omega}_{\mathrm{z}}^{+-\mathrm{H}}$, and the EVD $\tilde{\mathbf{A}} = \mathbf{U}_{\tilde{\mathbf{A}}}\mathbf{\Lambda}_{\tilde{\mathbf{A}}}\mathbf{U}_{\tilde{\mathbf{A}}}^{\mathrm{H}}$, the $\mathbf{Q}_{\mathrm{z}}'$ that minimizes (C.18), or equivalently (C.17), subject to $\mathrm{Tr}\{\mathbf{Q}_{\mathrm{z}}'\} \leq P_{\mathrm{z}}$ can be found as

$$\mathbf{Q}_{\mathrm{z}}' = \mathbf{U}_{\tilde{\mathbf{A}}}\sqrt{\frac{1}{\lambda}\mathbf{\Lambda}_{\tilde{\mathbf{A}}} + \frac{1}{4}\mathbf{\Lambda}_{\tilde{\mathbf{A}}}^2}\mathbf{U}_{\tilde{\mathbf{A}}}^{\mathrm{H}} - \mathbf{\Omega}_{\mathrm{z}}^{+-1}\big(\frac{1}{2}\tilde{\mathbf{B}} + \sigma^2\mathbf{I}\big)\mathbf{\Omega}_{\mathrm{z}}^{+-\mathrm{H}} \tag{C.19}$$

under the condition that the above $\mathbf{Q}_{\mathrm{z}}'$ is PSD. Here λ is chosen such that $\mathrm{Tr}\{\mathbf{Q}_{\mathrm{z}}'\} = P_{\mathrm{z}}$.

C.4 Proof of Theorem 5.2

The proof of Theorem 5.2 follows the same route as the proof of Theorem 5.1 in Section C.4 till the expression (C.18). Then, using (5.18), the $\bar{R}^{\mathrm{J}}$ in (C.18) can be rewritten as

$$\begin{aligned}
\bar{R}^{\mathrm{J}} &= \log\Big|\mathbf{I} + \tilde{\mathbf{A}}(\mathbf{Q}_{\mathrm{z}}' + \sigma^2\mathbf{\Omega}_{\mathrm{z}}^{+-1}\mathbf{\Omega}_{\mathrm{z}}^{+-\mathrm{H}})^{-1}\Big| \\
&= \log\left|\mathbf{I} + \begin{bmatrix}\mathbf{U}_{\tilde{\mathbf{A}}1} & \mathbf{U}_{\tilde{\mathbf{A}}2}\end{bmatrix}\begin{bmatrix}\mathbf{\Lambda}_{\tilde{\mathbf{A}}}^{+} & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{bmatrix}\begin{bmatrix}\mathbf{U}_{\tilde{\mathbf{A}}1}^{\mathrm{H}}\\ \mathbf{U}_{\tilde{\mathbf{A}}2}^{\mathrm{H}}\end{bmatrix}\mathbf{Q}_{\mathrm{z}}''^{-1}\right| \\
&= \log\left|\mathbf{I} + \begin{bmatrix}\mathbf{\Lambda}_{\tilde{\mathbf{A}}}^{+} & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{bmatrix}\left(\begin{bmatrix}\mathbf{U}_{\tilde{\mathbf{A}}1}^{\mathrm{H}}\\ \mathbf{U}_{\tilde{\mathbf{A}}2}^{\mathrm{H}}\end{bmatrix}\mathbf{Q}_{\mathrm{z}}''\begin{bmatrix}\mathbf{U}_{\tilde{\mathbf{A}}1} & \mathbf{U}_{\tilde{\mathbf{A}}2}\end{bmatrix}\right)^{-1}\right| \\
&= \log\left|\mathbf{I} + \begin{bmatrix}\mathbf{\Lambda}_{\tilde{\mathbf{A}}}^{+} & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{bmatrix}\begin{bmatrix}\mathbf{U}_{\tilde{\mathbf{A}}1}^{\mathrm{H}}\mathbf{Q}_{\mathrm{z}}''\mathbf{U}_{\tilde{\mathbf{A}}1} & \mathbf{U}_{\tilde{\mathbf{A}}1}^{\mathrm{H}}\mathbf{Q}_{\mathrm{z}}''\mathbf{U}_{\tilde{\mathbf{A}}2}\\ \mathbf{U}_{\tilde{\mathbf{A}}2}^{\mathrm{H}}\mathbf{Q}_{\mathrm{z}}''\mathbf{U}_{\tilde{\mathbf{A}}1} & \mathbf{U}_{\tilde{\mathbf{A}}2}^{\mathrm{H}}\mathbf{Q}_{\mathrm{z}}''\mathbf{U}_{\tilde{\mathbf{A}}2}\end{bmatrix}^{-1}\right| \\
&= \log\left|\mathbf{I} + \begin{bmatrix}\mathbf{\Lambda}_{\tilde{\mathbf{A}}}^{+} & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{bmatrix}\begin{bmatrix}\mathbf{F}_1^{-1} & \mathbf{F}_{12}\\ \mathbf{F}_{21} & \mathbf{F}_2^{-1}\end{bmatrix}\right| \\
&= \log\Big|\mathbf{I} + \mathbf{\Lambda}_{\tilde{\mathbf{A}}}^{+}\mathbf{F}_1^{-1}\Big|
\end{aligned} \tag{C.20}$$

where $\mathbf{Q}_{\mathrm{z}}'' \triangleq \mathbf{Q}_{\mathrm{z}}' + \sigma^2\mathbf{\Omega}_{\mathrm{z}}^{+-1}\mathbf{\Omega}_{\mathrm{z}}^{+-\mathrm{H}}$ in the second step. The result on block matrix inversion is used in the last step [92], in which

$$\mathbf{F}_1 \triangleq \mathbf{F}_1^1 - \mathbf{F}_1^2 \tag{C.21}$$

with $\mathbf{F}_1^1$ and $\mathbf{F}_1^2$ given by

$$\mathbf{F}_1^1 \triangleq \mathbf{U}_{\tilde{\mathbf{A}}1}^{\mathrm{H}}\mathbf{Q}_{\mathrm{z}}''\mathbf{U}_{\tilde{\mathbf{A}}1} \tag{C.22}$$

