Resource Allocation for Cooperative Transmission in Wireless Network

NG, Cho Yiu

A Thesis Submitted in Partial Fulfilment
of the Requirements for the Degree of
Doctor of Philosophy
in
Information Engineering

The Chinese University of Hong Kong
August 2010

Abstract of thesis entitled:

Resource Allocation for Cooperative Transmission in Wireless Network

Submitted by NG, Cho Yiu

for the degree of Doctor of Philosophy

at The Chinese University of Hong Kong in Aug 2010

In this thesis, different codes and resource allocation algorithms for cooperative transmissions are proposed. Briefly speaking, in cooperative transmission, a number of wireless nodes form a coalition in which they exchange and cooperatively transmit messages. As a result, the order of diversity can be increased without installing additional antennas.

Firstly, *cooperative orthogonal-division channel* is defined and two cooperative transmission schemes based on dirty-paper coding and superposition code are proposed and compared through simulations. Simulation Results show the significant improvement over the pure direct transmission schemes. Although one cooperative transmission scheme achieves a slightly larger rate region, the other scheme has a much simpler implementation so the remaining parts of the thesis focus on this scheme. The outage performance of this scheme is also compared with a simplified Han-Kobayashi scheme through simulations. Simulation results illustrate the significant improvement in the diversity gain of this scheme over the Han-Kobayashi scheme.

Next, a weighted sum rate maximization algorithm is proposed. There are two purposes of this algorithm. Firstly, this algorithm is adopted to find the Pareto-optimal points of the boundary of the achievable rate region through simulations. Secondly, this algorithm can be extended to solve the max-min fairness problem and the joint utility maximization algorithm by the proposed framework.

After that, the cooperative transmission scheme is extended for the scenario of more than

two source-destination pairs. One objective is to investigate the relationship between the diversity order and the number of source-destination pairs. This is done by considering the sum power minimization problem. A pricing game is derived to provide a distributed implementation. At Nash Equilibrium of the game, the total transmission power is minimized. Simulation results show the rapid convergence of the game and its adaptation to channel fluctuations. It also shows that the cooperative transmission scheme achieves full diversity order.

However, it is noted that the complexity of implementing superposition code, which is a building block of the cooperative transmission code, is very high when there are many users in the network. Hence, another time-division multiplex (TDM) based cooperative transmission scheme is proposed. Similar to the superposition code based scheme, there is a pricing game which can provide a distributed sum power minimization. Simulation results also show that the game has high convergence rate and it can adapt to changes of channel conditions efficiently. In addition, this cooperative transmission scheme also achieves full diversity order.

Apart from replacing a superposition code based cooperative transmission scheme by a TDM based scheme, the implementation can be simplified by introducing a partner selection scheme to the nodes. In that network, the cooperative transmission code still uses superposition code as the building block. Instead of relaying the messages from all other nodes, in this new scheme, the source nodes only relay the messages for their assigned partners. A natural question is: How can we assign the partners to the source nodes such that the total transmission power is minimized. The problem is solved in two phases. Firstly, we solve the sum power minimization problem for each pair of nodes. In some cases, this problem has closed-form solutions while for the other cases, a simple iterative algorithm can solve this problem.

With this information, we can assign the partners by Gabow's algorithm, which solves the maximum weighted matching problem that is mapped from the original partner selection problem. Nonetheless, it is noted that when the number of users is very large, it involves a large amount of the communication and computational cost to solve the sum power minimization problem for each pair of nodes as well as the partner selection problem. Therefore, the *Grouping Algorithm* is proposed to reduce the aforementioned implementation cost. Simulation results show that the optimal algorithm and the Grouping Algorithm can achieve full diversity order. Moreover, although the Grouping Algorithm is sub-optimal in general, it costs only 1dB of the sum power more than the optimal algorithm.

This thesis is ended with some future research directions.

博士論文題目為《在無線網絡中協作式傳送的資源分配》的摘要：
作者：吳祖堯
所屬院校：香港中文大學

本論文將探討不同的協作式傳送碼和資源分配算法。簡言之，在協作式傳送中，不同的無線節點組成一個群組，並彼此交換與協調地傳送訊息。結果，我們無須安裝額外的天線也能提高多樣階數。

首先，「協作式正交分段頻道」會被定義。在這個頻道中，兩個分別基於髒紙編碼和疊複碼的協作式傳送方法會被討論和用電腦模擬比較。電腦模擬結果顯示這兩個方法比純直接傳送有顯著的改善。雖然其中一個方法可以得到稍為大的傳輸率區域，但是另一個方法比較容易實踐，所以本論文會集中討論後者。電腦模擬亦比較這個方法和一個被簡法的 Han-Kobayashi 方法的中斷率。結果是本論文建議的方法能達到更顯著的多樣階數。

然後，一個加權總傳輸率算法被建議。這個算法有兩個目的。第一，電腦模擬可以採用這個算法來找出在傳輸率區域邊緣的帕累托最優點。第二，它也可以被本論提及的架構擴展來解決最大最小公平性問題和聯合效用最大化問題。

跟著，這個協作式傳送方法被擴展成更多的傳送源和目的地。這個擴展的其中一個目的是探討多樣階數和傳送源數目的關係。本論文以總功率最小化問題來探討這個關係。一個價格博弈和分散式實踐法被推論來解決這個問題。在這個價格博弈的納什均衡中，總能量會被最小化。電腦模擬結果顯示這個方法有高的聚合率和能夠適應頻度上的變化。此外，這個傳送方法能達到最大的多階數。

不過，當網路中有很多用戶，疊複碼的複雜性便會很高。所以，本論文也提出另一個基於時分覆用的協作式傳送方法。這個方法也有一個價格博弈把總功率最小化。電腦模擬結果也顯示這個博弈有很高的聚合率和對頻道的適應性。此外，它也能達到最大的多樣階數。

另一個減低複雜性的方法就是採用夥伴選擇法。我們依舊使用疊複碼作為協作式傳送碼的構成要素，但是每一個傳送節點只會它的夥伴節點轉達它的訊息。我們很自然地會有的問題是：「如何為這些節點分配夥伴以最小化這個網路的總功率？」這個問題可分成兩部份來解決。首先，我們要計算如果每一對節點夥伴起來，它們的最小總功率是多少。在某些情況下，可以用一個封閉解來計算出來。其他的情況下，可以用一個簡單的迭代算法來計算出來。

接著，我們可以用 Gabow 算法來分配夥伴。不過，當用戶的數目很大時，這會涉及很多的通訊和運算成本來解決這個總功率最小化問題。所以，本論文也提出「分組算法」來減少上述的成本。電腦模擬顯示最優化算法和分組算法皆能達到最大多樣階數。雖然分組算法不保證有最優化解，但是它的總功率只是多了 1dB。

本論文的結尾是未來的研究方向。

Contents

List of Figures

List of Tables

Acknowledgement

Behind my works in this thesis, there have been many people supporting me in various manner. I would like to take this opportunity to thank them.

I am greatly indebted to my supervisor, Professor Tat Ming Lok, for his continuous guidance and support. I do appreciate his encouragement and patience throughout my postgraduate studies. His advices have not only stimulated me a much deeper thinking of my research topic, but they have also made my research life more enjoyable. It is grateful to have a supervisor, who has treated me as a friend and helped me walking through the valleys of my postgraduate journey.

I would like to thank Professor Tan Foon Wong for his insightful advices during our collaborations. He has provided me a lot of hints how to improve my work on theoretical, protocol-design and performance evaluation issues. I express my gratitude to Professor Chi Wan Sung and Dr. Kenneth Shum. I have enjoyed and learnt a lot from their discussions about multi-user information theory and convex analysis. I also thank Professor Sidharth Jaggi. He has explained a lot of complicated concepts in network coding and information theory in a much simpler way. I have improved my analytical and mathematical skills during our discussions.

I am heartily thankful to the Sir Edward Youde Memorial Fund Council. The Sir Edward Youde Memorial Fellowship is not only a financial support, but also an encouragement to me as well as an acknowledgement of the value of my research to the society. I would like to show my gratitude to my colleagues in Radio Communication Laboratory. They are Mr. Billy Tam, Ms. Mandy Zheng, Mr. Kenan Zhou, Mr. Leo Lam and Mr. Wengang Mi. I enjoy our discussions on research and other leisure things. Also, the mutual encouragements are great supports during the difficult times in the research. I am grateful to have a friend, Mr. Ho Ting Cheng, who always updates me with interesting research literature. Also, his comments and encouragements have played an important role in my whole research life.

I owe my deepest gratitude to my parents and my brother. Their endless love and concerns are warm supports in every moment of my research. I wholeheartedly thank for the prayers and encouragements of my brothers and sisters in both Evangelical Community Church in Hong Kong and Valley Chinese Baptist Church in Arizona. Their prayers have brought the necessary power and strength for my research.

Finally and the most importantly, I have to thank my heavenly Father. He understands my needs and weaknesses. Instead of bringing an extremely tough or extremely smooth path for me, He has tailor-made a perfect training for me in this research journey. This is not only a training for better research skills, but also the determination to face difficulties.

Chapter 1

Introduction

1.1 Motivation

In wireless communications, fading is an important factor which limits the capacity of a point-to-point link. One of the well-known solutions is to introduce spatial diversity by installing multiple transmit or receive antennas. As a result, there are a higher number of transmit-receive pairs of antennas with independent fading so that the outage probability is significant reduced. This is the rationale behind multiple-input multiple-output (MIMO) systems [29].

In order to have statistically independent fading for the transmit-receive pairs, there is a minimum spacing requirement between the antennas. However, in many practical scenarios, we can only have a limited number of antennas installed in each terminal. For example, firefighters and soldiers require their mobile devices to be portable. Therefore, mobile devices cannot be too large so we can only have a limited number of antennas. This limits the achievable order of spatial diversity. This motivates us to design new techniques to increase the spatial diversity order without installing extra antennas. Cooperative transmission is one example.

Due to the broadcast nature of the wireless transmission medium, the transmitting nodes can listen and help forwarding the messages to the intended destination nodes. As a result, a set of additional paths are generated for each source-destination pair. These paths experience independent fading. This explains how cooperative transmission can further increase the order of spatial diversity.

Coding schemes and the protocols between the transmitting nodes are necessary and important components for practical cooperative transmission systems. This is the motivation behind the works in this thesis.

1.2 Overview of Contributions

To begin with, I consider cooperative transmission between two source and destination so that the achievable rate region is easier to be visualized. Two cooperative transmission schemes are proposed and compared through simulations. Although one scheme achieves a slightly larger rate region, the coding and resource allocation schemes for the other cooperative transmission scheme are much simpler. Therefore, only the latter cooperative transmission scheme is considered in the remaining parts of the thesis.

The weighted sum rate maximization algorithm for this cooperative transmission scheme is proposed. Apart from utilizing it to visualize the achievable rate region through simulations, we can apply it to solve the max-min fairness problem. I also show how the weighted sum rate maximization algorithm can be extended to solve the joint utility maximization problem, which is a more general class of problems.

After that, I extend the coding scheme for the case of two source-destination pairs to more source-destination pairs. A pricing game is proposed to have a distributed implementation for sum power minimization. Simulation results show that the proposed cooperative transmission code can achieve full diversity order. Also, the convergence rate to the Nash Equilibrium is shown to be high.

The building blocks of the cooperative transmission code is superposition code. The encoding and decoding process become too complex for practical implementations when there are a large number of source-destination pairs. This motivates me to propose a time-division multiplex (TDM) based cooperative transmission scheme. Another pricing game is proposed for a distributed sum power minimization for this scheme. Although the TDM based cooperative transmission scheme consumes more transmission power than the superposition code based scheme, simulation results show that it still achieves full diversity order. Also, the convergence rate to the Nash Equilibrium is fast enough for practical use.

Apart from using a TDM based cooperative transmission scheme, I propose a partner selection scheme to alleviate the complex coding issues in superposition code based scheme. Instead of relaying the messages for all other source nodes, each source node is assigned to at most one partner source node. The partnered nodes use the superposition code based cooperative transmission scheme. The objective is also sum power minimization.

The problem is solved in two steps. Firstly, for each pair of source nodes, we compute their minimum total transmission power if they cooperate. In some cases, the closed-form solutions of rate allocation of a given pair of partnered nodes are provided. For the remaining cases, a simple but efficient iterative algorithm is proposed to optimize the rate allocation.

Then, the partner selection problem is mapped to the maximum weighted matching problem in graph theory, which can be solved with polynomial time complexity. For large number of node pairs, I propose the grouping algorithm to reduce the computational and communication overhead of the partner selection algorithm. Simulation results show that the reduction of the overhead only costs at most 1dB of the total transmission power.

1.3 Outline of Thesis

Some preliminary knowledge of this thesis is discussed in Chapter 2. In Chapter 3, some literature about relay channel, cooperative transmission and game theoretical research in wireless networks is surveyed. In Chapter 4, two cooperative transmission schemes for two source-destination pairs are proposed. The resource allocation algorithms of one cooperative transmission scheme are also investigated in this chapter. In Chapter 5, superposition coding based and TDM based cooperative transmission schemes for more than two source-destination pairs are considered. The pricing games of the two cooperative transmission schemes proposed in that chapter are also provided. In Chapter 6, the partner selection problem is studied. The thesis is ended by a summary of contributions and some proposed future research directions in Chapter 7.

Chapter 2

Preliminaries

In this chapter, there is a brief overview of some preliminary concepts which are applied in this thesis. It begins with an introduction of *Gaussian broadcast channels*, which are the building blocks of the cooperative transmission schemes proposed in this thesis, and some important properties we use in the succeeding chapters. Then, the concept of *best-reply potential game* and its convergence to pure strategy Nash Equilibrium are introduced. This chapter is ended with some useful optimization theorems.

2.1 Gaussian Broadcast Channel

2.1.1 Single Channel Case

A *Gaussian broadcast channel* is an information-theoretic generalization of a single-cell downlink channel. To illustrate the main idea, we firstly consider a time invariant channel with no inter-symbol interference (ISI). The transmitter would like to send a distinct message to each of the K users. Let $M_i \in \{1, 2, \ldots, 2^{R_i}\}$ be the message for user i. The transmitter encodes the messages $\mathbf{M} = (M_1, M_2, \ldots, M_K)$ to a complex codeword $\{x(\mathbf{M})[1], x(\mathbf{M})[2], \ldots, x(\mathbf{M})[n]\}$. Let $y_i[j]$ be the received symbols of user i. The received symbols are given by

$$y_i[j] = h_i x[j] + w_i[j], \quad i = 1, 2, \ldots, K, j = 1, 2, \ldots, n \tag{2.1}$$

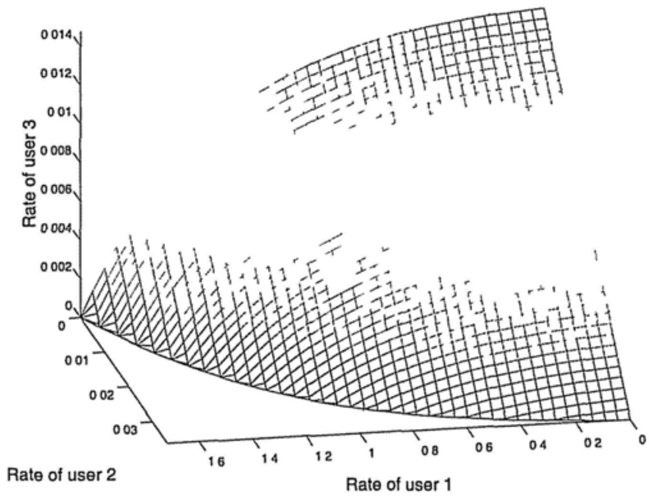

Figure 2 1 Capacity Region of a 3-user Broadcast Channel.

where h_i is the channel gain from the transmitter to user i and $w_i[j]$ is the Gaussian white noise process at user i with zero mean and variance σ_i^2. The codewords have to satisfy the following power constraint.

$$\frac{1}{n}\sum_{i=1}^{n}|x(\mathbf{M})[i]|^2 \leq P, \quad \forall \mathbf{M}. \tag{2.2}$$

User i decodes the received codeword to a message $\tilde{M}_i$. Decoding error is defined to be the event that there exists an i such that $1 \leq i \leq K$ and $\tilde{M}_i \neq M_i$. The rate vector $(R_1, R_2, \ldots, R_K)$ is *achievable* if for any $\epsilon > 0$, there exists a sufficiently large n such that the decoding error probability is upper bounded by ϵ The *capacity region* of this degraded Gaussian broadcast channel is the closure of the set achievable rate vectors.