$$\mathbf{F}_1^2 \triangleq \mathbf{U}_{\tilde{\mathbf{A}}1}^{\mathrm{H}}\mathbf{Q}_{\mathrm{z}}''\mathbf{U}_{\tilde{\mathbf{A}}2}(\mathbf{U}_{\tilde{\mathbf{A}}2}^{\mathrm{H}}\mathbf{Q}_{\mathrm{z}}''\mathbf{U}_{\tilde{\mathbf{A}}2})^{-1}\mathbf{U}_{\tilde{\mathbf{A}}2}^{\mathrm{H}}\mathbf{Q}_{\mathrm{z}}''\mathbf{U}_{\tilde{\mathbf{A}}1} \tag{C.23}$$

and

$$\mathbf{F}_{12} \triangleq -(\mathbf{U}_{\tilde{\mathbf{A}}1}^{\mathrm{H}}\mathbf{Q}_{\mathrm{z}}''\mathbf{U}_{\tilde{\mathbf{A}}1})^{-1}\mathbf{U}_{\tilde{\mathbf{A}}1}^{\mathrm{H}}\mathbf{Q}_{\mathrm{z}}''\mathbf{U}_{\tilde{\mathbf{A}}2}\mathbf{F}_2^{-1} \tag{C.24}$$

$$\mathbf{F}_{21} \triangleq -(\mathbf{U}_{\tilde{\mathbf{A}}2}^{\mathrm{H}}\mathbf{Q}_{\mathrm{z}}''\mathbf{U}_{\tilde{\mathbf{A}}2})^{-1}\mathbf{U}_{\tilde{\mathbf{A}}2}^{\mathrm{H}}\mathbf{Q}_{\mathrm{z}}''\mathbf{U}_{\tilde{\mathbf{A}}1}\mathbf{F}_1^{-1} \tag{C.25}$$

$$\mathbf{F}_2 \triangleq \mathbf{U}_{\tilde{\mathbf{A}}2}^{\mathrm{H}}\mathbf{Q}_{\mathrm{z}}''\mathbf{U}_{\tilde{\mathbf{A}}2} - \mathbf{U}_{\tilde{\mathbf{A}}2}^{\mathrm{H}}\mathbf{Q}_{\mathrm{z}}''\mathbf{U}_{\tilde{\mathbf{A}}1}(\mathbf{U}_{\tilde{\mathbf{A}}1}^{\mathrm{H}}\mathbf{Q}_{\mathrm{z}}''\mathbf{U}_{\tilde{\mathbf{A}}1})^{-1}\mathbf{U}_{\tilde{\mathbf{A}}1}^{\mathrm{H}}\mathbf{Q}_{\mathrm{z}}''\mathbf{U}_{\tilde{\mathbf{A}}2}. \tag{C.26}$$

Recalling the optimization problem (5.6), it can be seen from the last step of (C.20) that $\bar{R}^{\mathrm{J}}$ is not minimized if the trace of $\mathbf{F}_1$ can be increased under the jammer's power constraint. Therefore, a necessary condition for minimizing (C.20) is that the trace of $\mathbf{F}_1$ is maximized given the trace constraint of $\mathbf{Q}_{\mathrm{z}}'$.

Considering the fact that $\mathrm{Tr}\{\mathbf{U}_{\tilde{\mathbf{A}}1}^{\mathrm{H}}\mathbf{Q}_{\mathrm{z}}''\mathbf{U}_{\tilde{\mathbf{A}}1}\} \leq \mathrm{Tr}\{\mathbf{Q}_{\mathrm{z}}''\}$ and that $\mathbf{F}_1^2$ is PSD, maximizing $\mathrm{Tr}\{\mathbf{F}_1\}$ requires that $\mathbf{Q}_{\mathrm{z}}''$ must have the following form

$$\mathbf{Q}_{\mathrm{z}}'' = \mathbf{U}_{\tilde{\mathbf{A}}1}\mathbf{D}_{\mathrm{x}}\mathbf{U}_{\tilde{\mathbf{A}}1}^{\mathrm{H}} \tag{C.27}$$

in which $\mathbf{D}_{\mathrm{x}}$ is a $r_{\tilde{\mathbf{A}}} \times r_{\tilde{\mathbf{A}}}$ PSD matrix to be determined. The matrix $\mathbf{D}_{\mathrm{x}}$ should satisfy the constraint $\mathrm{Tr}\{\mathbf{D}_{\mathrm{x}}\} \leq P_{\mathrm{z}} + \sigma^2\mathrm{Tr}\{\mathbf{\Omega}_{\mathrm{z}}^{+-1}\mathbf{\Omega}_{\mathrm{z}}^{+-\mathrm{H}}\}$.

Using (C.27), $\mathbf{F}_1^2$ is $\mathbf{0}$ and $\mathbf{F}_1$ in (C.21) is equal to $\mathbf{D}_{\mathrm{x}}^{-1}$. Consequently, (C.20) can be rewritten as

$$R^{\mathrm{J}} = \log\left|\mathbf{I} + \mathbf{\Lambda}_{\tilde{\mathbf{A}}}^{+}\mathbf{D}_{\mathrm{x}}^{-1}\right|. \tag{C.28}$$

Therefore, the matrix $\mathbf{Q}_{\mathrm{z}}''$ in (C.27) corresponds to spreading the power (including jamming power and noise power) on the eigen-channels corresponding to the positive eigenvalues of $\tilde{\mathbf{A}}$. Indeed, 'spilling' power on the null space of $\tilde{\mathbf{A}}$ cannot be optimal.

Using the result from Lemma 5.1, the optimal $\mathbf{D}_{\mathrm{x}}$ is given as

$$\mathbf{D}_{\mathrm{x}} = \sqrt{\frac{1}{\lambda}\mathbf{\Lambda}_{\tilde{\mathbf{A}}}^{+} + \frac{1}{4}\mathbf{\Lambda}_{\tilde{\mathbf{A}}}^{+\,2}} - \frac{1}{2}\mathbf{\Lambda}_{\tilde{\mathbf{A}}}^{+}. \tag{C.29}$$

Accordingly, the optimal $\mathbf{Q}'$ is given as

$$\mathbf{Q}_{\mathrm{z}}' = \mathbf{U}_{\tilde{\mathbf{A}}1}\sqrt{\frac{1}{\lambda}\mathbf{\Lambda}_{\tilde{\mathbf{A}}}^{+} + \frac{1}{4}\mathbf{\Lambda}_{\tilde{\mathbf{A}}}^{+\,2}}\mathbf{U}_{\tilde{\mathbf{A}}1}^{\mathrm{H}} - \frac{1}{2}\mathbf{U}_{\tilde{\mathbf{A}}1}\mathbf{\Lambda}_{\tilde{\mathbf{A}}}^{+}\mathbf{U}_{\tilde{\mathbf{A}}1}^{\mathrm{H}} - \sigma^2\mathbf{\Omega}_{\mathrm{z}}^{+-1}\mathbf{\Omega}_{\mathrm{z}}^{+-\mathrm{H}} \tag{C.30}$$

if the above $\mathbf{Q}'$ is positive semi-definite (PSD), where λ are chosen such that $\mathrm{Tr}\{\mathbf{Q}_z'\} = P_{\mathrm{z}}$.