Without loss of generality, suppose $\frac{\sigma_1^2}{|h_1|^2} < \frac{\sigma_2^2}{|h_2|^2} < \ldots < \frac{\sigma_K^2}{|h_K|^2}$. In [18, Section 14.6], it is shown that the capacity region is given by

$$0 \leq R_i \leq \log\left(1 + \frac{|h_i|^2 p_i}{\sum_{j=1}^{i-1}|h_i|^2 p_j + \sigma_i^2}\right), \quad i = 1, 2, \ldots, K. \tag{2.3}$$

As an illustration, the capacity region of a 3-user broadcast channel is depicted in Fig. 2 1

In [9], it is shown that the achievable rate region by time-division multiple access (TDMA) and frequency division multiple access (FDMA) is a proper subset of the capacity region of the degraded Gaussian broadcast channel. It shows that the time sharing of any two rate vectors of the capacity region must result a rate vector within the interior of the capacity region. This implies that the capacity region is strictly convex. The strict convexity is crucial for the convergence of the proposed algorithms in this thesis.

2.1.2 Coding Schemes

There are several different encoding schemes which can achieve the above capacity region. In this chapter, I will briefly introduce two codes, namely, *superposition coding* and *dirty paper coding*, which will be applied in the cooperative transmission schemes outlined in the succeeding chapters. Detailed descriptions and discussions can be found in [8] and [111].

Firstly, the rate vectors on the boundary can be achieved by superposition code at the transmitter and *successive interference cancelation* at each receiver. We firstly encode each user' s message separately and then superpose the codewords together. In (2.3), the rate of user i suffers from the interference from the users with higher power gain and the background noise. We encode the message of that user at the same rate as in a single-user Gaussian channel with the noise power equal to the interference-and-noise power in this case. This is called *superposition coding.*

To decode its message, user i decodes the messages of the users with smaller power gain first. This is possible because user i has better signal-to-noise ratio (SNR) for that message than the user with smaller power gain. Therefore, user i can decode the message successfully. After decoding all these messages, it cancels the interference in the received signal and decodes its own message. This is called *successive interference cancelation.*

Alternatively, we can encode the messages by dirty paper coding which does not require the receivers to know the message from the "weaker" users before decoding his/her own message bits. The decoders do not even need to know the codebooks of other users. The advantage is to provide greater flexibility in the design of the decoding method.

Table 2.1: Steps for nested lattice coding

1: Map M_3 to a codeword $x_3(M_3, U_3)[j]$, using the dither U_3 as one of the inputs.
2: Interpret $x_3(M_3, U_3)[j]$ as interference known non-causally, M_2 is dirty-paper coded on top of $x_3(M_3, U_3)[j]$, and we obtain $x_2(M_2, M_3, U_2, U_3)[j]$.
3: By treating $x_2(M_2, M_3, U_2, U_3)[j]$ and $x_3(M_3, U_3)[j]$ as non-causally known interference, we apply dirty-paper coding to M_1 and obtain the codeword $x_1(M_1, M_2, M_3, U_1, U_2, U_3)[j]$.

An efficient implementation of dirty paper coding, called *nested lattice coding* [112] is discussed in this chapter. I illustrate the idea by the following three-user example. Suppose $|h_1|^2 > |h_2|^2 > |h_3|^2$, so that user 1 has the best channel condition and user 3 is the worst user. The transmitter signal of length n is the sum of three signals

$$x[j] = x_1[j] + x_2[j] + x_3[j] \tag{2.4}$$

with the power of x_i equal to equal to $\alpha_i P$, for $i = 1, 2, 3$ and non-negative real numbers α_1, α_2 and α_3 which sum to one. The encoding requires three pseudo-random "dithers", denoted by U_1, U_2 and U_3, as inputs. For $i = 1, 2, 3$, the dither U_i is a shared randomness between the transmitter and user i, and is obtained by a pseudo-random generator. To encode the three independent messages, which are M_1, M_2 and M_3, we follow the three steps in Table 2.1.

As noted in [112, Section VI], although $x_2[j]$ is functionally dependent on $x_3[j]$, $x_2[j]$ is statistically independent of $x_3[j]$, due to the presence of the dither. Also, $x_1[j]$ is statistically independent of $x_2[j]$ and $x_3[j]$.

2.1.3 Parallel Channel Case

In the previous sub-section, all the messages are transmitted over a single channel. The results for the parallel channel case is provided in this sub-section. Also, the maximum weighted sum rate algorithm over parallel Gaussian broadcast channel is briefly introduced here. This algorithm is one of the component in the algorithms in Chapter 4. For further details, please refer to [44, 101].

Consider a parallel broadcast channel with N parallel channels. One typical example is an orthogonal frequency-division multiple access (OFDMA) system with N subcarriers. Let $n_i^{(j)}$ and $h_i^{(j)}$ be the noise power and the link gain from the transmitter to user i in channel j respectively. We denote the vectors $(n_1^{(j)}, n_2^{(j)}, \ldots, n_K^{(j)})$ and $(|h_1^{(j)}|^2, |h_2^{(j)}|^2, \ldots, |h_K^{(j)}|^2)$ by $\mathbf{n}^{(j)}$ and $\mathbf{g}^{(j)}$ respectively. Let $C_b(\mathbf{n}, \mathbf{g}, P)$ be the capacity region of a degraded Gaussian broadcast channel with noise power vector $\mathbf{n}$, power gain vector $\mathbf{g}$ and transmission power P. If the total transmission power is $\bar{P}$, the capacity region is given by [101]

$$C(\bar{P}) = \bigcup_{P_j : \sum_{j=1}^{N} P_j = \bar{P}} \sum_{j=1}^{N} C_b(\mathbf{n}^{(j)}, \mathbf{g}^{(j)}, P_j) \tag{2.5}$$

where for two sets A and B, $A + B \equiv \{\mathbf{u} + \mathbf{v} : \mathbf{u} \in A, \mathbf{v} \in B\}$. That means, to achieve any Pareto-optimal rate vector of the capacity region, we only need to allocate the total power to each channel and perform the superposition coding and successive interference cancelation described in the previous subsection separately for each channel. Notice that the capacity region of a parallel broadcast channel is also strictly convex.

In Chapter 4, we need an algorithm for the following weighted sum rate maximization problem in parallel Gaussian broadcast channel.

$$\max_{0 \le p_i^{(j)} \le \bar{P}} \sum_{i=1}^{K} \mu_i R_i \tag{2.6}$$

subject to

$$(R_1, R_2, \ldots, R_K) \in C(\bar{P}) \tag{2.7}$$

where $p_i^{(j)}$ is the power allocated to user i in channel j and μ_i's are given positive weights. According to [101, Theorem 3.2], the algorithm is outlined as follows. Let

$$u_i^{(j)}(z) = \frac{\mu_i}{2\left(\frac{n_i^{(j)}}{|h_i^{(j)}|^2} + z\right)} - \lambda, \quad 1 \le i \le K, 1 \le j \le N \tag{2.8}$$

$$u^{(j)*}(z) = \left[\max_{1 \le i \le K} u_i^{(j)}(z)\right]^+, \quad 1 \le j \le N \tag{2.9}$$

where $[x]^+ = \max\{x, 0\}$ and λ is the solution of the following equation

$$\sum_{j=1}^{N}\left[\max_{1\le i\le K}\left(\frac{\mu_i}{\lambda} - \frac{n_i^{(j)}}{|h_i^{(j)}|^2}\right)\right]^+ = \bar{P} \tag{2.10}$$

which can be easily solved by bisection method.

Also, let

$$\mathcal{A}_i^{(j)} = \left\{z : [0,\infty) : u_i^{(j)}(z) = u^{(j)*}(z)\right\}, \quad 1 \le j \le N. \tag{2.11}$$

The optimal rate and power allocated to user i at channel j are given by

$$R_i^{(j)*} = \int_{\mathcal{A}_i^{(j)}} \frac{1}{2\left(\frac{n_i^{(j)}}{|h_i^{(j)}|^2} + z\right)} dz \tag{2.12}$$

$$P_i^{(j)*} = \left|\mathcal{A}_i^{(j)}\right|. \tag{2.13}$$

The set $\mathcal{A}_i^{(j)}$ can be conveniently computed by the fact (see the proof of [101, Theorem 3.2]) that for $i \neq k$, $u_i^{(j)}(z)$ and $u_k^{(j)}(z)$ only intersect at one point and $u_i^{(j)}(z)$'s are monotonically decreasing functions of z. That is, $u_i^{(j)}(z) > u_k^{(j)}(z)$ on one side of the intersection point while $u_i^{(j)}(z) < u_k^{(j)}(z)$ on the other side of the intersection point. Therefore, after computing all the intersection points, we can obtain the intervals of z which $u_i^{(j)}(z) > u_k(j)(z)$ for all $k \neq i$. This is exactly the set $\mathcal{A}_i^{(j)}$.

2.2 Best-Reply Potential Games

Best-reply potential game is firstly introduced in [103]. In this chapter, we focus on the convergence of best-reply potential game to *pure strategy Nash Equilibria*. The detailed proof can be found in [47]. A brief summary is provided in this section.

Consider a game $G = [\mathcal{P}, \{\mathcal{R}_i\}, \{u_i\}]$ of M players. $\mathcal{P}$ is the player set and the players are indexed by the integers $1, 2, \ldots, M$. Each player i has its strategy set $\mathcal{R}_i$ and a utility function $u_i(R_i, \mathbf{R}_{-i})$, where R_i is the strategy of player i and $\mathbf{R}_{-i} = (R_1, R_2, \ldots, R_{i-1}, R_{i+1}, \ldots, R_M)$ is the vector of strategies of other players. This utility function denotes the degree of satisfaction of a particular player so player i aims to maximize u_i. Then, we have the following definition of *pure strategy Nash Equilibrium*.

Definition 2.1. *A strategy* $(R_1^*, R_2^*, \ldots, R_M^*)$ *is a* pure strategy Nash Equilibrium *if*

$$u_i(R_i, \mathbf{R}_{-i}^*) \le u_i(R_i^*, \mathbf{R}_{-i}^*), \quad i = 1, 2, \ldots, M, R_i \in \mathcal{R}_i. \tag{2.14}$$

Let $B_i(\mathbf{R}_{-i})$ be the *best response update* by player i, which is defined as

$$B_i(\mathbf{R}_{-i}) = \arg\max_{R_i \in \mathcal{R}_i} u_i(R_i, \mathbf{R}_{-i}). \tag{2.15}$$

Here, we assume that B_i is a well-defined function of $\mathbf{R}_{-i}$. That is, for all $\mathbf{R}_{-i}$, there is a unique R_i which maximizes u_i. In addition, a pure strategy Nash Equilibrium is a fixed point of the best response updates.

Below are the definitions of *potential game* and *best-reply potential game.*

Definition 2.2. *Let* $\mathbf{R} = (R_1, R_2, \ldots, R_M)$. *A game* $G = [\mathcal{P}, \{\mathcal{R}_i\}, \{u_i\}]$ *is a* potential game *if there exists a potential function* $\Phi(\mathbf{R})$ *such that*

$$\Phi(R_i^{(1)}, \mathbf{R}_{-i}) - \Phi(R_i^{(2)}, \mathbf{R}_{-i}) = u_i(R_i^{(1)}, \mathbf{R}_{-i}) - u_i(R_i^{(2)}, \mathbf{R}_{-i}), \quad R_i^{(1)}, R_i^{(2)} \in \mathcal{R}_i, 1 \le i \le M. \tag{2.16}$$

Definition 2.3. *A potential game* $G = [\mathcal{P}, \{\mathcal{R}_i\}, \{u_i\}]$ *with potential function* Φ *is a* best-reply potential game *if*

$$B_i(\mathbf{R}_{-i}) = \arg\max_{R_i \in \mathcal{R}_i} \Phi(R_i, \mathbf{R}_{-i}), \quad 1 \le i \le M. \tag{2.17}$$

In this chapter, we focus on the convergence to pure strategy Nash Equilibria of a best-reply potential game. To begin with, we have the definition of a *sequential best-reply path* as follows.

Definition 2.4. *A sequence of strategies* $(R_1^t, R_2^t, \ldots, R_M^t)_{t=0}^{\infty}$ *is a* sequential best-reply path *if for all* t, *there exists a set* $\tilde{P}^t \subseteq \mathcal{P}$ *such that*

$$R_i^t = \begin{cases} B_i(\mathbf{R}_{-i}^{t-1}), & \text{if } i \in \tilde{P}^t \\ R_i^{t-1}, & \text{otherwise.} \end{cases} \tag{2.18}$$

Notice that it is possible to have $R_i^t = R_i^{t-1}$ for some $i \in \tilde{P}^t$ in a sequential best-reply path. That is, after the best response update, the strategy of player i does not change. Among all sequential best-reply paths, in most practical situations, we are only interested in *admissible sequential best-reply paths* which are defined as follows.

Definition 2.5. *A sequential best-reply path is* admissible *if for all M successive periods, all players have performed their best response updates (even if the best response update is an identity function) at least once.*

In Chapter 5, we need the following theorem which is proved in [47, Theorem 2].

Theorem 2.1. *Suppose in a best-reply potential game $G = [\mathcal{P}, \{\mathcal{R}_i\}, \{u_i\}]$, $\mathcal{R}_i$'s are all compact, the best response updates $B_i(\mathbf{R}_{-i})$'s are all continuous functions and there is a unique pure strategy Nash Equilibrium. Then, any admissible sequential best-reply path converges to the unique pure strategy Nash Equilibrium.*

2.3 Theorems about Optimization Problems

In this section, some frequently used theorems related to optimization in this thesis are provided. Due to the page limitation, I skip the proofs of these theorems.

I begin with the *Envelope's Theorem.* It is useful to compute the first derivative of a function which is expressed as a parameterized optimization problem. One example of such kind of functions is the dual function of an optimization problem.

Theorem 2.2 (Envelope's Theorem [28]). *Let*

$$f^*(\mathbf{r}) = \max_{\mathbf{x}} f(\mathbf{x}, \mathbf{r}). \tag{2.19}$$

Assume the above optimization problem has unique optimal solution. Also, let

$$\mathbf{x}^*(\mathbf{r}) = \arg\max_{\mathbf{x}} f(\mathbf{x}, \mathbf{r}). \tag{2.20}$$

Then, we have

$$\frac{df^*(\mathbf{r})}{dr_i} = \left.\frac{\partial f(\mathbf{x}, \mathbf{r})}{\partial r_i}\right|_{\mathbf{x}=\mathbf{x}^*(\mathbf{r})}. \tag{2.21}$$

Envelope's Theorem can be generalized to the *Danskin's Theorem*. More specifically, Danskin's Theorem allows the cases that the parameterized optimization problems have more than one optimal solutions.