C.5 Proof of Lemma 5.3

The four-step procedure in Table 5.1 uses the sequential parametric convex approximation method [93]. The convergence of this method to optimality is proved in [93] assuming that the convex relaxations (in our case, the righthand side of (5.22b)) are "convex upper estimate functions" of the righthand side of the original nonconvex constraints (in our case, the righthand side of (5.21b)). Therefore, it is sufficient to prove that

$$\begin{aligned}\log\left|\mathbf{Q}'_{\mathrm{z}}+\mathbf{D}_0+\tilde{\mathbf{A}}\right| \;\leq\; & \log\left|\mathbf{Q}'^{\dagger}_{\mathrm{z}}+\mathbf{D}_0+\tilde{\mathbf{A}}\right|+\mathrm{Tr}\{\left(\mathbf{Q}'^{\dagger}_{\mathrm{z}}+\mathbf{D}_0+\tilde{\mathbf{A}}\right)^{-1}\mathbf{Q}'_{\mathrm{z}}\} \\ & -\mathrm{Tr}\{\left(\mathbf{Q}'^{\dagger}_{\mathrm{z}}+\mathbf{D}_0+\tilde{\mathbf{A}}\right)^{-1}\mathbf{Q}'^{\dagger}_{\mathrm{z}}\} \end{aligned} \tag{C.31}$$

for all $\mathbf{Q}'_{\mathrm{z}}$ and $\mathbf{Q}'^{\dagger}_{\mathrm{z}}$ which are positive definite and satisfy (5.21c), and that the righthand-side of (C.31) is convex and continuously differentiable with respect to $\mathbf{Q}'_{\mathrm{z}}$ given $\mathbf{Q}'^{\dagger}_{\mathrm{z}}$. It is not difficult to see that the latter condition is satisfied. Thus, we only need to prove the first point. Using Taylor expansion, it can be shown that the righthand-side of (C.31) is the tangent of the function $f(\mathbf{Q}'_{\mathrm{z}})=\log\left|\mathbf{Q}'_{\mathrm{z}}+\mathbf{D}_0+\tilde{\mathbf{A}}\right|$ at $\mathbf{Q}'_{\mathrm{z}}=\mathbf{Q}'^{\dagger}_{\mathrm{z}}$ [94]. Recalling the fact that the function $f(\mathbf{Q}'_{\mathrm{z}})=\log\left|\mathbf{Q}'_{\mathrm{z}}+\mathbf{D}_0+\tilde{\mathbf{A}}\right|$ is strictly concave when $\mathbf{Q}'_{\mathrm{z}}\succ 0$, it can be seen that (C.31) is satisfied for all valid $\mathbf{Q}'_{\mathrm{z}}$ and $\mathbf{Q}'^{\dagger}_{\mathrm{z}}$.

C.6 Proof of Theorem 5.3

First, the following train of equalities holds true and leads to the simplified expression for $\sum_i w_i R_i^{\mathrm{C}}$ in (5.35)

$$\begin{aligned}
\sum_i w_i R_i^{\mathrm{C}} &= \sum_i w_i \log\left(1+(1+\xi_i)^2\left(g_i^{-1}|h_{\mathrm{z}i}|^2\Big(\sum_{j\neq i}\xi_j^2 q_j^{\mathrm{v}}+\sigma_{\mathrm{z}}^2\Big)+g_i^{-1}\sigma_i^2\right)^{-1}\right)\\
&= \sum_i w_i \log\left(1+(1+\xi_i)^2\left(|h_i^{-1}h_{\mathrm{z}i}|^2(P_{\mathrm{z}}-\xi_i^2 q_i^{\mathrm{v}})+g_i^{-1}\sigma_i^2\right)^{-1}\right)\\
&= \sum_i w_i \log\left(1+(1+\xi_i)^2\left(\frac{1}{q_i^{\mathrm{v}}}(P_{\mathrm{z}}-\xi_i^2 q_i^{\mathrm{v}})+g_i^{-1}\sigma_i^2\right)^{-1}\right)\\
&= \sum_i w_i \log\left(1+(1+\xi_i)^2\left(\frac{P_{\mathrm{z}}}{q_i^{\mathrm{v}}}-\xi_i^2+\frac{\sigma_i^2}{g_i}\right)^{-1}\right)\\
&= \sum_i w_i \log\left(1+(1+\xi_i)^2\big(\gamma_i-\xi_i^2+\rho_i\big)^{-1}\right) \qquad \text{(C.32)}
\end{aligned}$$

where $\rho_i \triangleq \sigma_i^2/g_i$. The second row in (C.32) uses the fact that the jammer uses full power, i.e., $\xi_i^2\sum_i q_i^{\mathrm{v}}+\sigma_{\mathrm{z}}^2 = P_{\mathrm{z}}$. It can be seen that γ_i is the ratio of the maximum jamming power and the power that is required to completely cancel the signal from the ith transmitter. Therefore, the range of ξ_i to be considered is $\xi_i \in [-\min\{1,\sqrt{\gamma_i}\},0]$. We prove the theorem by showing that the Hessian matrix of $\sum_i w_i R_i^{\mathrm{C}}$ is PSD with respect to ξ_i in the above interval for all i.