Theorem 2.3 (Danskin's Theorem [21]). *Let Z be a compact subset of $\mathbb{R}^m$, and let $\phi : \mathbb{R}^n \times Z \to \mathbb{R}$ be continuous and such that $\phi(\cdot, z) : \mathbb{R}^n \to \mathbb{R}$ is convex for each $z \in Z$.*

1. *The function $f : \mathbb{R}^n \to \mathbb{R}$ given by*

$$f(x) = \max_{z \in Z} \phi(x, z) \tag{2.22}$$

is convex and has directional derivative given by

$$f'(x; y) = \max_{z \in Z(x)} \phi'(x, z; y), \tag{2.23}$$

where $\phi'(x, z; y)$ is the directional derivative of the function $\phi(\cdot, z)$ at x in the direction y, and $Z(x)$ is the set of maximizing points in (2.22)

$$Z(x) = \left\{ \bar{z} \,\middle|\, \phi(x, \bar{z}) = \max_{z \in Z} \phi(x, z) \right\}. \tag{2.24}$$

In particular, if $Z(x)$ consists of a unique point $\bar{z}$ and $\phi(\cdot, \bar{z})$ is differentiable at x, then f is differentiable at x, and $\nabla f(x) = \nabla_x \phi(x, \bar{z})$, where $\nabla_x \phi(x, \bar{z})$ is the vector with components

$$\frac{\partial \phi(x, \bar{z})}{\partial x_i}, \quad i = 1, 2, \ldots, n. \tag{2.25}$$

2. *If $\phi(\cdot, z)$ is differentiable for all $z \in Z$ and $\nabla_x \phi(x, \cdot)$ is continuous on Z for each x, then*

$$\partial f(x) = conv\{\nabla_x \phi(x, z) \,|\, z \in Z(x)\}, \quad \forall x \in \mathbb{R}^n. \tag{2.26}$$

Next, we have the following simplified version of Berge's *Maximum Theorem* in [7, p.116]. It is used to determine whether a function, which can again be expressed as an optimization problem, is continuous.

Theorem 2.4 (Special Case of Berge's Maximum Theorem). *Using the notation as in Danskin's theorem, if $\mathcal{Y} \subset \mathbb{R}^M$ is a compact set, then both $\phi(\mathbf{x}, \mathbf{z}(\mathbf{x}))$ and $\mathbf{z}(\mathbf{x})$ are continuous functions of $\mathbf{x}$.*

Finally, we have the convergence results of *blocked Gauss-Seidel algorithm* [33]. In Gauss-Seidel algorithm [10], we optimize each decision variable by fixing the values of other variables in a round robin manner. In blocked Gauss-Seidel algorithm, we optimize each disjoint subset of decision variables in a round robin manner. Although both methods provide a convenient means to derive distributed algorithms, they may not converge to the optimal solution in general even when the optimization problem is convex (see [79, p.53-54] for the explanations).

Consider the following optimization problem:

$$\min f(\mathbf{x}) \tag{2.27}$$

subject to

$$\mathbf{x} \in \mathcal{X} = \mathcal{X}_1 \times \mathcal{X}_2 \times \ldots \times \mathcal{X}_m \subseteq \mathbf{R}^n \tag{2.28}$$

where

1. $f : \mathbf{R}^n \to \mathbf{R}$ is a continuously differentiable function,

2. every $\mathcal{X}_i \subseteq \mathbf{R}^{n_i}$ is closed, nonempty and convex,

3. $\sum_{i=1}^{m} n_i = n$ and

4. $\mathbf{x} = (\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_m) \in \mathbf{R}^n$ with $\mathbf{x}_i \in \mathcal{X}_i$ for all i.

In each block Gauss-Seidel iteration, we perform the following updates:

$$x_i^{k+1} = \arg\min_{y_i \in \mathcal{X}_i} f\left(x_1^{k+1}, x_2^{k+1}, \ldots, x_{i-1}^{k+1}, y_i, x_{i+1}^k, x_{i+2}^k, \ldots, x_m^k\right) \tag{2.29}$$

where x_i^k is the value of x_i in the k-th iteration.

We need the following definitions before stating the convergence result.

Definition 2.6. *A differentiable function f is* pseudoconvex *if it satisfies the property: $\nabla f(x)(y-x) \geq 0$ implies $f(y) \geq f(x)$.*

Note that a differentiable and convex function is a pseudoconvex function.

Definition 2.7. *The $r-$*level set *of a function f relative to $\mathcal{X}$ is the set $\{x \in \mathcal{X} : f(x) \leq r\}$.*

The sufficient conditions for convergence and the optimality of block Gauss-Seidel algorithm are stated in [33, Proposition 6] which is quoted below.

Proposition 2.5. *Suppose that f is pseudoconvex on $\mathcal{X}$ and all level sets are compact. Then, the sequence $\{x^k\}$ generated by the block Gauss-Seidel algorithm has limit points and every limit point $\bar{x}$ of $\{x^k\}$ is a global minimizer of f.*

Chapter 3

Literature Survey

In this chapter, we will review some previous works about relay channel and cooperative transmission. In this thesis, we classify relay channel and cooperative transmission as follows. In relay channel, the relaying nodes do not have their own messages to be transmitted. In cooperative transmission, the relaying nodes are also source nodes in general.

3.1 Relay Channel

The relay channel is typically defined as a channel in which there is one source node, one destination node and a number of relay nodes. The relay nodes can overhear the signals from the source node and help forwarding the messages. The most primitive form, which comprises of only one relay node, is depicted in Figure 3.1 and is thoroughly studied in [17, 24]. The authors in [17] provide the capacity bounds for the general relay channel and compute the capacity for degraded relay channels. Bounds for the capacity of a general relay channel can also be found in [24]. The author also provides a necessary and sufficient condition for a positive capacity.

In the models of the aforementioned manuscripts, the relay node is assumed to operate in full-duplex mode, i.e. it can transmit and receive at the same time and frequency. However, the transmit signal is typically 100dB higher than the received signal. This requires very precise and expensive circuitry to realize a full-duplex relay node. It motivates the studies of

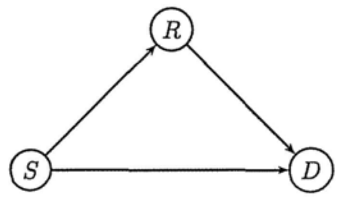

Figure 3.1: System model of relay channel with one destination node.

half-duplex relay nodes [42, 53, 66]. In these works, the relay nodes operates in time-division manner. For a given time period, the relay node is in the receive mode for a fraction of time and in the transmit mode for the remaining time. Similarly, the relay node may operate in frequency-division manner in the sense that different portion of the frequency spectrum is used for transmit and receive modes [32, 40, 55].

Amplify-and-forward (AF) and decode-and-forward (DF) are the two major types of relaying techniques [57]. In AF, the relay multiplies the received signal from the source by a complex factor and forwards it to the destination. In DF, the relay first buffers and decodes the message. Then, it re-encodes the the message and sends it to the destination node. While AF has the advantage of simple implementation and preservation of soft information, DF has the advantage of error correction at the relay nodes. The comparisons of these two approaches in practical systems are provided in [70, 110]. The authors in [70] point out that AF and DF can outperform the other depending on the underlying channel condition. If the link between the source and relay is worse than the other links, AF offers higher capacity. Otherwise, DF outperforms AF. In [110], the authors compare AF and DF with turbo codes and they discover that AF and DF are practically on par with one another.

In low SNR regime, the capacity of DF approaches is limited by the high bit error rate at the relay nodes. In AF approaches, the amplified noise becomes the major component in the relay node's transmitted signal. Scaling laws of capacity for AF approaches in low SNR and wideband regimes are studied in [104]. In [57], incremental AF approach is developed. In that approach, the destination node has feedback to the source node. In [5], bursty AF scheme is proposed and it is shown to achieve the optimal outage performance at asymptotically low

SNR regime.

All these schemes consider the cases that channel state information is not available at the nodes. Nevertheless, when the links suffer from slow fading, it is possible to estimate the link gains. In [42, 107], the authors consider cooperative protocols which make use of the channel quality information to further increase the achievable rate. Bounds of ergodic capacity are obtained in [42]. In [107], the authors propose a DF approach which can achieve the maximum diversity-multiplexing tradeoff.

3.2 Cooperative Transmission

In this thesis, I consider cooperative transmission schemes. The major difference from the above relay channel schemes is that in cooperative transmission, there are no pure relay nodes. Each source node overhears one another's message and forward the message to the intended destination.

The earliest model is the one introduced in [106]. A two-user multiple access (MAC) channel is considered. The two encoders are connected by communication links with finite capacities so that they can cooperatively encode and transmit their messages. In [6, 51, 60], the authors also consider the case that all the messages are transmitted from cooperating sources to the same destination nodes. This type of cooperative transmission schemes are also known as *cooperative MAC* (e.g. [6, 13, 35, 49, 51, 60, 99] and the references therein). In [99], a distributed CMAC scheme based on superposition coding and the corresponding sum power minimization algorithm are proposed. In the CMAC scheme considered in [99], a source node relays messages from other nodes to the (single) destination by time-sharing its link to the destination among the messages. In the network with distributed source-destination pairs considered in Chapter 5, a source node needs to relay messages from other nodes over distinct links to different destinations. This complication renders the solution approach and the distributed resource allocation algorithm in [99] inapplicable to the network of interest in that chapter. CMAC in [35] and [49] require centralized implementation. However, it is not practical to extend their approaches to our scenario when the number of source-destination pairs is large. Thus, in Chapter 5, we propose a pricing game for distributed implementation. Weighted sum rate maximization for two-user scenario is studied in [51]. In [6, 13], frame error rate minimization for two users are investigated. Fading cooperative MAC is considered in [89] where channel state information is unknown at the transmitters. In this thesis, I consider the scenario that such information is available to the transmitters. As argued in the previous section, the assumption of available channel state information is valid in some practical scenarios such as slow fading channel.

Inner and outer bounds for cooperative transmission for two source-destination pairs is studied in [41]. Based on these bounds, the author characterizes the capacity gains for transmitter cooperations and receiver cooperations at high SNR regime. In [45], the outage performance of cooperative transmission with collocated destination is investigated. The authors consider a type of cooperative transmission called *coded cooperation*. The basic idea is that each user, instead of repeating the received bits (either via amplification or decoding) tries to transmit incremental redundancy for its partner.

However, the works mentioned in the above paragraphs assume full-duplex relaying by the source nodes which is not practical. Half-duplex cooperative transmission are motivated. In [48], DF transmitter cooperations and receiver cooperations for two source nodes and two destination nodes are compared. The authors conclude that the improvement of capacity by transmitter cooperations is much more significant than receiver cooperations. Hence, in this thesis, I concentrate on transmitter cooperation schemes. In [71], two half-duplex schemes with two source nodes and two destinations are proposed and their sum capacities are studied. Each source node can transmit only for a fraction of time and listen for the remaining portion of time. In Chapter 4, each source node is allocated a disjoint set of parallel channels instead of time-sharing the same frequency band. This setting provides a more convenient node cooperations in some wireless networks like the OFDMA networks, in which each node is allocated a disjoint subset of subcarriers. Also, I study the whole achievable rate region of my cooperative transmission scheme. Furthermore, in my model in Chapter 5, there are more source and destination nodes. Sum power minimization problem in that setting is considered.

Another half-duplex cooperative transmission scheme for two source-destination pairs is considered in [93]. The authors make use of fact that a source node S_1 has the full knowledge of the message $M_{1,r}$, which is relayed by another source node S_2. Then, S_2 chooses a codebook for its message $M_{2,r}$, which will be relayed by S_1, according to $M_{1,r}$. As a result, the effect of the interfering signal of $M_{1,r}$ to $M_{2,r}$ is eliminated. However, it requires a much larger codebook for the transmitting and receiving nodes. This problem is more serious in the scenario considered in Chapter 5, which consists of more source-destination pairs. In one of the proposed scheme described in Chapter 4, it requires a relatively smaller codebook and it is more practical to be extended for larger number of source-destination pairs.

In [96], the authors perform an asymptotic analysis of the achieved diversity order of *cooperative space-time coding.* However, they heavily rely on the assumption that the inter-user channel has high power gains. In the cooperative transmission schemes proposed in this thesis, I do not make this assumption and the overall achievable diversity order is the full diversity order.

Half duplex cooperative transmission schemes for more than two source-destination pairs are considered in [23, 45, 61, 64, 83]. In [23, 45, 64], the authors focus on the frame error rate of the users under individual power constraints. In [45], the authors assume that channel state information is unavailable at the transmitters and in one of their protocols, the receivers only have the rough channel state information of the incoming links. In this thesis, I consider the case that the transmitters have the knowledge of the channel state information of their outgoing links. In the pricing game of Chapter 5, by exchanging pricing information, the transmitters can optimize the rates of their outgoing links without knowing the channel state information of other links. In [61], cooperative transmission for ad-hoc networks with hybrid ARQ is considered to maximize the total throughput. The authors in [83] aim to maximize the total throughput by cooperative transmission in random access networks. In Chapters 5 and 6, sum power minimization problem is considered.

In [57], both AF and DF half-duplex cooperative transmission schemes are proposed for the case that channel state information is unavailable. Since in [56], it is shown that the sum capacity cannot be improved in full-duplex cooperations when transmitters have no channel state information and the fading is ergodic, the authors in [57] focus on the delay-limited case and nonergodic case. Similarly, in [58], cooperative transmission protocol where channel state information is only available at the receiver nodes is considered. In the following chapters, I consider the case that the channel condition can be estimated.

In [27, 50], half-duplex cooperative transmission scheme with analog network coding is considered. However, as shown by the authors in [25], the limitations of precisions of real (or complex) number processing can increase the condition number of the network transform matrix and potentially reduce the achievable rates substantially.

Some cooperative transmission coding schemes become too complicated to be implemented when a large number of source nodes involve. This is the motivation behind some *partner selection schemes*. Briefly speaking, the source nodes choose their partners and perform the two-user cooperative transmission. In [80], the authors consider partner selection schemes to reduce the bit error performance. However, the bit error performance of the underlying cooperative transmission scheme heavily depend on the reception quality of the relay node because of the absence of the direct path transmission. The partner selection scheme proposed in Chapter 6 uses the cooperative transmission scheme in Chapter 4, which in general comprises of the direct path transmission as well as the relay path transmission.

Besides [80], partner selection problem is studied in [36, 49, 65]. In these works, the relay nodes are pure relay nodes which do not have their own messages to transmit. But I consider a more general case that no nodes are pure relay nodes in this chapter. In [35, 95], partner selection schemes for uplink cooperations are proposed but in the system model in Chapter 6, the destination nodes are not co-located.

3.3 Game Theory in Wireless Networks

Pricing games for interfering links are proposed in [15, 43, 85, 98]. Their objectives are maximizing certain utility functions of the received signal to interference plus noise ratio (SINR) subject to maximum transmission power constraints. This is in some sense "dual" to the problem in Chapter 5 that minimizes the total transmission power under the minimum rate (which is a utility function of the received SINR) constraint. Other power control games for interfering links are investigated in [3, 39, 63]. In these models, the transmitters do not relay others' messages and the receivers simply treat the interference as noise. However, when a transmit-receiver pair suffers from deep fading, the source node has to increase the transmission power significantly to compensate for the fading loss and the interference from other nodes so that the rate requirements can be met. As a result, other nodes experience greater interference and they may need to raise their transmission powers as well. In cooperative transmissions, messages are more likely to be forwarded over the paths with higher power gains so that smaller transmission power is needed.

Various games [36, 46, 67, 92] have been proposed for cooperative transmissions. However, the algorithms are designed for different objectives. In [92], the authors propose a pricing game for AF user cooperations to improve individual frame success rate per unit energy. However, the protocol cannot be applied to our case that aims to minimize the total transmission power under individual rate requirements in a network with DF user cooperations. As mentioned in previous sections, one advantage of DF user cooperations is the opportunity of error correction and signal regeneration by the relaying nodes. The protocol in [46] requires an additional Stackelberg leader node for price computation which is not needed in the protocol proposed in this thesis. In [36], the objective, which is different from ours, is to minimize the maximum transmission power under received SNR requirements. In [67], a coalition game is proposed to maximize the sum rates of a wireless cluster.

Chapter 4

Achievable Rates in Cooperative Orthogonal-Division Channel

In this chapter, I consider cooperative transmission between two source-destination pairs. Exchange of data is allowed between the two source nodes. In addition to the direct transmission link from the source to the intended destination, there is a two-hop relay path that sends the data via the neighboring source node. The bandwidth is partitioned into two parts, and each part is solely utilized by one source node, such that the transmissions from the two sources are orthogonal to each other. In this way, the interference channel is reduced to two independent broadcast channels. The bandwidth of each source node is divided into orthogonal sub-channels, and results from parallel broadcast channel is used to find the optimal allocation of power and rate to each links. I propose an iterative algorithm that maximizes the weighted sum rate. It can also be applied to obtain the boundary of the achievable rate region. The achievable rate region in low signal-to-noise ratio regime is also characterized and the rate allocation problem becomes much simpler. Finally, I extend the proposed weighted sum rate maximization algorithm to solve a more general joint utility maximization problem.

The chapter is organized as follows. In Section 4.1, I describe the system model. The two proposed transmission protocols are discussed in Section 4.2. Some mathematical properties of these two protocols are provided in the same section as well. Numerical examples are presented in Section 4.3. An algorithm that maximizes the weighted sum rate of one proposed cooperative transmission scheme is proposed in Section 4.4. An extension of this algorithm for the parallel channel case is also provided in this section. In Section 4.5, I show that when the signal-to noise ratio (SNR) is low, for instance, in wide-band systems, some simplifications are possible. One application of the weighted sum rate maximization algorithm is to achieve the max-min fairness, which is detailed in Section 4.6. In Section 4.7, I show how to extend the proposed weighted sum rate maximization algorithm to a more general joint utility maximization algorithm.

Part of the contents in this chapter can also be found in [76, 77].

4.1 System Model and Notations

Consider a wireless network with two transmitter-receiver pairs, denoted by (S_1, D_1) and (S_2, D_2). Node S_1 wants to transmit data to D_1 and node S_2 to D_2. Assume that the transmissions of S_1 and S_2 are on two orthogonal channels, each of bandwidth $B_W/2$ Hz[1], so that the total bandwidth is B_W Hz. Such orthogonality allows the source nodes to transmit their signals and overhear each other's signals simultaneously.

For $i = 1, 2$, consider the transmission of S_i. In the following, let $j = 3 - i$, so that $\{i, j\} = \{1, 2\}$. For example, if $i = 2$, then $j = 1$. The power gain of the link from S_i to S_j, D_i, and D_j by a_i, b_i, and c_i, respectively. I will call a_i, b_i and c_i the *cooperative*, *direct* and *cross* link gains, respectively. Let the two-sided power spectral density of white noise experienced at each receiver be $N_0/2$ W/Hz. In a period of T seconds, each orthogonal channel has $B_W T$ real degrees of freedom [31, p.177], when $B_W T$ is large. In the channel where S_i is transmitting, the received channel symbols at S_j, D_i and D_j at time t, for

[1] For the convenience of presentation, I consider the equal bandwidth case. The results in this chapter can be easily extended to the unequal bandwidth case.

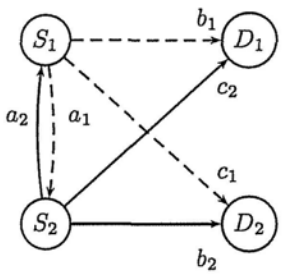

Figure 4.1: System Model. The dashed lines correspond to the channel with S_1 as the sender, and the solid lines correspond to the channel with S_2 as the sender. The labels of the arrows are the associated link gains.

$t = 1, 2, \ldots, B_W T$, are respectively

$$Y_{S_iS_j}[t] = \sqrt{a_i} X_{S_i}[t] + Z_{S_iS_j}[t] \tag{4.1}$$

$$Y_{S_iD_i}[t] = \sqrt{b_i} X_{S_i}[t] + Z_{S_iD_i}[t] \tag{4.2}$$

$$Y_{S_iD_j}[t] = \sqrt{c_i} X_{S_i}[t] + Z_{S_iD_j}[t], \tag{4.3}$$

where $X_{S_i}[t]$ is the transmitted symbol, and $Z_{S_iS_j}[t]$, $Z_{S_iD_i}[t]$ and $Z_{S_iD_j}[t]$ are additive white Gaussian noise (AWGN) with mean zero and variance $N_0/2$ (Fig. 4.1). Let P_i be the maximum transmission power of S_i, for $i = 1, 2$, i.e., the transmitted signal satisfies

$$\frac{1}{B_W T} \sum_{t=1}^{B_W T} (X_{S_i}[t])^2 \leq P_i. \tag{4.4}$$

This model is called the *cooperative orthogonal-division channel.* The transmission rate from S_i to D_i is denoted by R_i, where $i = 1, 2$. In order to simplify notations, I normalize the power such that $N_0/2 = 1$, and assume that all noise powers are equal to 1.

I will use $\mathbb{R}_+^n$ to stand for the first orthant in the n-dimensional Euclidean space, i.e., it consists of the vectors with non-negative components. Let $C(x) \triangleq 0.5 \log_2(1 + x)$ denote the Shannon capacity formula. Vectors are typeset in boldface.

4.2 Transmission Schemes

I propose two cooperative transmission schemes in this section. The two schemes have the same encoding function, and differ in the the decoding processing. The encoding at the two source nodes are the same in both schemes. For the encoding, note that the channel described by (4.1), (4.2) and (4.3) is a broadcast channel [18]. Nested lattice code or superposition code, which are described in Section 2.1, are adopted in these Gaussian broadcast channels, which are the building blocks of the proposed cooperative transmission schemes. After describing the common encoding process, I characterize the achievable rates by the two decoding methods.

4.2.1 Encoding at the Two Source Nodes

In the proposed transmission protocol for the cooperative orthogonal-division channel. For $i = 1, 2$, source node i splits its own data stream into two streams. The first stream is sent directly through the direct link between S_i and D_i. The second one is sent through a two-hop path, from S_i to the opposite source node S_j, and then from S_j to the intended destination D_i. Node S_j acts as a relay node, and re-encode the message to be forwarded. In other words, the proposed schemes can be classified as *partial decoding-and-forward* schemes; the relay node decodes and forwards only part of the message from the source node. For $i = 1, 2$, let r_{id} denote the rate of data through the direct path, and r_{ir} the rate of data through the two-hop path. Here, the subscripts "$_d$" and "$_r$" signify "direct path" and "relay path" respectively.

Time is divided into $B + 1$ time slots, and each slot contains a codeword of length L. For $i = 1, 2$, the data from S_i is divided into $2B$ parts: $b_{id}(n)$ and $b_{ir}(n)$ for $n = 1, 2, \ldots, B$. For each n, $b_{id}(n)$ consists of Lr_{id} bits, and is transmitted through the direct path from S_i to D_i; $b_{ir}(n)$ consists of Lr_{ir} bits, and is decoded and re-encoded by the opposite source node S_j.

From the viewpoint of S_1, for example, it has to transmit three data streams: the first one is direct transmission to its intended receiver D_1; the second one is transmission to its relay node, S_2; the third one is forwarding the data from S_2 to D_2. In time slot n, source S_1

Slot 1	Slot 2	Slot 3	Slot 4
$b_{1r}(1)$	$b_{1r}(2)$	$b_{1r}(3)$	
$b_{1d}(1)$	$b_{1d}(2)$	$b_{1d}(3)$	
	$b_{2r}(1)$	$b_{2r}(2)$	$b_{2r}(3)$

Slot 1	Slot 2	Slot 3	Slot 4
$b_{2r}(1)$	$b_{2r}(2)$	$b_{2r}(3)$	
$b_{2d}(1)$	$b_{2d}(2)$	$b_{2d}(3)$	
	$b_{1r}(1)$	$b_{1r}(2)$	$b_{1r}(3)$

Figure 4.2: Illustration of the encoding process ($B = 3$). The first row indicates the messages encoded by S_1, and the second row the messages by S_2.

transmits the codeword

$$\boldsymbol{x}_{S_1}(b_{1r}(n), b_{1d}(n), \hat{b}_{2r}(n-1)), \tag{4.5}$$

where $\hat{b}_{2r}(n-1)$ denotes the decoded message $b_{2r}(n-1)$ from the previous time slot. Here, $\boldsymbol{x}_{S_1}$ is a codeword as in (2.4) from nested lattice coding, with three messages as inputs (the dithers are not shown for notational simplicity). Similarly, in time slot n, source S_2 transmits

$$\boldsymbol{x}_{S_2}(b_{2r}(n), b_{2d}(n), \hat{b}_{1r}(n-1)). \tag{4.6}$$

We initialize the encoding process by setting $b_{1r}(0)$ and $b_{1r}(0)$ to some known and constant bit strings, the all-zero bit strings for instance, of length Lr_{1r} and Lr_{2r} respectively. Likewise, the encoding process terminates in time slot $B+1$, by setting $b_{ir}(B+1)$ and $b_{id}(B+1)$ to some pre-defined bit strings, for $i = 1, 2$. An illustration of the encoding for $B = 3$ is shown in Fig. 4.2.

Note that there is a loss of data rate by a factor of $B/(B+1)$, which tends to 1 as B tends to infinity. Hence, this loss of data rate is negligible when B is large.

4.2.2 Achievable Rates

Scheme 1

In the first proposed scheme, the received signals from the two orthogonal channels are decoded *separately*, using the broadcast channel decoding algorithm. The decoded data are put together to form the original message. In the sub-channel with S_1 as the source node, the rate triple (r_{1r}, r_{1d}, r_{2r}) is constrained by the capacity region of the corresponding Gaussian BC with link gains a_1, b_1 and c_1. Similarly, in the second sub-channel with S_2 as source node, the rate triple (r_{2r}, r_{2d}, r_{1r}) is limited by the Gaussian BC with link gains a_2, b_2 and c_2. I call this *Scheme 1* and characterize the achievable rates as follows.

Theorem 4.1 (Scheme 1). *The rate pair $(R_1, R_2) \in \mathbb{R}_+^2$ is achievable by separately processing the received signals from the two orthogonal channels, if*

$$R_1 = r_{1d} + r_{1r} \tag{4.7}$$

$$R_2 = r_{2d} + r_{2r} \tag{4.8}$$

$$(r_{1r}, r_{1d}, r_{2r}) \in 0.5 \cdot \mathcal{C}_{BC}(a_1, b_1, c_1, P_1) \tag{4.9}$$

$$(r_{2r}, r_{2d}, r_{1r}) \in 0.5 \cdot \mathcal{C}_{BC}(a_2, b_2, c_2, P_2). \tag{4.10}$$

(the product of a real number x and a set $\mathcal{S}$ is defined as $\{xy : y \in \mathcal{S}\}$.)

Proof. Equations (4.7) and (4.8) say that the total data rate is the sum of rates of the direct path and the relay path. The condition (4.9) and (4.10) mean that the rate vectors (r_{1r}, r_{1d}, R_{2r}) and (r_{2r}, r_{2d}, R_{1r}) are both feasible. The decoding and re-encoding at the two sources node is thus performed with arbitrarily small probability of error. Note that there is a factor of 0.5 in (4.9) and (4.10), because the total bandwidth is divided into two equal halves, one for each broadcast channel. □

I call that (R_1', R_2') *Pareto dominates* (R_1, R_2) if $R_1' \geq R_1$ and $R_2' \geq R_2$. A rate pair (R_1, R_2) is said to be *Pareto optimal* if it is not Pareto dominated by other achievable rate pairs, i.e., if (R_1', R_2') is an achievable rate pair that Pareto dominates (R_1, R_2), then

$(R_1, R_2) = (R_1', R_2')$. In the remaining of this sub-section, I will aim at characterizing the Pareto optimal rate pairs (Theorem 4.4).

Let the maximum data rate between S_1 and D_1 be denoted by $R_1^{\max}$. Formally, $R_1^{\max}$ is defined as the maximum R_1 such that (R_1, R_2) is achievable for some R_2. When R_1 is maximized, S_1 will not forward data from S_2. It is because, if on the contrary S_1 allocates some power for the relay link between S_1 and D_2, S_1 can move some power to the direct link between S_1 and D_1 and thus increase the data rate R_1. Let the maximum value of $r_{12} + R_{11}$ subject to $(r_{12}, R_{11}, 0) \in \mathcal{C}_1(P_1)$ be achieved by $(\hat{r}_{12}, \hat{R}_{11}, 0)$. I first consider the case that $(0, 0, \hat{r}_{12}) \in \mathcal{C}_2(P_2)$. Then S_2 can help S_1 by forwarding at a rate of $\hat{r}_{12}$. In the mean time, S_2 can transmit data to D_2 at a rate of $\hat{R}_{22}$, which is the maximum value of R_{22} such that $(0, R_{22}, \hat{r}_{12}) \in \mathcal{C}_2(P_2)$. In this case, the maximum R_1 equals $\hat{r}_{12} + \hat{R}_{11}$, and the vertical line segment between $(\hat{r}_{12} + \hat{R}_{11}, 0)$ and $(\hat{r}_{12} + \hat{R}_{11}, \hat{R}_{22})$ belongs to the achievable rate region. Now consider the second case that the data rate $\hat{r}_{12}$ between S_2 and D_1 is not supported by S_2, i.e., if $(0, 0, \hat{r}_{12}) \notin \mathcal{C}_2(P_2)$. In this case, I let $\hat{R}_{21}$ be the maximal data rate such that $(0, 0, \hat{R}_{21}) \in \mathcal{C}_2(P_2)$. Then $R_1^{\max}$ can be obtained by maximizing $r_{12} + R_{11}$ subject to the constraints $(r_{12}, R_{11}, 0) \in \mathcal{C}_1(P_1)$ and $r_{12} \leq \hat{R}_{21}$.

In a similar fashion, I can obtain the maximal value of R_2, which will be denoted by $R_2^{\max}$.

Before I proceed to characterize some properties of the achievable rate region, I have the following definitions.

Definition 4.1. *For a subset $S \subset \mathbb{R}_+^n$, a point $\boldsymbol{x} \in S$ is in the* relative interior, *or an* interior point, *of S if there is an open ball B in $\mathbb{R}^n$ centered at $\boldsymbol{x}$, such that $B \cap \mathbb{R}_+^n \subset S$.*

As an example, consider the region in $\mathbb{R}_+^2$ defined by $x + y \leq 2$. The points $(0, 0)$ and $(0, 1)$ are within the relative interior, but $(0, 2)$ and $(1, 1)$ are not.

Definition 4.2. *In a rate region, a point is a* boundary point, *or lies on the boundary, if it is not inside the relative interior. A subset S in $\mathbb{R}_+^n$ is said to be* strictly convex, *if for any two distinct points $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$ in S, then the linear combination $\alpha \boldsymbol{x}_1 + (1 - \alpha)\boldsymbol{x}_2$ lies within the relative interior of S for all $0 < \alpha < 1$.*

Given two non-negative weights w_1 and w_2, not both equal to zero, the problem of maximizing the weighted sum $w_1 R_1 + w_2 R_2$ over the achievable rate region will be considered in Section 4.4. Since the achievable rate region is a closed and bounded set, the maximum weighted sum by a rate pair (R_1, R_2) in the rate region exists.

The relation between Pareto optimal point, boundary point and weighted sum rate is summarized in the next proposition.

Proposition 4.2. *For a rate pair (R_1, R_2) in the achievable rate region satisfying $R_1 < R_1^{\max}$ and $R_2 < R_2^{\max}$, the followings are equivalent: (i) there exists two positive weights w_1 and w_2 such that the corresponding weighted sum is attained by (R_1, R_2), (ii) (R_1, R_2) is a boundary point and (iii) (R_1, R_2) is Pareto optimal.*

The proof is relegated to the Appendix A.1.