Denote $v_{i1} \triangleq \gamma_i-\xi_i^2+\rho_i+(1+\xi_i)^2$ and $v_{i2} \triangleq \gamma_i-\xi_i^2+\rho_i$. Then $\sum_i w_i R_i^{\mathrm{C}} = \sum_i w_i\big(\log v_{i1}-\log v_{i2}\big)$. The first-order and seconder-order derivatives of $\sum_i w_i R_i^{\mathrm{C}}$ with respect to ξ_i are given as[1]

$$\frac{\partial\sum_i w_i R_i^{\mathrm{C}}}{\partial\xi_i} = \frac{2w_i}{v_{i1}}-\frac{-2w_i\xi_i}{v_{i2}} \qquad \text{(C.33)}$$

$$\begin{aligned}
\frac{\partial^2\sum_i w_i R_i^{\mathrm{C}}}{\partial\xi_i^2} &= -\frac{4w_i}{v_{i1}^2}+\frac{2w_i}{v_{i2}}+\frac{4w_i\xi_i^2}{v_{i2}^2}\\
&= \frac{2w_i}{v_{i1}^2 v_{i2}^2}\left(v_{i1}^2 v_{i2}+2\xi_i^2 v_{i1}^2-2v_{i2}^2\right). \qquad \text{(C.34)}
\end{aligned}$$

[1] A constant multiplier $1/\ln 2$ is neglected.

Define $\phi_i \triangleq \gamma_i + \rho_i$. Then $v_{i2} = \phi_i - \xi_i^2$ and (C.34) can be rewritten as

$$\begin{aligned}\frac{\partial^2 \sum\limits_i w_i R_i^{\mathrm{C}}}{\partial \xi_i^2} &= \frac{2w_i}{v_{i1}^2 v_{i2}^2}\Big(v_{i1}^2(\phi_i + \xi_i^2) - 2(\phi_i - \xi_i^2)^2\Big) \\ &= \frac{2w_i}{v_{i1}^2 v_{i2}^2}\Big((\phi_i - \xi_i^2)\big(v_{i1}^2 - 2(\phi_i - \xi_i^2)\big) + 2\xi_i^2 v_{i1}^2\Big). \quad \text{(C.35)}\end{aligned}$$

Denote $l_i \triangleq (\phi_i - \xi_i^2)\big(v_{i1}^2 - 2(\phi_i - \xi_i^2)\big) + 2\xi_i^2 v_{i1}^2$. Using the fact that $v_{i1} = \phi_i + 2\xi_i + 1$, l_i can be expressed as

$$\begin{aligned} l_i &= (\phi_i - \xi_i^2)\big(\phi_i^2 + 4\xi_i^2 + 1 + 4\phi_i\xi_i + 2\phi_i + 4\xi_i - 2(\phi_i - \xi_i^2)\big) + 2\xi_i^2 v_{i1}^2 \\ &= (\phi_i - \xi_i^2)\big(6\xi_i^2 + 4\xi_i + 4\phi_i\xi_i + \phi_i^2 + 1\big) + 2\xi_i^2 v_{i1}^2 \quad \text{(C.36)}\end{aligned}$$

Moreover, using the fact that $v_{i1}^2 = (v_{i2} + (1 + \xi_i)^2)^2$, the last item in the above equation can be expanded as

$$2\xi_i^2 v_{i1}^2 = 2\xi_i^2 v_{i2}^2 + 4\xi_i^2(1 + \xi_i)^2 v_{i2} + 2\xi_i^2(1 + \xi_i)^4 \quad \text{(C.37)}$$

Substituting (C.37) back into the expression l_i (C.36) and using the fact that $v_{i2} = \phi_i - \xi_i^2$, we obtain

$$\begin{aligned} l_i &= (\phi_i - \xi_i^2)\Big(6\xi_i^2 + 4\xi_i + 4\phi_i\xi_i + \phi_i^2 + 1 + 2\xi_i^2(\phi_i - \xi_i^2) + 4\xi_i^2(1 + \xi_i)^2\Big) + 2\xi_i^2(1 + \xi_i)^4 \\ &= (\phi_i - \xi_i^2)\big(2\xi_i^4 + 8\xi_i^3 + 10\xi_i^2 + 4\xi_i + 2\phi_i\xi_i^2 + 4\phi_i\xi_i + \phi_i^2 + 1\big) + 2\xi_i^2(1 + \xi_i)^4 \\ &= (\phi_i - \xi_i^2)\Big(2(1 + \xi_i)^4 - 2\xi_i^2 - 4\xi_i + 2\phi_i\xi_i^2 + 4\phi_i\xi_i + \phi_i^2 - 1\Big) + 2\xi_i^2(1 + \xi_i)^4 \\ &= (\phi_i - \xi_i^2)\Big(2(1 + \xi_i)^4 + 2(\phi_i - 1)(\xi_i + 1)^2 + (\phi_i - 1)^2\Big) + 2\xi_i^2(1 + \xi_i)^4 \\ &= 2(\phi_i - \xi_i^2)\left(\left((1 + \xi_i)^2 + \frac{\phi_i - 1}{2}\right)^2 + \frac{(\phi_i - 1)^2}{4}\right) + 2\xi_i^2(1 + \xi_i)^4 \quad \text{(C.38)}\end{aligned}$$

Substituting (C.38) back into (C.35), we have

$$\frac{\partial^2 \sum\limits_i w_i R_i^{\mathrm{C}}}{\partial \xi_i^2} = \frac{2w_i}{v_{i1}^2 v_{i2}^2}\left(2(\phi_i - \xi_i^2)\left(\left((1 + \xi_i)^2 + \frac{\phi_i - 1}{2}\right)^2 + \frac{(\phi_i - 1)^2}{4}\right) + 2\xi_i^2(1 + \xi_i)^4\right). \quad \text{(C.39)}$$

Since $\phi_i = \gamma_i + \rho_i$, it can be seen that the above second order derivative is always non-negative if $-\min\left(\sqrt{\gamma_i}, 1\right) \leq \xi_i \leq 0$. It is also not difficult to see that

$$\frac{\partial^2 \sum\limits_i w_i R_i^{\mathrm{C}}}{\partial \xi_i \partial \xi_j} = 0 \quad \forall j \neq i, \forall i. \quad \text{(C.40)}$$

Therefore, the Hessian matrix of $\sum_i w_i R_i^{\mathrm{C}}$ with respect to ξ_i's is diagonal and PSD. This completes the proof for Theorem 5.3.

$\sim$

Appendix D

Proofs for Chapter 6

D.1 Proof of Theorem 6.1

First we prove the necessity of (6.5). The expectation of (6.1) with respect to t_j can be found as

$$\mathbb{E}_{t_j}\{u_i(t_i,t_j)\} = \int_{\mathcal{S}_j} \frac{b_{ii}^2}{\sigma_2^2 + b_{ji}^2(1-t_j)} f(t_j)\, dt_j + t_i \int_{\mathcal{S}_j} \kappa_j(t_j) f(t_j)\, dt_j \tag{D.1}$$

where

$$\kappa_j(t_j) = \left(\frac{b_{ii}^1}{\sigma_1^2 + b_{ji}^1 t_j} - \frac{b_{ii}^2}{\sigma_2^2 + b_{ji}^2(1-t_j)} \right). \tag{D.2}$$

In order to satisfy (6.3) in this game, it is necessary that

$$\int_{t_j \in \mathcal{S}_j} \kappa_j(t_j) f_j(t_j)\, dt_j = 0. \tag{D.3}$$

Since $\kappa_j(t_j)$ is a decreasing function of t_j on $[0,1]$ with

$$\kappa_j(0) < 0 \Big|_{\frac{b_{ii}^1}{b_{ii}^2} < \frac{\sigma_1^2}{\sigma_2^2 + b_{ji}^2}} \; ; \; \kappa_j(1) > 0 \Big|_{\frac{b_{ii}^1}{b_{ii}^2} > \frac{\sigma_1^2 + b_{ji}^1}{\sigma_2^2}}, \tag{D.4}$$

there is no strategy $f_j(t_j)$ that satisfies (D.3) if (6.5) is not satisfied. It can be shown that there exists only one MSNE which is actually a pure strategy Nash equilibrium (NE) if (6.5) is not satisfied.

Now we prove the sufficiency of (6.5). If (6.5) is satisfied for user i, then there exists a point $t_j^0 \in (0,1)$ such that $\kappa_j(t_j^0) = 0$ and $\kappa_j(t_j) > 0, \forall t_j < t_j^0$; $\kappa_j(t_j) <$

$0, \forall t_j > t_j^0$. It can be proved that for any given $\epsilon_j^1 > 0$ and $\epsilon_j^2 > 0$ such that $\bar{\mathcal{S}}_j = [t_j^0 - \epsilon_j^1, t_j^0 + \epsilon_j^2] \subseteq [0,1]$, there exists at least one distribution $\bar{f}_j(t_j)$ defined on $\bar{\mathcal{S}}_j$ which satisfies

$$\int_{t_j \in \bar{\mathcal{S}}_j} \kappa_j(t_j) \bar{f}_j(t_j)\, dt_j = 0. \tag{D.5}$$

In this case, (D.1) can be rewritten as

$$\mathbb{E}_{t_j}\{u_i(t_i, t_j)\} = \int_{\bar{\mathcal{S}}_j} \frac{b_{ii}^2}{\sigma_2^2 + b_{ji}^2(1 - t_j)} \bar{f}_j(t_j)\, dt_j \tag{D.6}$$

which satisfies (6.3) for user i. Moreover, condition (6.4) is inherently satisfied if user j uses $\bar{f}_j(t_j)$ because $\mathbb{E}_{t_j}\{u_i(t_i, t_j)\}$ does not depend on $f_i(t_i)$ on [0,1]. Since there are infinitely many different ϵ_j^1 and ϵ_j^2, which satisfy $\epsilon_j^1 > 0$, $\epsilon_j^2 > 0$ and $[t_j^0 - \epsilon_j^1, t_j^0 + \epsilon_j^2] \subseteq [0,1]$, there must be infinitely many distributions $\bar{f}_j(t_j)$ which satisfy (D.5). Denote the set of all such $\bar{f}_j(t_j)$ as $\bar{\Delta}_{f_j}$. Since it is the same case for user j, it can be concluded that $\bar{\Delta}_{f_1}$ and $\bar{\Delta}_{f_2}$ both have infinitely many elements if (6.5) is satisfied. Moreover, any strategy profile $\{\bar{f}_1(t_1), \bar{f}_2(t_2)\}$ that satisfies $\bar{f}_1(t_1) \in \bar{\Delta}_{f_1}$ and $\bar{f}_2(t_2) \in \bar{\Delta}_{f_2}$ constitutes an MSNE. Therefore, the game has infinitely many MSNE upon the satisfaction of (6.5).

D.2 Proof of Theorem 6.2

Assume that the most efficient MSNE is $\{\tilde{f}_1(t_1), \tilde{f}_2(t_2)\}$ and the support of $\tilde{f}_i(t_i)$ is $\tilde{\mathcal{S}}_i$. From Theorem 6.1, it can be seen that $\tilde{f}_j(t_j)$ is the distribution which maximizes $\int_{t_j \in \mathcal{S}_j} \hat{\kappa}_j(t_j) f(t_j)\, dt_j$ among all distributions subject to (D.3), where

$$\hat{\kappa}_j(t_j) = \frac{b_{ii}^2}{\sigma_2^2 + b_{ji}^2(1 - t_j)} \tag{D.7}$$

is a strictly convex and increasing function on [0,1]. Denote $\check{\kappa}_j(t_j) = \kappa_j(t_j) + \hat{\kappa}_j(t_j)$, then

$$\check{\kappa}_j(t_j) = \frac{b_{ii}^1}{\sigma_1^2 + b_{ji}^1 t_j} \tag{D.8}$$

and $\check{\kappa}_j(t_j)$ is a strictly convex and decreasing function on [0,1]. Then (D.3) can be rewritten as

$$\int_{t_j \in \mathcal{S}_j} \hat{\kappa}_j(t_j) f_j(t_j)\, dt_j = \int_{t_j \in \mathcal{S}_j} \check{\kappa}_j(t_j) f_j(t_j)\, dt_j. \tag{D.9}$$

Therefore, among all distributions the distribution $\tilde{f}_j(t_j)$ maximizes $\int_{\mathcal{S}_j} \hat{\kappa}_j(t_j) f(t_j)\, dt_j$ and $\int_{\mathcal{S}_j} \check{\kappa}_j(t_j) f(t_j)\, dt_j$ simultaneously subject to the condition

$$\int_{t_j \in \tilde{\mathcal{S}}_j} \hat{\kappa}_j(t_j) \tilde{f}_j(t_j)\, dt_j = \int_{t_j \in \tilde{\mathcal{S}}_j} \check{\kappa}_j(t_j) \tilde{f}_j(t_j)\, dt_j. \tag{D.10}$$