Proposition 4.3. *Assume all the normalized noise variances are distinct. Let (R_1, R_2) and (R_1', R_2') be two distinct Pareto optimal rate pairs in the achievable rate region. Then, for $0 < \lambda < 1$, the linear combination $\lambda(R_1, R_2) + (1-\lambda)(R_1', R_2')$ lies within the relative interior of the achievable rate region.*

Proof. I denote the achievable rate region by $\mathcal{X}$. Let $(r_{12}, R_{11}, R_{12}) \in \mathcal{C}_1(P_1)$ and $(r_{21}, R_{22}, R_{21}) \in \mathcal{C}_2(P_2)$ be the rate allocation associated with (R_1, R_2) and let $(r_{12}', R_{11}', R_{12}') \in \mathcal{C}_1(P_1)$ and $(r_{21}', R_{22}', R_{21}') \in \mathcal{C}_2(P_2)$ be the rate allocation associated with (R_1', R_2').

I claim that either $r_{12} \neq r_{12}'$ or $r_{21} \neq r_{21}'$, or both. Suppose on the contrary that $r_{12} = r_{12}'$ and $r_{21} = r_{21}'$. For fixed r_{12} and r_{21}, let $R_{11}^{\max}$ be the maximal R_{11} such that $(r_{12}, R_{11}, r_{21}) \in \mathcal{C}_1(P_1)$, and let $R_{22}^{\max}$ be the maximal R_{22} such that $(r_{21}, R_{22}, r_{12}) \in \mathcal{C}_2(P_2)$. Since both rate pairs (R_1, R_2) and (R_1', R_2') are Pareto optimal, we must have $R_{11} = R_{11}' = R_{11}^{\max}$ and $R_{22} = R_{22}' = R_{22}^{\max}$. This contradicts that the two rate pairs are distinct, and proves the claim.

In the rest of the proof, I will assume without loss of generality that $r_{12} \neq r_{12}'$.

Let $(r_{12}'', R_{11}'', R_{12}'')$ be the convex combination

$$\lambda(r_{12}, R_{11}, R_{12}) + (1-\lambda)(r_{12}', R_{11}', R_{12}') \tag{4.11}$$

and $(r''_{21}, R''_{22}, R''_{21})$ be

$$\lambda(r_{21}, R_{22}, R_{21}) + (1-\lambda)(r'_{21}, R'_{22}, R'_{21}). \tag{4.12}$$

By definition of Pareto optimality, we cannot have $R_2 = R'_2 = R_2^{\max}$, otherwise one of (R_1, R_2) and (R'_1, R'_2) will Pareto dominates the other. From our earlier discussion about the calculation of $R_2^{\max}$, at least one of r_{21} and r'_{21} is positive. This implies that $r''_{21} > 0$. Since $r_{12} \neq r'_{12}$, we also have $r''_{12} > 0$.

By our assumption that $r_{12} \neq r'_{12}$, and thus $R_{21} \neq R'_{21}$, the two rate triples (r_{12}, R_{11}, R_{12}) and $(r'_{12}, R'_{11}, R'_{12})$ are distinct in $\mathcal{C}_1(P_1)$, and the two rate triples (r_{21}, R_{22}, R_{21}) and $(r'_{21}, R'_{22}, R'_{21})$ are distinct in $\mathcal{C}_2(P_2)$. Because the normalized noise variances are distinct, both $\mathcal{C}_1(P_1)$ and $\mathcal{C}_2(P_2)$ are strictly convex [100]. As a result, I conclude that the rate triple $(r''_{12}, R''_{11}, R''_{12})$ is an interior point of $\mathcal{C}_1(P_1)$, and $(r''_{21}, R''_{22}, R''_{21})$ is an interior point of $\mathcal{C}_2(P_2)$. Because $r''_{12} > 0$ and $r''_{21} > 0$, we can find a 2-dimensional open disc $D_\epsilon = \{(x, y) : x^2 + y^2 < \epsilon^2\}$ with sufficiently small radius ϵ such that

$$(r''_{12} + x, R''_{11}, R''_{12} + y) \in \mathcal{C}_1(P_1) \tag{4.13}$$

and

$$(r''_{21} + y, R''_{22}, R''_{21} + x) \in \mathcal{C}_2(P_2) \tag{4.14}$$

for all $(x, y) \in D_\epsilon$. Note that $(r''_{12} + x, R''_{11}, R''_{12} + y)$ and $(r''_{21} + y, R''_{22}, R''_{21} + x)$ satisfy the conditions in Theorem 4.1. Consequently, $(R''_{11} + R''_{21} + x, R''_{22} + R''_{21} + y) = (R''_1 + x, R''_2 + y)$ is achievable for all $(x, y) \in D_\epsilon$. This shows that (R''_1, R''_2) is an interior point of the achievable rate region. □

We are now in a position to characterize the Pareto optimal rate allocations.

Theorem 4.4. *Assume all the normalized noise variances are distinct. Let $(\tilde{R}_1, \tilde{R}_2)$ be a Pareto optimal rate pair. Then one of the followings holds:*

(i) $\tilde{R}_1 = R_1^{\max}$.

(ii) $\tilde{R}_2 = R_2^{\max}$.

(iii) There are two positive weights w_1 and w_2 such that $(\tilde{R}_1, \tilde{R}_2)$ is the unique rate pair that maximizes the weighted sum rate $w_1 R_1 + w_2 R_2$.

Proof. I will show that if $\tilde{R}_1 < R_1^{\max}$ and $\tilde{R}_2 < R_2^{\max}$, then (iii) holds. The existence of weights w_1 and w_2 such that $w_1R_1 + w_2R_2$ is maximized at $(R_1, R_2) = (\tilde{R}_1, \tilde{R}_2)$ follows from Prop. 4.2. I only need to prove that $(\tilde{R}_1, \tilde{R}_2)$ is the unique solution to the weighted sum maximization.

Suppose that two distinct rate pairs (R_1, R_2) and (R_1', R_2') attain the maximum weighted sum rate associated with weight w_1 and w_2. By Proposition 4.2, both (R_1, R_2) and (R_1', R_2') are Pareto optimal. Then by Proposition 4.2, each convex combination $(R_1'', R_2'') = \lambda(R_1, R_2) + (1-\lambda)(R_1', R_2')$ with $0 < \lambda < 1$ is a relative interior point of the achievable rate region. Hence, there is a sufficiently small $\epsilon > 0$ such that $(R_1''+\epsilon, R_2''+\epsilon)$ is also achievable. This contradicts that the value of the weighted sum rate is $w_1R_1 + w_2R_2 = w_1R_1' + w_2R_2'$. □

In view of Theorem 4.4, the following definition is well-defined. For non-negative weights w_1 and w_2, not both zero, let $(\check{R}_1(w_1, w_2), \check{R}_2(w_1, w_2))$ be the Pareto optimal rate pair that maximizes $w_1R_1 + w_2R_2$ over the achievable rate region. When $w_1 = 0$ and $w_2 > 0$, $\check{R}_2(0, w_2) = R_2^{\max}$ and $\check{R}_1(0, w_2)$ is the largest R_1 such that $(R_1, R_2^{\max})$ is achievable. When $w_2 = 0$ and $w_1 > 0$, $\check{R}_1(w_1, 0) = R_1^{\max}$ and $\check{R}_2(w_2, 0)$ is the largest R_2 such that $(R_1^{\max}, R_2)$ is achievable. If $w_1 > 0$ and $w_2 > 0$, then $(\check{R}_1(w_1, w_2), \check{R}_2(w_1, w_2))$ is the unique rate pair that maximizes the corresponding weighted sum rate.

By varying the weights w_1 and w_2, we can trace out the boundary of the achievable rate region. An efficient iterative algorithm for computing the optimal weighted sum rate is devised in Section 4.4.

Scheme 2

In the second transmission scheme, the received signals from the two orthogonal channels are *jointly* processed in the decoding of the relayed message. For $n = 1, 2, \ldots, B$, the message $b_{1r}(n)$ is encoded by $\boldsymbol{x}_{S_1}(b_{1r}(n), b_{1d}(n), \hat{b}_{2r}(n-1))$ in the n-th time slot of the first channel, and $\boldsymbol{x}_{S_2}(b_{2r}(n+1), b_{2d}(n+1), \hat{b}_{1r}(n))$ in the $(n+1)$-st time slot of the second channel. The major difference from Scheme 1 is on the decoding function. The decoding function takes these two received signals as inputs and estimates $b_{1r}(n)$. Likewise, the decoding of $b_{2r}(n)$ is based on the received signals $\boldsymbol{x}_{S_2}(b_{2r}(n), b_{2d}(n), \hat{b}_{1r}(n-1))$ and $\boldsymbol{x}_{S_1}(b_{1r}(n+1), b_{1d}(n+1), \hat{b}_{2r}(n))$. This transmission scheme with joint processing of the signals from the two channels is called *Scheme 2*.

Since this cooperative transmission scheme is primarily effective when the link between the two source nodes are good, I will only state the rate region by Scheme 2 for the case where $a_i > b_i$ and $a_i > c_i$, for $i = 1, 2$. Let $\mathbb{I}$ be the indicator function defined by

$$\mathbb{I}(z) = \begin{cases} 1 & \text{if } z \text{ is true} \\ 0 & \text{otherwise.} \end{cases} \tag{4.15}$$

Theorem 4.5 (Scheme 2). *Suppose $a_i > b_i$ and $a_i > c_i$ for $i = 1, 2$, and assume without loss of generality that $b_i \neq c_i$, for $i = 1, 2$. The rate pair $(R_1, R_2) \in \mathbb{R}_+^2$ is achievable by Scheme 2 if it satisfies* (4.7), (4.8) *and the following conditions*

$$r_{1r} \leq \frac{1}{2} C(a_1 \alpha_1 P_1) \tag{4.16}$$

$$r_{2r} \leq \frac{1}{2} C(a_2 \beta_1 P_2) \tag{4.17}$$

$$r_{1d} \leq \frac{1}{2} C\Big(\frac{b_1 \alpha_2 P_1}{1 + b_1 P_1(\alpha_1 + \mathbb{I}(c_1 > b_1)\alpha_3)}\Big) \tag{4.18}$$

$$r_{2d} \leq \frac{1}{2} C\Big(\frac{b_2 \beta_2 P_2}{1 + b_2 P_2(\beta_1 + \mathbb{I}(c_2 > b_2)\beta_3)}\Big) \tag{4.19}$$

$$r_{1r} \leq \frac{1}{2} C(b_1 \alpha_1 P_1) + \frac{1}{2} C\Big(\frac{c_2 \beta_3 P_2}{1 + c_2 P_2(\beta_1 + \mathbb{I}(b_2 > c_2)\beta_2)}\Big) \tag{4.20}$$

$$r_{2r} \leq \frac{1}{2} C(b_2 \beta_1 P_2) + \frac{1}{2} C\Big(\frac{c_1 \alpha_3 P_1}{1 + c_1 P_1(\alpha_1 + \mathbb{I}(b_1 > c_1)\alpha_2)}\Big) \tag{4.21}$$

for some non-negative real numbers α_1, α_2, α_3, β_1, β_2, and β_3 such that $\alpha_1 + \alpha_2 + \alpha_3 = \beta_1 + \beta_2 + \beta_3 = 1$.

Proof. The first two conditions in (4.16) and (4.17) ensure that the source nodes are able to decode the relay message from the opposite source node, so that the re-encoding is error-free. The conditions in (4.18) and (4.19) guarantee that the direct part of the data, $b_{id}(n)$, for $i = 1, 2$ and $n = 1, 2, \ldots, B$, can be sent reliably to the destination through the direct link. The last two conditions are the rate constraints for the decoding of the relay part of the data, $b_{ir}(n)$.

The first terms in (4.20) is the signal-to-noise ratio (SNR) of the signal

$$b_1 \cdot x_{S_1}(b_{1r}(n), b_{1d}(n), \hat{b}_{2r}(n-1) + \mathbf{z}_{S_1 D_1}, \tag{4.22}$$

received by D_1 in the first channel, where $\mathbf{z}_{S_1 D_1}$ denotes the noise vector with each component independently distributed according to the standard normal distribution. In the second term in (4.20), the fraction represents the SNR of

$$c_2 \cdot x_{S_2}(b_{2r}(n+1), b_{2d}(n+1), \hat{b}_{1r}(n) + \mathbf{z}_{S_2 D_1}. \tag{4.23}$$

By standard argument from information theory, we see that the data rate on the right hand side of (4.20) is achievable after maximizing the mutual information between input and output. Similar comment goes with (4.21). □

Note that if R_{1d} and R_{2d} are restricted to zero, then Scheme 2 is the same as the transmission scheme in [27].

In the first proposed scheme, the decoding algorithm for the Gaussian BC is employed, and the capacity region $\mathcal{C}_{BC}$ appears in the statement of Theorem 4.1. In the second proposed scheme, in which the signals from the two orthogonal channels are jointly processed, the resulting rate region is strictly better than the one in the first scheme. However, the decoding complexity also increases accordingly.

4.3 Performance Comparison

4.3.1 Achievable Rate Region

I compare the rate region achieved by Schemes 1 and 2 with a cut-set outer bound and the capacity region of the strong interference channel.

From [54], I have the following cut-set outer bound for the achievable rates. The derivation is straightforward and omitted. I refer the readers to [54] for more details.

Proposition 4.6 (Outer Bound). *A rate pair* (R_1, R_2) *is achievable in the cooperative* *orthogonal-division channel only if it satisfies*

$$R_1 \le 0.5 \cdot C((a_1^2 + c_1^2)P_1) \tag{4.24}$$

$$R_1 \le 0.5 \cdot C(b_1^2 P_1) + 0.5C(c_2^2 P_2) \tag{4.25}$$

$$R_2 \le 0.5 \cdot C((a_2^2 + c_2^2)P_2) \tag{4.26}$$

$$R_2 \le 0.5 \cdot C(c_2^2 P_1) + 0.5C(b_2^2 P_2) \tag{4.27}$$

$$R_1 + R_2 \le 0.5 \cdot [C((b_1^2 + c_1^2)P_1) + C((b_2^2 + c_2^2)P_2)]. \tag{4.28}$$

In non-cooperative transmission scheme, the link between source nodes S_1 and S_2 is ignored, and the channel reduces to the Gaussian *interference channel* (GIC). The two source nodes can transmit simultaneously, and the received symbols at the two destination nodes are

$$Y_{D_1}[t] = \sqrt{b_1} X_{S_1}[t] + \sqrt{c_2} X_{S_2}[t] + Z_{D_1}[t] \tag{4.29}$$

$$Y_{D_2}[t] = \sqrt{b_2} X_{S_2}[t] + \sqrt{c_1} X_{S_1}[t] + Z_{D_2}[t]. \tag{4.30}$$

For fair comparison, I consider interference channel with the same bandwidth as in Scheme 1 and Scheme 2, i.e., the total bandwidth is B_W, and the same total power constraint. In a period of T seconds, the number of real degrees of freedom is $2B_WT$. To make the total power constraint the same as Scheme 1 and 2, the average power of S_i in each channel symbol is $P_i/2$, for $i = 1, 2$.