First, we prove that $\tilde{\mathcal{S}}_j \subseteq \{0,1\}, \forall i$. Assume that there exists $t_j^{'}$ such that $0 < t_j^{'} < 1$, $t_j^{'} \in \tilde{\mathcal{S}}_j$ and $\tilde{f}_j(t_j)$ defined on $\tilde{\mathcal{S}}_j$ maximizes $\int_{\mathcal{S}_j} \hat{\kappa}_j(t_j) f(t_j)\, dt_j$ among all possible $f_j(t_j)$ and satisfies (D.10). Since both $\hat{\kappa}_j(t_j)$ and $\check{\kappa}_j(t_j)$ are strictly convex, we can write that

$$\begin{aligned}
\int_{t_j \in \tilde{\mathcal{S}}_j} \hat{\kappa}_j(t_j) \tilde{f}_j(t_j)\, dt_j < & \int_{t_j \in \tilde{\mathcal{S}}_j / \{t_j^{'}\}} \hat{\kappa}_j(t_j) \tilde{f}_j(t_j)\, dt_j \\
& + (1 - t_j^{'}) \tilde{f}_j(t_j^{'}) \hat{\kappa}_j(0) + t_j^{'} \tilde{f}_j(t_j^{'}) \hat{\kappa}_j(1) \qquad \text{(D.11)} \\
\int_{t_j \in \tilde{\mathcal{S}}_j} \check{\kappa}_j(t_j) \tilde{f}_j(t_j)\, dt_j < & \int_{t_j \in \tilde{\mathcal{S}}_j / \{t_j^{'}\}} \check{\kappa}_j(t_j) \tilde{f}_j(t_j)\, dt_j \\
& + (1 - t_j^{'}) \tilde{f}_j(t_j^{'}) \check{\kappa}_j(0) + t_j^{'} \tilde{f}_j(t_j^{'}) \check{\kappa}_j(1). \qquad \text{(D.12)}
\end{aligned}$$

The above inequalities imply that both the left-hand side and the right-hand side of (D.10) can be increased by setting $\tilde{f}_j(t_j^{'}) = 0$ and transferring the probability densities $(1 - t_j^{'}) \tilde{f}_j(t_j^{'})$ and $t_j^{'} \tilde{f}_j(t_j^{'})$ to $t_j = 0$ and $t_j = 1$, respectively. Let $t \in [0,1]$ and denote the increases on the left-hand sides of (D.11) and (D.12) via transferring the probability densities $(1-t) \tilde{f}_j(t_j^{'})$ and $t \tilde{f}_j(t_j^{'})$ to $t_j = 0$ and $t_j = 1$ as $\hat{\delta}_j(t)$ and $\check{\delta}_j(t)$, respectively. If $\hat{\delta}_j(t_j^{'}) = \check{\delta}_j(t_j^{'})$, then (D.10) is still satisfied after the above transferring of probability densities. Note that $\hat{\kappa}_j(t_j)$ is strictly increasing and $\check{\kappa}_j(t_j)$ is strictly decreasing on $[0,1]$. Therefore, if $\hat{\delta}_j(t_j^{'}) > \check{\delta}_j(t_j^{'})$, then there exist $\epsilon > 0$ and $\dot{t}_j \in [t_j^{'} - \epsilon, t_j^{'})$ such that $\dot{t}_j \in (0, t_j^{'})$ and $\hat{\delta}_j(\dot{t}_j) = \check{\delta}_j(\dot{t}_j) > 0$. Similarly, if $\hat{\delta}_j(t_j^{'}) < \check{\delta}_j(t_j^{'})$, then there exist $\epsilon^{'} > 0$ and $\ddot{t}_j \in (t_j^{'}, t_j^{'} + \epsilon^{'}]$ such that $\ddot{t}_j \in (t_j^{'}, 1)$ and $\hat{\delta}_j(\ddot{t}_j) = \check{\delta}_j(\ddot{t}_j) > 0$. In any of the above three cases, (D.10) can be satisfied and at the same time both sides of (D.10) can be increased. Thus, $\tilde{f}_j(t_j)$ defined on any $\tilde{\mathcal{S}}_j$ that includes $t_j^{'} \in (0,1)$ cannot be the distribution which maximizes $\int_{\mathcal{S}_j} \hat{\kappa}_j(t_j) f(t_j)\, dt_j$ subject to (D.10). Therefore, $\tilde{\mathcal{S}}_j \subseteq \{0,1\}$. It is the same for $\tilde{\mathcal{S}}_i$.

Second, assume that $\tilde{f}_j(t_j) = \xi_j \delta(t_j) + (1 - \xi_j) \delta(t_j - 1)$ where $0 \leq \xi_j \leq 1$.

Then

$$\int_{t_j \in \tilde{\mathcal{S}}_j} \hat{\kappa}_j(t_j)\tilde{f}_j(t_j)\,dt_j = \xi_j \frac{b_{ii}^2}{\sigma_2^2 + b_{ji}^2} + (1-\xi_j)\frac{b_{ii}^2}{\sigma_2^2} \tag{D.13}$$

$$\int_{t_j \in \tilde{\mathcal{S}}_j} \check{\kappa}_j(t_j)\tilde{f}_j(t_j)\,dt_j = \xi_j \frac{b_{ii}^1}{\sigma_1^2} + (1-\xi_j)\frac{b_{ii}^1}{\sigma_1^2 + b_{ji}^1}. \tag{D.14}$$

Using the condition (D.10), ξ_j can be derived as in (6.7).