The problem of finding the capacity region for the GIC in general is currently open. However, the answer is known in some special cases, for instance, the capacity region under strong interference [86], and the optimal sum rate in the low-interference regime [4, 69, 90]. I will compare with the capacity region of the GIC in the strong interference case. Suppose that $P_1 = P_2$, and $c_i > b_i$ for $i = 1, 2$. A rate pair (R_1, R_2) is achievable in the Gaussian interference channel given in (4.29) and (4.30) if and only if

$$0 \le R_1 \le C(b_1 P_1/2) \tag{4.31}$$

$$0 \le R_2 \le C(b_2 P_2/2) \tag{4.32}$$

$$R_1 + R_2 \le \min\{C(\frac{b_1 P_1}{2} + \frac{c_1 P_2}{2}), C(\frac{b_2 P_2}{2} + \frac{c_1 P_1}{2})\}. \tag{4.33}$$

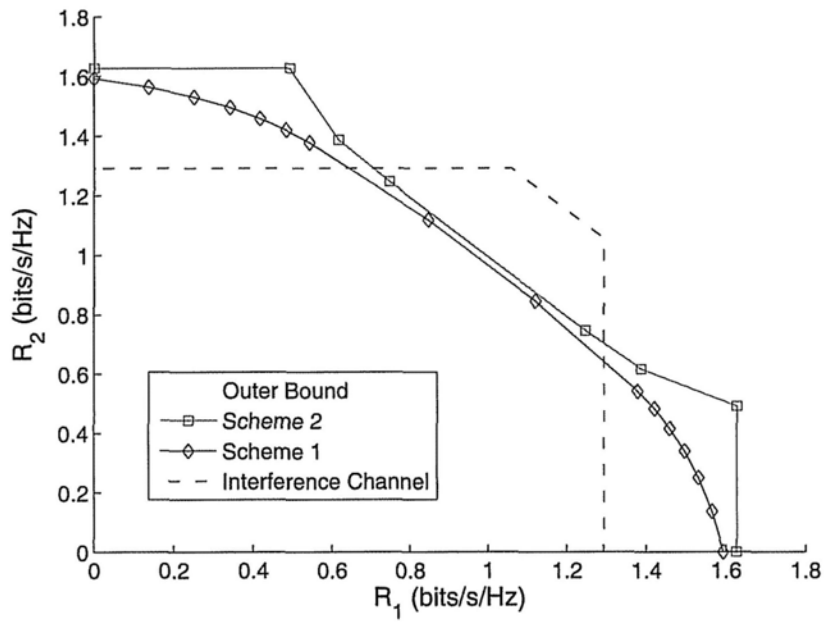

Figure 4.3: Comparison of rate regions when the cooperative link is not very strong.

It is remarked that there the Shannon formulae in (4.31) to (4.33) are not multiplied by 0.5, because the the system occupies the whole bandwidth of B_W Hz.

In Fig. 4.3, I plot the cut-set outer bound, the rate regions of the two proposed schemes, and the capacity region of the GIC, for the case with parameters $P_1 = P_2 = 10$, and $a_i = 3$, $b_i = 1$ and $c_i = 2$ for $i = 1, 2$. The points for Scheme 1 are obtained by maximizing the weighted sum $w_1 R_1 + w_2 R_2$ over the feasible rate region, for different choices of weights w_1 and w_2.

For scheme 2, an efficient algorithm which maximizes weighted sum rate is difficult to be obtained because in general, its achievable rate region is not convex. A general maximization algorithm, called the *branch-and-bound* method, is applied instead.

It can be seen from Fig. 4.3 that the capacity region of the non-cooperative GIC and the rate regions of Schemes 1 and 2 do not dominate each other; the GIC have larger sum rate, whereas two cooperative schemes can achieve better rate pairs which are more asymmetric.

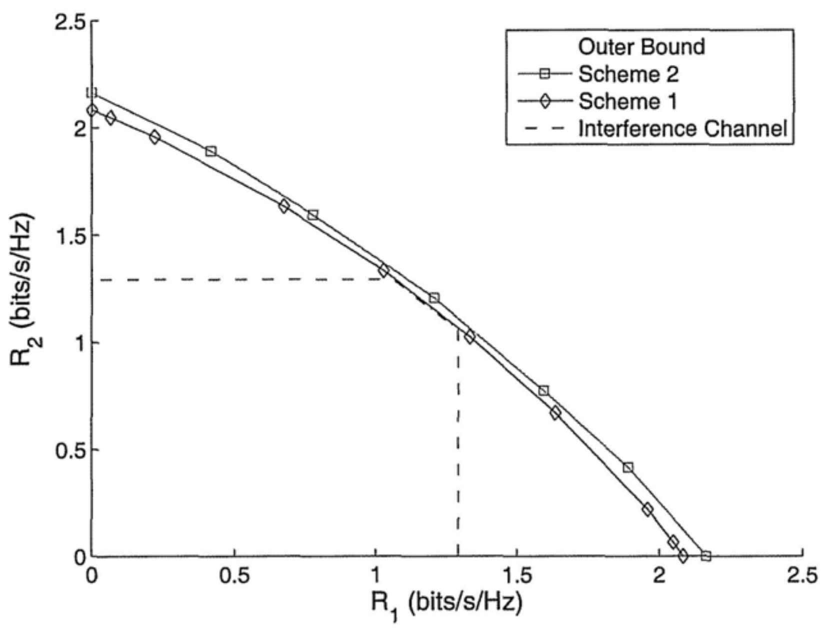

Figure 4.4: Comparison of rate regions when the cooperative link is strong.

It can also be observed that the region of Scheme 2 defined as in Theorem 4.5 is not convex in this setting. The corner points $(0.4912, 1.6269)$ and $(1.6269, 0.4912)$ in Scheme 2 are of special interest. The rate pair $(0.4912, 1.6269)$ is achieved by setting the power allocation vector $(\alpha_1, \alpha_2, \alpha_3)$ to $(0, 9/11, 2/11)$ and $(\beta_1, \beta_2, \beta_3)$ to $(1, 0, 0)$. In the first subchannel, S_1 uses $9/11$ of the total power for its own message to be sent through the direct link, and $2/11$ of the total power for forwarding S_2's through the link from S_1 to D_2. In the second subchannel, S_2 uses all of its power for the message to be relayed. There is no direct message from S_2. Time-sharing can be applied if we want to operate on a point which lies on the line segment between $(0.4912, 1.6269)$ and $(1.6269, 0.4912)$. However, this may require carrier-level synchronization among the nodes which would increase the implementation cost.

In Fig. 4.4, the cooperative link gain a_i is increased from 3 to 10, and plot the corresponding rate region. It is noted that the capacity region for the GIC remains unchanged, because it does not depend on the link gains a_1 and a_2. However, the rate regions for Scheme 1 and Scheme 2 become larger, and include the capacity region of the GIC as a subset. It is in accordance with the heuristics that transmitter-cooperation is effective when the link between the two source nodes is good. It can also be observed that the gap between Scheme 1 and Scheme 2 is small.

4.3.2 Outage Performance

Now, I consider the outage performance. Due to the expensive computational cost for Scheme 2, I only compare Scheme 1 and a simplified Han-Kobayashi scheme proposed in [26] for interference channel through simulations. According to [26, Corollary 1], the achievable rate region contains all rate pairs (R_1, R_2) satisfying

$$R_i \leq 2C\left(1+\frac{b_iP_i}{2}\right)-1, \quad i=1,2 \tag{4.34}$$

$$R_1+R_2 \leq \log_2\left(\frac{2c_1P_1+b_1P_1}{2}\right)+2C\left(\frac{2+b_2P_2}{c_1P_1}\right)-2 \tag{4.35}$$

$$R_1+R_2 \leq \log_2\left(\frac{2c_2P_2+b_2P_2}{2}\right)+2C\left(\frac{2+b_1P_1}{c_2P_2}\right)-2 \tag{4.36}$$

$$R_1+R_2 \leq 2C\left(\frac{c_2P_2}{2}+\frac{b_1}{c_1}\right)+2C\left(\frac{c_1P_1}{2}+\frac{b_2}{c_2}\right)-2 \tag{4.37}$$

$$2R_1+R_2 \leq 2C\left(\frac{b_1P_1+c_2P_2}{2}\right)+2C\left(\frac{c_1P_1}{2}+\frac{b_2}{c_2}\right)+2C\left(1+\frac{b_1}{c_1}\right)-3 \tag{4.38}$$

$$R_1+2R_2 \leq 2C\left(\frac{b_2P_2+c_1P_1}{2}\right)+2C\left(\frac{c_2P_2}{2}+\frac{b_1}{c_1}\right)+2C\left(1+\frac{b_2}{c_2}\right)-3 \tag{4.39}$$

The purpose is to compare the achievable order of diversity of these two schemes. Outage is defined as the event that the sum rate $R_1 + R_2$ is less than a given required sum rate R_T. In our simulations, R_T is chosen to be 1 and 2.

The power gains of all the links are independently and exponentially distributed with mean 1, which corresponds to the Rayleigh fading case. For simplicity, the transmission power of both source nodes are the same. I plot the outage probabilities of Scheme 1 and the simplified Han-Kobayashi scheme against the normalized SNR of each source node. If Γ is the common transmit SNR of the source nodes, the normalized SNR $\bar{\Gamma}$ is given by

$$\bar{\Gamma} = \frac{\Gamma}{2^{\frac{R_T}{2}} - 1}. \tag{4.40}$$

The denominator is the minimum required SNR of a source node at rate $\frac{R_T}{2}$ in the no fading case (i.e. an AWGN channel with power gain 1). The normalized SNR can be regarded as the additional amount of power (in dB) of each source node to combat against fading.

Let $\Phi_C(\bar{\Gamma})$ be the outage probability of Scheme 1 when the normalized transmit SNR is $\bar{\Gamma}$ respectively. I define the *sum rate diversity order* of Scheme 1 as follows. If there exists a real number r such that

$$\lim_{\bar{\Gamma}\to\infty} \Phi_C(\bar{\Gamma})\bar{\Gamma}^k = r, \tag{4.41}$$

the sum rate diversity order of Scheme 1 is said to be k. If the sum rate diversity order of Scheme 1 is k, in the log-log plot of the outage probability against the normalized SNR, the slope of the curve of Scheme 1 is $-k$ for sufficiently large normalized SNR. The sum rate diversity order of the simplified Han-Kobayashi scheme is defined in a similar manner.

The results are plotted in Fig. 4.5 and 4.6. Both figures are obtained from 1,000,000 trials of Monte Carlo simulations. In both cases, the outage probability of Scheme 1 is strictly smaller than the simplified Han-Kobayashi scheme. Besides that, when the normalized SNR is greater than 15dB, the slope of the curves for Scheme 1 is roughly -4 while the slope of the curves for the simplified Han-Kobayashi scheme is -2. The sum rate diversity order of the simplified Han-Kobayashi scheme is 2 because there is a multi-user diversity among the two source-destination pairs. One reason why the sum rate diversity order of Scheme 1 is twice of the one of the simplified Han-Kobayashi scheme is that in an interference channel, the achievable rate of a source-destination pair is limited by the power gain of their direct link. If the direct link is in deep fading, there is no alternative path for the message to be transmitted. On the contrary, in Scheme 1, if the direct link is in deep fading, part of the messages can be relayed by another source node, which is a new path with independent fading. Therefore, the sum rate diversity order of Scheme 1 is twice of the one of the simplified Han-Kobayashi scheme.

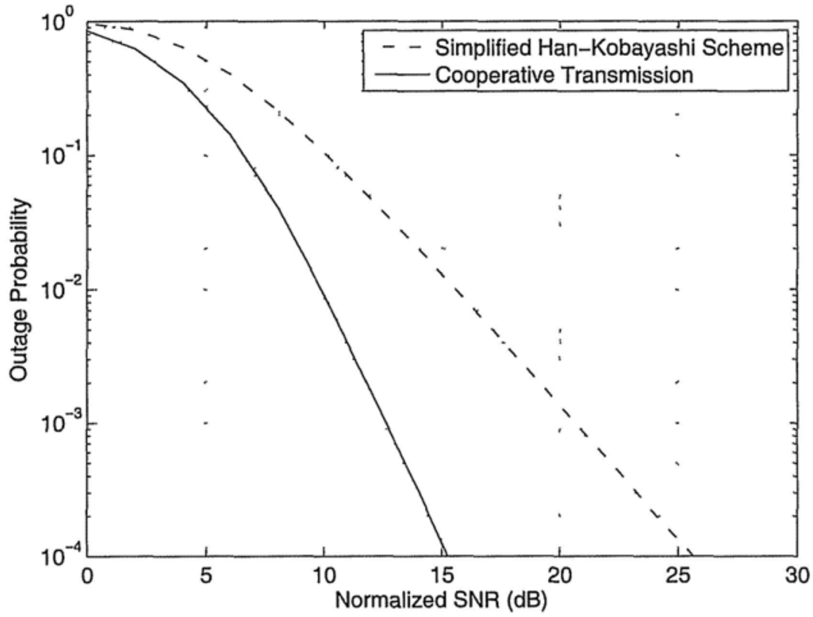

Figure 4.5: Outage probability of Scheme 1 and the simplified Han-Kobayashi scheme ($R_T = 1$).

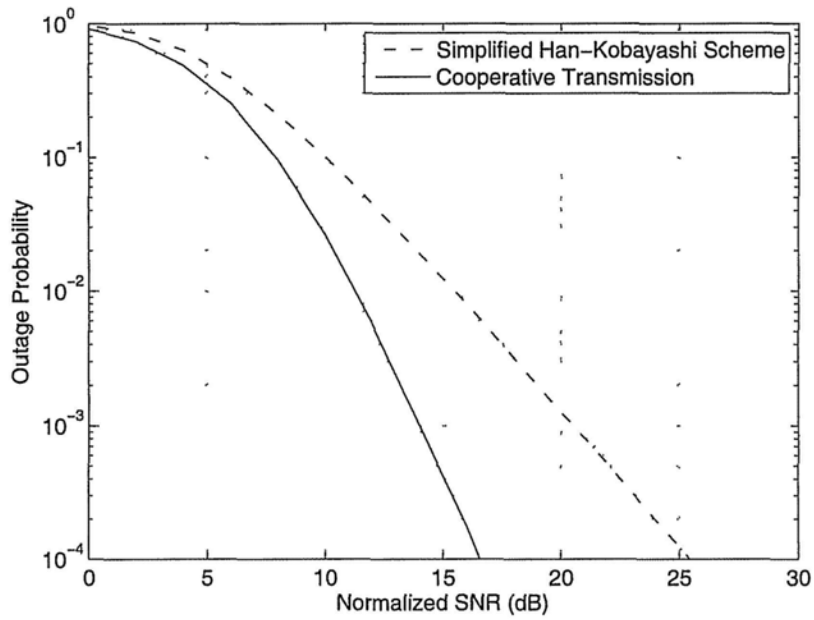

Figure 4.6: Outage probability of Scheme 1 and the simplified Han-Kobayashi scheme ($R_T = 2$).

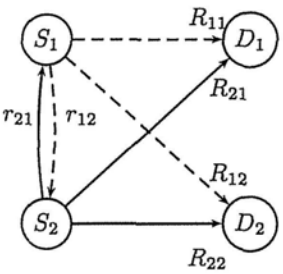

Figure 4.7: Data rates in the System Model

4.4 Iterative Algorithm for Maximizing Weighted Sum Rate

In the last section, it is shown via numerical examples that the difference between Scheme 1 and Scheme 2 is small. Since Scheme 1 has much smaller decoder complexity, for the remaining parts of the thesis, Scheme 1 and its extensions are considered. In this section, a fast weighted sum rate maximization algorithm is derived.

4.4.1 An Iterative Algorithm Based on Lagrangian

To determine the optimal weighted sum rate for given weights w_1 and w_2, an iterative algorithm based on Lagrangian duality is proposed. The rate constraints in Theorem 4.1 are slightly modified and the weighted sum maximization is formulated as follows:

$$\text{maximize } w_1 R_1 + w_2 R_2 \tag{4.42}$$

subject to

$$R_1 = R_{11} + R_{21} \tag{4.43}$$

$$R_2 = R_{22} + R_{12} \tag{4.44}$$

$$(r_{12}, R_{11}, R_{12}) \in 0.5 \cdot \mathcal{C}_{BC}(a_1, b_1, c_1, P_1) \tag{4.45}$$

$$(r_{21}, R_{22}, R_{21}) \in 0.5 \cdot \mathcal{C}_{BC}(a_2, b_2, c_2, P_2) \tag{4.46}$$

$$r_{12} = R_{21} \tag{4.47}$$

$$r_{21} = R_{12}. \tag{4.48}$$

The notations are illustrated in Fig. 4.7.

It is readily checked that the constraints (4.43) to (4.48) are equivalent to those in Theorem 4.1.

Let $\boldsymbol{r} \equiv (r_{12}, R_{11}, R_{12}, r_{21}, R_{22}, R_{21})$ be a rate vector and $\boldsymbol{\mu} \equiv (\mu_1, \mu_2)$ be a vector of Lagrange multipliers. After relaxing constraints (4.47) and (4.48), the following *partial Lagrangian* is formed:

$$L(\boldsymbol{r}, \boldsymbol{\mu}) \equiv w_1(R_{11} + R_{21}) + w_2(R_{22} + R_{12}) + \mu_1(r_{12} - R_{21}) + \mu_2(r_{21} - R_{12}) \tag{4.49}$$

for $\boldsymbol{r} \in \mathcal{R} \equiv \mathcal{C}_{BC}(a_1, b_1, c_1, P_1) \times \mathcal{C}_{BC}(a_2, b_2, c_2, P_2)$ and $\boldsymbol{\mu} \in \mathbb{R}^2$. We have the following *weak duality property*,

$$\max_{\boldsymbol{r}\in\mathcal{R}} \min_{\boldsymbol{\mu}\in\mathbb{R}^2} L(\boldsymbol{r}, \boldsymbol{\mu}) \leq \min_{\boldsymbol{\mu}\in\mathbb{R}^2} \max_{\boldsymbol{r}\in\mathcal{R}} L(\boldsymbol{r}, \boldsymbol{\mu}), \tag{4.50}$$

which holds in general for any function of two sets of variables [84, p.379]. The max-min value on the left hand side of (4.50) is precisely equal to the optimal weighted sum in (4.42). It is because $\min_{\boldsymbol{\mu}\in\mathbb{R}^2} L(\boldsymbol{r}, \boldsymbol{\mu})$ is equal to $w_1(R_{11} + R_{21}) + w_2(R_{22} + R_{12})$ if the constraints $r_{12} = R_{21}$ and $r_{21} = R_{12}$ are satisfied, and $-\infty$ otherwise.

Consider the right hand side of (4.50). Let

$$q(\boldsymbol{\mu}) \equiv \max_{\boldsymbol{r}\in\mathcal{R}} L(\boldsymbol{r}, \boldsymbol{\mu}), \tag{4.51}$$

which is called the *dual function*. For each $\boldsymbol{\mu}$, the value of $q(\boldsymbol{\mu})$ is an upper bound of the maximum weighted sum rate.

By rearranging the terms, the dual function $q(\boldsymbol{\mu})$ can be decomposed into $q(\boldsymbol{\mu}) = q_1(\boldsymbol{\mu}) + q_2(\boldsymbol{\mu})$ where

$$q_1(\boldsymbol{\mu}) \equiv \max_{r_{12}, R_{11}, R_{12}} \{\mu_1 r_{12} + w_1 R_{11} + (w_2 - \mu_2) R_{12}\} \tag{4.52}$$

$$q_2(\boldsymbol{\mu}) \equiv \max_{r_{21}, R_{22}, R_{21}} \{\mu_2 r_{21} + w_2 R_{22} + (w_1 - \mu_1) R_{21}\} \tag{4.53}$$

with the maxima taken over all $(r_{12}, R_{11}, R_{12}) \in \mathcal{C}_{BC}(a_1, b_1, c_1, P_1)$ and $(r_{21}, R_{22}, R_{21}) \in \mathcal{C}_{BC}(a_2, b_2, c_2, P_2)$ respectively. The computation of $q_1(\boldsymbol{\mu})$ and $q_2(\boldsymbol{\mu})$ amounts to the power allocation problems for maximizing the weighted sum rate in the broadcast channels with S_1 and S_2 as the sources respectively. Each of them can be solved by the greedy algorithm in [101], which is briefly described in Section 2.1.3. Let $r^*_{ij}(\boldsymbol{\mu})$ and $R^*_{ij}(\boldsymbol{\mu})$ be the resultant solution. Hence, given any $\boldsymbol{\mu}$, $q(\boldsymbol{\mu})$ can be solved readily.

I will use the following fundamental theorem from the theory of Lagrangian multipliers [84, Theorem 28.3]. A short proof is included here for the sake of completeness.

Theorem 4.7 ([84]). *If $\bar{\boldsymbol{\mu}}$ is a Lagrange multiplier vector such that*

$$r^*_{12}(\bar{\boldsymbol{\mu}}) = R^*_{21}(\bar{\boldsymbol{\mu}}) \tag{4.54}$$

$$r^*_{21}(\bar{\boldsymbol{\mu}}) = R^*_{12}(\bar{\boldsymbol{\mu}}), \tag{4.55}$$

then the corresponding rate vector

$$\bar{\boldsymbol{r}} \equiv (\bar{r}_{12}, \bar{R}_{11}, \bar{R}_{12}, \bar{r}_{21}, \bar{R}_{22}, \bar{R}_{21}) \tag{4.56}$$

$$= (r^*_{12}(\bar{\boldsymbol{\mu}}), R^*_{11}(\bar{\boldsymbol{\mu}}), R^*_{12}(\bar{\boldsymbol{\mu}}), r^*_{21}(\bar{\boldsymbol{\mu}}), R^*_{22}(\bar{\boldsymbol{\mu}}), R^*_{21}(\bar{\boldsymbol{\mu}})) \tag{4.57}$$

which maximizes $L(\boldsymbol{r}, \bar{\boldsymbol{\mu}})$, is an optimal solution to the weighted sum rate maximization problem in (4.42) *to* (4.48).

Proof. From the weak duality property (4.50), it suffices to prove that

$$q(\bar{\boldsymbol{\mu}}) = L(\bar{\boldsymbol{r}}, \bar{\boldsymbol{\mu}}) \leq \max_{\boldsymbol{r}\in\mathcal{R}} \min_{\boldsymbol{\mu}\in\mathbb{R}^2} L(\boldsymbol{r}, \boldsymbol{\mu}). \tag{4.58}$$

Putting (4.54) and (4.55) into the partial Lagrangian, we see that the equality

$$L(\bar{\boldsymbol{r}}, \boldsymbol{\mu}) = w_1(\bar{R}_{11} + \bar{R}_{21}) + w_2(\bar{R}_{22} + \bar{R}_{12}) = L(\bar{\boldsymbol{r}}, \bar{\boldsymbol{\mu}}) \tag{4.59}$$

holds for all $\boldsymbol{\mu} \in \mathbb{R}^2$. In particular, we get

$$L(\bar{\boldsymbol{r}}, \bar{\boldsymbol{\mu}}) = \min_{\boldsymbol{\mu}\in\mathbb{R}^2} L(\bar{\boldsymbol{r}}, \boldsymbol{\mu}), \tag{4.60}$$

which implies

$$L(\bar{\boldsymbol{r}}, \bar{\boldsymbol{\mu}}) \leq \max_{\boldsymbol{r}\in\mathcal{R}} \min_{\boldsymbol{\mu}\in\mathbb{R}^2} L(\boldsymbol{r}, \boldsymbol{\mu}). \tag{4.61}$$

Thus, we have the following saddle-point property,

$$\max_{\boldsymbol{r}\in\mathcal{R}} \min_{\boldsymbol{\mu}\in\mathbb{R}^2} L(\boldsymbol{r}, \boldsymbol{\mu}) = L(\bar{\boldsymbol{r}}, \bar{\boldsymbol{\mu}}) = \min_{\boldsymbol{\mu}\in\mathbb{R}^2} \max_{\boldsymbol{r}\in\mathcal{R}} L(\boldsymbol{r}, \boldsymbol{\mu}). \tag{4.62}$$

This proves that the optimal weighted sum rate is achieved when $\boldsymbol{r} = \bar{\boldsymbol{r}}$, with maximal value $L(\bar{\boldsymbol{r}}, \bar{\boldsymbol{\mu}})$. □

In view of Theorem 4.7, a vector of Lagrangian multipliers, $\boldsymbol{\mu}$, that satisfies the conditions in (4.54) and (4.55) is said to be *optimal.* Theorem 4.7 says that if $\bar{\boldsymbol{\mu}}$ is optimal, then the maximum weighted sum rate can be obtained from (4.59), and the associated optimal rate allocation is given by (4.57). In order to develop an iterative algorithm that computes the optimal $\boldsymbol{\mu}$, the following lemma, which is about some continuous and monotonic properties of $r^*_{ij}(\boldsymbol{\mu})$ and $R^*_{ij}(\boldsymbol{\mu})$, is needed.

Lemma 4.8. *Assume the noise power at the receivers of S_1, S_2, D_1 and D_2 are distinct. r^*_{12}, r^*_{21}, R^*_{12} and R^*_{21} are all continuous functions of $\boldsymbol{\mu}$. For $\mu_i \geq 0$, r^*_{ij} is an increasing function of μ_i and R^*_{ij} is a decreasing function of μ_i. For $\mu_i \leq w_i$, r^*_{ji} is an increasing function of μ_i and R^*_{ji} is a decreasing function of μ_i.*

Proof. See Appendix A.2. □

The heuristic behind Lemma 4.8 is as follows. In the maximization of weighted sum rate in (4.52), the weighting of r_{12} is μ_1. If we increase μ_1, r^*_{12} will increase and R^*_{12} will decrease. The coefficient of R_{12} in (4.52) is $w_2 - \mu_2$. If we increase μ_2, the weighting of R_{12} is decreased. As a consequence, R^*_{12} will decrease but r^*_{12} will increase. Similar heuristic applies to (4.53).

Based on the above results, the optimal weighted sum is computed using an alternating optimization. Firstly, μ_2 is fixed in the vector $\boldsymbol{\mu}$ and search for μ_1 such that (4.54) holds. Using the continuous and monotonic property in Lemma 4.8, this can be done by a simple binary search. Then, the first component μ_1 is fixed in $\boldsymbol{\mu}$ and find μ_2 such that (4.55) holds. Again, by Lemma 4.8, a binary search suffices. The algorithm is stated formally in Table 4.1. The value of ϵ in the algorithm of Table 4.1 is the error tolerance, and is set to a very small positive real number.

Lemma 4.9. *In each iteration t of the algorithm in Table 4.1,*

$$r^*_{12}(\boldsymbol{\mu}(t)) \geq R^*_{21}(\boldsymbol{\mu}(t)) \tag{4.63}$$

$$r^*_{21}(\boldsymbol{\mu}(t)) \leq R^*_{12}(\boldsymbol{\mu}(t)). \tag{4.64}$$

In addition, $\mu_1(t)$ is decreasing with t while $\mu_2(t)$ is increasing with t.

Table 4.1: Search for Optimal $\boldsymbol{\mu}$

$t = 0$
$\mu_1(0) \leftarrow w_1,\ \mu_2(0) \leftarrow 0$
while $|r_{12}^*(\boldsymbol{\mu}(t)) - R_{21}^*(\boldsymbol{\mu}(t)| > \epsilon$ or $|r_{21}^*(\boldsymbol{\mu}(t)) - R_{12}^*(\boldsymbol{\mu}(t)| > \epsilon$ **do**
  search for ν_1 so that $r_{12}^*(\nu_1, \mu_2(t)) = R_{21}^*(\nu_1, \mu_2(t))$
  $t \leftarrow t + 1$
  $\mu_1(t) \leftarrow \nu_1$
  search for ν_2 so that $r_{21}^*(\mu_1(t), \nu_2) = R_{12}^*(\mu_1(t), \nu_2)$
  $\mu_2(t) \leftarrow \nu_2$
end while

Proof. See Appendix A.3. □

Using the above lemmas, the convergence of the algorithm can then be proved.

Theorem 4.10. *In the algorithm in Table 4.1, $\mu_1(t)$ and $\mu_2(t)$ converge to the optimal solution.*

Proof. I first show that $\boldsymbol{\mu}_1(t)$ and $\boldsymbol{\mu}_2(t)$ converges. For the convergence of $\boldsymbol{\mu}_1(t)$, I consider two cases. In the first case, suppose that $\mu_1(t) < 0$ for some t, say at $t = t'$. We then have $r_{12}^*(\boldsymbol{\mu}(t')) = 0$. By (4.63) in Lemma 4.9,

$$0 = r_{12}^*(\boldsymbol{\mu}(t')) \geq R_{21}^*(\boldsymbol{\mu}(t')) \geq 0 \tag{4.65}$$

which means $r_{12}^*(\boldsymbol{\mu}(t')) = R_{21}^*(\boldsymbol{\mu}(t')) = 0$. This implies that $\mu_1(t) = \mu_1(t')$ for all $t \geq t'$. Hence $\mu_1(t)$ converges in this case. In the second case, suppose that $\mu_1(t) \geq 0$ for all t, i.e., $\mu_1(t)$ is lower bounded by 0. It has been shown in Lemma 4.9 that $\mu_1(t)$ is a decreasing function of t. Hence, if $\mu_1(t)$ is lower bounded by 0, it converges. Similarly, by considering the cases of $\mu_2(t)$ being or not being upper bounded by w_2, I can show that $\mu_2(t)$ is convergent.

After establishing the convergence of $\mu_1(t)$ and $\mu_2(t)$, let the limit of $\mu_1(t)$ and $\mu_2(t)$ be $\bar{\mu}_1$ and $\bar{\mu}_2$ respectively. Because $r_{12}^*(\mu_1(t+1), \mu_2(t)) = R_{21}^*(\mu_1(t+1), \mu_2(t))$ for all $t \geq 0$ by construction, and r_{12}^* and R_{21}^* are continuous by Lemma 4.8, I obtain $r_{12}^*(\bar{\mu}_1, \bar{\mu}_2) =$

$R^*_{21}(\bar{\mu}_1, \bar{\mu}_2)$. Therefore (4.54) holds with $\boldsymbol{\mu}_0 = (\bar{\mu}_1, \bar{\mu}_2)$. By the same argument and the fact that $r^*_{21}(\mu_1(t), \mu_2(t)) = R^*_{12}(\mu_1(t), \mu_2(t))$ for all $t \geq 1$, I can establish the equality in (4.55). The optimality of the solution then follows immediately from Theorem 4.7. □

For positive w_1 and w_2, after running the algorithm in Table 4.1, the limit of $\mu_1(t)$ and $\mu_2(t)$ is obtained. Let $\bar{\mu}$ be the limit of $(\mu_1(t), \mu_2(t))$. The optimal rate allocation which maximize $w_1 R_1 + w_2 R_2$ is denoted as

$$\check{R}_1(w_1, w_2) \equiv R^*_{11}(\bar{\boldsymbol{\mu}}) + R^*_{21}(\bar{\boldsymbol{\mu}}) \tag{4.66}$$

$$\check{R}_2(w_1, w_2) \equiv R^*_{22}(\bar{\boldsymbol{\mu}}) + R^*_{12}(\bar{\boldsymbol{\mu}}). \tag{4.67}$$

4.4.2 Extension to Parallel Channel Case

Scheme 1 is then generalized to frequency selective channel, represented by a bank of parallel Gaussian channels. To be more specific, consider the case that the source nodes employ multi-channel transmissions over disjoint and orthogonal frequency bands. Node S_1 has N_1 parallel sub-channels, whereas S_2 has N_2. If each parallel sub-channel represents one frequency carrier, then the model can be used to represent orthogonal frequency division multiplex (OFDM) transceivers. Let the bandwidth of each subchannel by B, so that the total bandwidth is equal to $B_W = B(N_1 + N_2)$.

Consider the transmission of S_i. Denote the power gain of subchannel k from S_i to S_j, D_i, and D_j by $a_i^{(k)}$, $b_i^{(k)}$, and $c_i^{(k)}$, respectively. Let P_i be the maximal transmission power of S_i, $i = 1, 2$. The total power P_i is split into N_i parts, $P_i^{(k)}$, $k = 1, 2, \ldots, N_i$, such that $P_i^{(k)}$ is the power associated with the k-th sub-channel of S_i, and

$$P_i^{(1)} + P_i^{(2)} + \ldots + P_i^{(N_i)} = P_i. \tag{4.68}$$

Each source node is associated with a parallel BC channel. The capacity region of the parallel BC channel for S_i is

$$\mathcal{C}_i^{\|}(P_i) \equiv \bigcup_{\sum_k P_i^{(k)} = P_i} \left\{ \boldsymbol{v}_1 + \ldots + \boldsymbol{v}_{N_i} : \boldsymbol{v}_k \in \mathcal{C}_{BC}(a_i^{(k)}, b_i^{(k)}, c_i^{(k)}, P_i^{(k)}) \right\}, \tag{4.69}$$

with the union taken over all power allocations satisfying (4.68).