D.3 Proof of Theorem 6.3

Denote $\Omega_N = \{1, \ldots, N\}$ as the set of all channels and define $\Phi_i^0 = \{k \in \Omega_N | \exists l \in \Omega_N \neq k : \nu_i^1(k) \leq \nu_i^2(l)\}$. If $\Phi_i^0 \neq \emptyset$, the first iteration of Step 2 of the algorithm deletes Φ_1^0 from $\Delta_{i=1}$ and increases $\nu_2^2(k)$ to b_{22}^k/σ_k^2, $\forall k \in \Phi_1^0$. In the first iteration of Step 3, in consequence, the set of channels not satisfying the inequalities $\nu_2^1(k) > \nu_2^2(l), \forall l \in \Delta_{i=2} \neq k$ for user 2 can be potentially extended to $\Phi_2^1 = \Phi_2^0 + \bar{\Phi}_2^1$, where $\bar{\Phi}_2^1$ denotes the extra set of channels which do not satisfy the above inequalities due to the deletion of Φ_1^0 from $\Delta_{i=1}$ in Step 2. The deletion of Φ_2^1 from $\Delta_{i=2}$ in Step 3 could break the inequalities of $\nu_1^1(k) > \nu_1^2(l), \forall l \in \Delta_{i=1} \neq k$ on certain channels in $\Delta_{i=1}$ (which has been updated in Step 2) and the process potentially repeats as Step 2 and Step 3 iterate. Denote $\bar{\Phi}_i^q, q \geq 1$ as the set of channels which do not satisfy the aforementioned inequalities for user i due to the deletion of $\bar{\Phi}_j^{q-1}$ (if $q \neq 1$) or Φ_j^0 (if $q = 1$) from Δ_j in the preceding step. Note that $\bar{\Phi}_i^q = \emptyset$ if $\bar{\Phi}_j^{q-1} = \emptyset$. According to the definition of Φ_i^0 and $\bar{\Phi}_i^q$, it follows that $\Phi_1^2 = \bar{\Phi}_1^1 + \bar{\Phi}_1^2$, and iteratively $\Phi_i^q = \bar{\Phi}_i^{q-1} + \bar{\Phi}_i^q, q = 2, 4, \ldots, q^{\max}$ for $i = 1$ and $q = 1, 3, \ldots, q^{\max} + 1$ for $i = 2$. Here $q^{\max} = \min\big(r | r \in \{0, 2, 4, \ldots, N\}, \Phi_2^{r+1} = \emptyset\big)$.

Proof of i) At any iteration of the algorithm, if $\hat{k} \in \Phi_i^q$, then $\exists p \neq \hat{k} \in \Omega_N$ such that $\nu_i^2(p) \geq \nu_i^2(\hat{k})$. Otherwise there exists l such that $\nu_i^1(\hat{k}) \leq \nu_i^2(l)$ and $\nu_i^2(l) < \nu_i^2(\hat{k})$. In consequence, it leads to $\nu_i^2(\hat{k}) > \nu_i^1(\hat{k})$ which is impossible. Thus, deleting any $\hat{k} \in \Phi_i^q$ will not change $\max\limits_{k \in \Delta_i} \nu_i^2(k)$. Therefore, the result of checking the inequalities $\nu_i^1(k) > \nu_i^2(l), \forall l \in \Delta_i \neq k$ for any other channel, i.e. for $k \neq \hat{k}$, will not be affected. In conclusion, the ordering of channels is irrelevant to the result of the algorithm.

Now consider the ordering of users. When the algorithm starts from user 1, the

sequences of deletions are $\Phi_1^0, \bar{\Phi}_1^1 + \bar{\Phi}_1^2, \bar{\Phi}_1^3 + \bar{\Phi}_1^4, \ldots, \bar{\Phi}_1^{q^{\max}-1} + \bar{\Phi}_1^{q^{\max}}$ for user 1 and $\Phi_2^0 + \bar{\Phi}_2^1, \bar{\Phi}_2^2 + \bar{\Phi}_2^3, \ldots, \bar{\Phi}_2^{q^{\max}} + \bar{\Phi}_2^{q^{\max}+1}$ for user 2 through all iterations of Step 2 and Step 3, respectively. Here $\bar{\Phi}_2^{q^{\max}} + \bar{\Phi}_2^{q^{\max}+1} = \Phi_2^{q^{\max}+1} = \emptyset$ according to the definitions of Φ_i^q and $q^{\max}$. In this case, the outputs of the algorithm are $\Delta_{i=1}^1 = \Omega_N - \cup_q \Phi_1^q, q = 2, 4, \ldots, q^{\max}$ and $\Delta_{i=2}^1 = \Omega_N - \cup_q \Phi_2^q, q = 1, 3, \ldots, q^{\max} + 1$. If the ordering of users is changed or, equivalently, if the algorithm starts from user 2, the sequences of deletions change to $\Phi_2^0, \bar{\Phi}_2^1 + \bar{\Phi}_2^2, \bar{\Phi}_2^3 + \bar{\Phi}_2^4, \ldots, \bar{\Phi}_2^{q^{\max}-1} + \bar{\Phi}_2^{q^{\max}}, \bar{\Phi}_2^{q^{\max}+1}$ for user 2 and $\Phi_1^0 + \bar{\Phi}_1^1, \bar{\Phi}_1^2 + \bar{\Phi}_1^3, \ldots, \bar{\Phi}_1^{q^{\max}} + \bar{\Phi}_1^{q^{\max}+1}$ for user 1, respectively. Here $\bar{\Phi}_1^{q^{\max}+1} = \emptyset$ since $\bar{\Phi}_2^{q^{\max}} = \emptyset$, while $\bar{\Phi}_2^{q^{\max}} = \emptyset$ because $\Phi_2^{q^{\max}+1} = \bar{\Phi}_2^{q^{\max}} + \bar{\Phi}_2^{q^{\max}+1} = \emptyset$. Note that $\Phi_i^0, \forall i$ and $\bar{\Phi}_i^q, \forall i, \forall q$ keep unchanged regardless of the ordering of users according to their definitions. The outputs of the algorithm in this case are $\Delta_{i=2}^2 = \Omega_N - \cup_q \Phi_2^q, q = 2, 4, \ldots, q^{\max}$ and $\Delta_{i=1}^2 = \Omega_N - \cup_q \Phi_1^q, q = 1, 3, \ldots, q^{\max}+1$. Using the facts that $\Phi_2^{q^{\max}+1} = \emptyset$ and $\bar{\Phi}_1^{q^{\max}+1} = \emptyset$, it can be shown that $\Delta_{i=1}^1 = \Delta_{i=1}^2 = \Omega_N - \cup_{s=1}^{s=q^{\max}} \bar{\Phi}_1^s - \Phi_1^0$ and $\Delta_{i=2}^1 = \Delta_{i=2}^2 = \Omega_N - \cup_{s=1}^{s=q^{\max}-1} \bar{\Phi}_2^s - \Phi_2^0$ if $q^{\max} \geq 2$ and $\Delta_{i=1}^1 = \Delta_{i=1}^2 = \Omega_N - \Phi_1^0$ and $\Delta_{i=2}^1 = \Delta_{i=2}^2 = \Omega_N$ if $q^{\max} = 0$. Therefore, the ordering of users is irrelevant.

Proof of ii) According to the algorithm and the definition of Γ_i, $t_j^k = 0, \forall k \in \Gamma_i$. In the algorithm, $t_j^k = 0$ occurs together with setting $\nu_i^2(k) = b_{ii}^k/\sigma_k^2$ at all times. Thus, $\nu_i^2(k) = \nu_i^1(k), \forall k \in \Gamma_i$. Meanwhile, the inequalities $\nu_i^1(k) > \nu_i^2(l), \forall l \in \Delta_i \neq k$ must be satisfied $\forall k \in \Delta_i$ for user i at the output of the algorithm. If $L(\Gamma_i) \geq 2$, then there exist $\hat{l}, \check{l} (\hat{l} \neq \check{l})$ such that the inequalities $\nu_i^1(\hat{l}) > \nu_i^2(\check{l}) = \nu_i^1(\check{l})$ and $\nu_i^1(\check{l}) > \nu_i^2(\hat{l}) = \nu_i^1(\hat{l})$ are satisfied at the same time, which is impossible. Thus $L(\Gamma_i) \leq 1$.

Proof of iii) It can be shown that the channel indexes removed from $\Delta_i = 1$ and $\Delta_i = 2$ in Steps 2 and 3 correspond to the channels which must not be used for user 1 and user 2, respectively, in any MSNE. It can also be shown that one MSNE, in which both users end up allocating $P_i^{\max}$ on one channel in the output Δ_i exits, if $L(\Delta_{i=1}) = 1$ or $L(\Delta_{i=2}) = 1$. Given the above two facts, it follows that a unique MSNE exists if $L(\Delta_{i=1}) = 1$ or $L(\Delta_{i=2}) = 1$. It proves the sufficiency of the uniqueness condition of MSNE and the necessity of the condition for the existence of infinitely many MSNE at the same time. Now assume that $L(\Delta_i) > 1, \forall i$. Denote

$L_i = L(\Delta_i)$, $T_i = \Delta_i(L_i)$ and $\tilde{\Delta}_i = \{\Delta_i(1), \ldots, \Delta_i(L_i - 1)\}$. Then $E_{\mathbf{t}_j}\{u_i(\mathbf{t}_i, \mathbf{t}_j)\}$ at the output of the algorithm, denoted as $E_{\mathbf{t}_j}\{u_i\}$, can be written as

$$\begin{aligned} E_{\mathbf{t}_j}\{u_i\} &= \int_{\mathcal{S}_j} \left(\sum_{k \in \tilde{\Delta}_i} \frac{b_{ii}^k t_i^k}{\sigma_k^2 + b_{ji}^k t_j^k} + \frac{b_{ii}^{T_i}(1 - \sum_{k \in \tilde{\Delta}_i} t_i^k)}{\sigma_{T_i}^2 + b_{ji}^{T_i}(\zeta - \sum_{k \in \tilde{\Delta}_i} t_j^k)} \right) f_j(\mathbf{t}_j) d\mathbf{t}_j \\ &= \int_{\mathcal{S}_j} \sum_{k \in \tilde{\Delta}_i} t_i^k \iota^k f_j(\mathbf{t}_j) d\mathbf{t}_j + \int_{\mathcal{S}_j} \frac{b_{ii}^{T_i}}{\sigma_{T_i}^2 + b_{ji}^{T_i}(\zeta - \sum_{k \in \tilde{\Delta}_i} t_j^k)} f_j(\mathbf{t}_j) d\mathbf{t}_j \end{aligned} \tag{D.15}$$

where $\zeta = 1 - \sum_{k \in \Gamma_j} t_j^k$ is the total power that user j allocates on the channels represented by the indexes in Δ_i and

$$\iota^k = \frac{b_{ii}^k}{\sigma_k^2 + b_{ji}^k t_j^k} - \frac{b_{ii}^{T_i}}{\sigma_{T_i}^2 + b_{ji}^{T_i}(\zeta - \sum_{k \in \tilde{\Delta}_i} t_j^k)}. \tag{D.16}$$

In order to satisfy (6.9), it is required in this game that $\int_{\mathcal{S}_j} \sum_{k \in \tilde{\Delta}_i} t_i^k \iota^k f_j(\mathbf{t}_j) = 0$. The minimum of $\sum_{k \in \tilde{\Delta}_i} t_i^k \iota^k$ as a function of $t_j^k, \forall k \in \tilde{\Delta}_i$ is $\min\left(t_i^k \iota^k, \forall k \in \tilde{\Delta}_i|_{t_j^k = 1}\right)$. Denote $\Upsilon = \{k \,|\, k \in \Delta_i \cap \Delta_j\}$, then Υ is nonempty given that $L(\Delta_i) > 1$ as assumed, and $L(\Gamma_i) \leq 1$ as proved in the proof of statement ii). It can be shown that $\min(\{) t_i^k \iota^k, \forall k \in \tilde{\Delta}_i|_{t_j^k = 1}\} < 0$ if $t_j^k = 1, k \in \Upsilon$ since $\iota^k|_{t_j^k = 1} < 0$ and $t_i^k > 0, \forall k \in \tilde{\Delta}_i$. Moreover, if $T_i = \Delta_i(L_i) \in \Upsilon$, which can always be satisfied since the elements of Δ_i can be ordered in any manner with no effect on anything else, then it holds that $\lim_{t_j^k \to 0, \forall k \neq T_i} t_i^k \iota^k > 0, \forall k \in \tilde{\Delta}_i$. It follows that the sets $\Lambda_j^1 = \{\mathbf{t}_j | t_j^k = 0, \forall k \notin \Delta_j, \sum_{k \in \Delta_j} t_j^k = 1, \text{and } \sum_{k \in \tilde{\Delta}_i} t_i^k \iota^k < 0\}$ and $\Lambda_j^2 = \{\mathbf{t}_j | t_j^k = 0, \forall k \notin \Delta_j, \sum_{k \in \Delta_j} t_j^k = 1, \text{and } \sum_{k \in \tilde{\Delta}_i} t_i^k \iota^k > 0\}$ are both nonempty. Then similar to the proof for Theorem 6.1, it can be shown that there exist infinitely many $f_j(\mathbf{t}_j)$, each of which satisfies $\int_{\mathcal{S}_j} \sum_{k \in \tilde{\Delta}_i} t_i^k \iota^k f_j(\mathbf{t}_j) = 0$. Moreover, similar to the proof for Theorem 6.1, condition (6.10) is inherently satisfied upon the satisfaction of (6.9).

$\sim$