Similar to the single channel case, each source node S_i splits its own data stream into two streams: one direct to its intended receiver D_i and the other through a two-hop path with the other source node as relay. Each stream is divided into many blocks, each of which contains a codeword. The relay node decodes the received block from S_i, and then re-encodes and forwards a new block to D_i.

By a similar argument in Theorem 4.1, a rate pair

$$(R_1, R_2) = (r_{1r} + r_{1d}, r_{2r} + r_{2d}) \tag{4.70}$$

(with unit bits/s/Hz) is achievable if it satisfies

$$(r_{1r}, r_{1d}, r_{2r}) \in \frac{1}{N_1 + N_2} \mathcal{C}_1^{\|}(P_1) \tag{4.71}$$

$$(r_{2r}, r_{2d}, r_{1r}) \in \frac{1}{N_1 + N_2} \mathcal{C}_2^{\|}(P_2). \tag{4.72}$$

As in the single-channel case, the weighted sum rate maximization problem of this scheme can be solved by decomposing it into two weighted sum rate maximization problems for the two parallel Gaussian BCs. We can use the greedy algorithm in [101], which is applicable to the parallel Gaussian BCs.

4.5 Rate Allocation in Low SNR Regime

The system is in the low SNR regime when bandwidth is large. When noise levels are very high, time-sharing is asymptotically optimal in the broadcast channel [59]. In fact, by applying the Taylor series expansion

$$\ln(1 + x) = x - x^2/2 + x^3/3 + \ldots \tag{4.73}$$

The capacity region of the k-th sub-channel of S_i can be approximated as

$$r_{ij}^{(k)} \leq \frac{\alpha_{i1}^{(k)}}{\ln 2} \frac{P_i^{(k)}}{n_{i1}^{(k)}/a_i^{(k)}} \tag{4.74}$$

$$R_{ii}^{(k)} \leq \frac{\alpha_{i2}^{(k)}}{\ln 2} \frac{P_i^{(k)}}{n_{i2}^{(k)}/b_i^{(k)}} \tag{4.75}$$

$$R_{ij}^{(k)} \leq \frac{\alpha_{i3}^{(k)}}{\ln 2} \frac{P_i^{(k)}}{n_{i3}^{(k)}/c_i^{(k)}} \tag{4.76}$$

where $\alpha_{i1}^{(k)} + \alpha_{i2}^{(k)} + \alpha_{i3}^{(k)} = 1$. The boundary of the capacity region can be approximated by the plane

$$n_{i1}^{(k)} \frac{r_{ij}^{(k)}}{a_i^{(k)}} + n_{i2}^{(k)} \frac{R_{ii}^{(k)}}{b_i^{(k)}} + n_{i3}^{(k)} \frac{R_{ij}^{(k)}}{c_i^{(k)}} = \frac{P_i^{(k)}}{\ln 2}. \tag{4.77}$$

Note that the fraction $n_{ij}^{(k)}$ is interpreted as the spectral density of white noise. From (4.77), the achievable rate vectors on the boundary can be obtained as linear combinations of the three vertices of the capacity region. This amounts to time-sharing of the sub-channel among the three streams of data, thus no complicated coding and interference cancelation is required at low SNR.

Putting all N_i sub-channels together, the capacity region of the parallel broadcast channel $C_i(P_i)$ is also well approximated by a polyhedron.

Theorem 4.11. *In the low SNR regime, the capacity region of the parallel broadcast channel $C_i(P_i)$, $i = 1, 2$, can be approximated by*

$$\Big\{(r_{ij}, R_{ii}, R_{ij}) \in \mathbb{R}_+^3 : \chi_i r_{ij} + \psi_i R_{ii} + \varphi_i R_{ij} \leq \frac{P_i}{\ln 2}\Big\}, \tag{4.78}$$

where

$$\chi_i \equiv \min\{n_{i1}^{(k)} a_i^{(k)}) : k = 1, 2, \ldots, N_i\} \tag{4.79}$$

$$\psi_i \equiv \min\{n_{i2}^{(k)} b_i^{(k)}) : k = 1, 2, \ldots, N_i\} \tag{4.80}$$

$$\varphi_i \equiv \min\{n_{i3}^{(k)} c_i^{(k)}) : k = 1, 2, \ldots, N_i\}. \tag{4.81}$$

Proof. The statement amounts to the following claim: For $k = 1, 2, \ldots, K$, let $X_k(s_k)$ be the polyhedron

$$\{(x, y, z) \in \mathbb{R}_+^3 : p_k x + q_k y + r_k z \leq s_k\}, \tag{4.82}$$

where p_k, q_k, r_k and s_k are positive constants for $k = 1, \ldots, K$. The set

$$\mathcal{A} \equiv \bigcup_{s_1+\ldots+s_K=s} \{\boldsymbol{v}_1 + \ldots + \boldsymbol{v}_k : \boldsymbol{v}_k \in X_k(s_k)\} \tag{4.83}$$

is equal to

$$\mathcal{B} \equiv \{(x, y, z) \in \mathbb{R}_+^3 : px + qy + rz \leq s\}, \tag{4.84}$$

where $p \equiv \min_{1\leq k\leq K} p_k$, $q \equiv \min_{1\leq k\leq K} q_k$, $r \equiv \min_{1\leq k\leq K} r_k$ and $s = s_1 + \ldots + s_K$.

It is clear that any point in $\mathcal{A}$ is also in $\mathcal{B}$, i.e., $\mathcal{A} \subseteq \mathcal{B}$. Conversely, suppose that $p = p_{k_1}$, $q = q_{k_2}$ and $r = r_{k_3}$, and (x, y, z) be a point in $\mathcal{B}$. I will complete the proof by considering several cases.

First, suppose that k_1, k_2 and k_3 are distinct. I write (x, y, z) as the sum of $(x, 0, 0)$, $(0, y, 0)$ and $(0, 0, z)$. We have

$$(x, 0, 0) \in X_{k_1}(px) \tag{4.85}$$

$$(0, y, 0) \in X_{k_2}(qy) \tag{4.86}$$

$$(0, 0, z) \in X_{k_3}(rz). \tag{4.87}$$

Since $px + yz + qz \leq s$, I can conclude that $(x, y, z) \in \mathcal{A}$.

Second, consider the case $k_1 = k_2 \neq k_3$. We write (x, y, z) as $(x, y, 0) + (0, 0, z)$, and note that

$$(x, y, 0) \in X_{k_1}(px + qy) \tag{4.88}$$

$$(0, 0, z) \in X_{k_3}(rz). \tag{4.89}$$

Again, I can conclude that $(x, y, z) \in \mathcal{A}$.

The remaining cases can be treated similarly and are omitted. □

The implication of the above theorem is that for any transmission node, though having a group of sub-channels, only uses the best sub-channel, in the sense that its effective noise power is smallest. For example, S_1 sends to S_2 only via the sub-channel that minimizes $n_{i1}^{(k)} a_i^{(k)}$.

Hence, the achievable rate region for the cooperative orthogonal-division channel at low SNR can be approximated by (4.7) (4.8) (4.47) (4.48) and

$$\chi_1 r_{12} + \psi_1 R_{11} + \varphi_1 R_{12} \leq \frac{P_1}{\ln 2} \tag{4.9'}$$

$$\chi_2 r_{21} + \psi_2 R_{22} + \varphi_2 R_{21} \leq \frac{P_2}{\ln 2}. \tag{4.10'}$$

Its boundary is piece-wise linear. Consider the weighted sum rate maximization problem again. The objective function is $w_1 R_1 + w_2 R_2$. This problem reduces to a linear programming problem, and can be solved by standard techniques. On the other hand, a closer observation can further simplify the problem.

Given any weights w_1 and w_2, let $r_{ij}^{\dagger}$, $R_{ii}^{\dagger}$ and $R_{ij}^{\dagger}$ be the corresponding optimal solutions of r_{ij}, R_{ii} and R_{ij} respectively. Since the objective function increases with $R_{11}^{\dagger}$ and $R_{22}^{\dagger}$, equalities must hold in (4.9') and (4.10'), and we have

$$R_{ii}^{\dagger} = \frac{1}{\psi_i}\left(\frac{P_i}{\ln 2} - \chi_i r_{ij}^{\dagger} - \varphi_i r_{ji}^{\dagger}\right), \tag{4.90}$$

for $i = 1, 2$. Then, substitute $R_{11}^{\dagger}$ and $R_{22}^{\dagger}$ by (4.90) in the objective function. Removing the constant terms, the weighted sum rate maximization problem is transformed into the following linear programming problem:

$$\max_{r_{12}, r_{21} \geq 0} \alpha_1 r_{12} + \alpha_2 r_{21} \tag{4.91}$$

subject to

$$\frac{\chi_1}{\varphi_1} r_{12} + r_{21} \leq \frac{P_1}{(\ln 2)\varphi_1} \tag{4.92}$$

$$r_{12} + \frac{\chi_2}{\varphi_2} r_{21} \leq \frac{P_2}{(\ln 2)\varphi_2} \tag{4.93}$$

where for $i = 1, 2$,

$$\alpha_i \equiv w_i\left(1 - \frac{\chi_i}{\psi_i}\right) - \frac{w_j \varphi_j}{\psi_j}. \tag{4.94}$$

Note that there are only two variables in this linear programming problem. Recall that r_{ij} is the information rate from S_i to S_j. The objective function is a weighted sum between r_{12} and r_{21}, where the weight, α_i depends on the normalized effective noise powers in the relay paths. For example, consider α_1. The effective noise in the link from S_1 to S_2 is χ_1 and the effective noise in the link from S_2 to D_1 is φ_2. The value of α_1 depends on these two terms.

From the problem formulation, it is straightforward to see that:

1. If $\alpha_1, \alpha_2 < 0$, $r_{12}^{\dagger} = r_{21}^{\dagger} = 0$.

2. If $\alpha_i > 0$, $\alpha_j < 0$, $r_{ij}^{\dagger} = r_{i,max}$ and $r_{ji}^{\dagger} = 0$ where

$$r_{i,max} = \frac{1}{\ln 2} \min \left\{ \frac{P_i}{\chi_i}, \frac{P_j}{\varphi_j} \right\}. \tag{4.95}$$

Intuitively, α_i measures how good the relay path for S_i is. The above results say that when α_i is too small, or more precisely, less than zero, the relay path should not be used.

For $\alpha_1, \alpha_2 > 0$, the optimal solution is given by the following theorem:

Theorem 4.12. *Let*

$$\underline{m} = \min \left\{ \frac{\chi_1}{\varphi_1}, \frac{\varphi_2}{\chi_2} \right\}, \tag{4.96}$$

$$\overline{m} = \max \left\{ \frac{\chi_1}{\varphi_1}, \frac{\varphi_2}{\chi_2} \right\}, \tag{4.97}$$

$$\Lambda = \chi_1\chi_2 - \varphi_1\varphi_2, \tag{4.98}$$

$$\tilde{r}_{12} = \frac{1}{\ln 2} \frac{\chi_2 P_1 - \varphi_1 P_2}{\Lambda}, \tag{4.99}$$

$$\tilde{r}_{21} = \frac{1}{\ln 2} \frac{\chi_1 P_2 - \varphi_2 P_1}{\Lambda}. \tag{4.100}$$

In the low SNR regime, if $\alpha_1, \alpha_2 > 0$,

1. If $\frac{\alpha_1}{\alpha_2} \leq \underline{m}$, $r_{12}^{\dagger} = 0$ and $r_{21}^{\dagger} = r_{2,max}$.

2. If $\frac{\alpha_1}{\alpha_2} \geq \overline{m}$, $r_{12}^{\dagger} = r_{1,max}$ and $r_{21}^{\dagger} = 0$.

3. If $\underline{m} \leq \frac{\alpha_1}{\alpha_2} \leq \overline{m}$ and $\chi_1\chi_2 \neq \varphi_1\varphi_2$ (i.e. $\underline{m} < \overline{m}$), there is at most one of $\tilde{r}_{12}$ and $\tilde{r}_{21}$

being negative and

$$r_{12}^{\dagger} = \begin{cases} \tilde{r}_{12}, & \text{if } \tilde{r}_{12}, \tilde{r}_{21} \geq 0 \\ r_{1,max}, & \text{if } \tilde{r}_{21} < 0 \\ 0, & \text{otherwise} \end{cases} \tag{4.101}$$

$$r_{21}^{\dagger} = \begin{cases} \tilde{r}_{21}, & \text{if } \tilde{r}_{12}, \tilde{r}_{21} \geq 0 \\ r_{2,max}, & \text{if } \tilde{r}_{12} < 0 \\ 0, & \text{otherwise.} \end{cases} \tag{4.102}$$

Proof. See Appendix A.4. □

Although the expression for the optimal solution looks complicated, the idea is simple. First of all, note that at the point $(\tilde{r}_{12}, \tilde{r}_{21})$, equalities hold in both (4.92) and (4.93). According to Theorem 4.12, if both components are positive, then it is the optimal solution. In that case, both $R_{11}^{\dagger}$ and $R_{22}^{\dagger}$ are equal to zero, meaning that there is no traffic through the direct path. If any of the components is non-positive, then the optimal solution is located at the boundary.

4.6 Max-min Fairness

Apart from maximizing the weighted sum rate, one application of that algorithm is to achieve max-min fairness. We say that the rate pair (R_1^M, R_2^M) is *max-min fair* if it satisfies the following two properties: (a) the minimum of R_1^M and R_2^M is maximized. (b) If $R_1^M \geq R_2^M$, then R_1^M is the maximum rate of user 1 given $R_2 = R_2^M$; if $R_2^M \geq R_1^M$, then R_2^M is the maximum rate of user 2 given $R_1 = R_1^M$. It follows from the definition that max-min fair rate pair is Pareto optimal. As I have already devised an algorithm for the computation of Pareto optimal rate allocation, I will leverage it and compute max-min fair rate allocation.

Noted that in general, max-min fairness may not be the same as symmetric fairness. A rate pair is symmetric fair if it is in the form (R, R) with R chosen as large as possible such that (R, R) remains achievable. For example, consider the set in $\mathbb{R}_+^2$ consisting of points (x, y)

such that $y \leq 1$ and $x + y \leq 3$. Then the point $(1, 1)$ is symmetric fair but not max-min fair. The max-min fair point in this region is $(2, 1)$.

The computation of the max-min fair rate pair is based on the following observation.

Proposition 4.13. *For fixed w_2, $\check{R}_1(w_1, w_2)$ is a non-decreasing function of w_1 but it is a non-increasing function of w_2, meanwhile, for fixed w_1, $\check{R}_2(w_1, w_2)$ is a non-increasing function of w_1 but it is a non-decreasing function of w_2. In particular, the function*

$$\triangle R(w) \equiv \check{R}_1(w, 1 - w) - \check{R}_2(w, 1 - w) \tag{4.103}$$

defined for $0 \leq w \leq 1$, is non-decreasing.

Proof. To show that $\check{R}_1(w_1, w_2)$ is non-decreasing in w_1 for fixed w_2, I suppose on the contrary that there exists w_1, w_2 and δ such that $\delta > 0$ and $\check{R}_1(w_1 + \delta, w_2) < \check{R}_1(w_1, w_2)$. Then,

$$(w_1 + \delta)\check{R}_1(w_1 + \delta, w_2) + w_2\check{R}_2(w_1 + \delta, w_2) \tag{4.104}$$

$$\leq w_1\check{R}_1(w_1, w_2) + w_2\check{R}_2(w_1, w_2) + \delta\check{R}_1(w_1 + \delta, w_2) \tag{4.105}$$

$$< w_1\check{R}_1(w_1, w_2) + w_2\check{R}_2(w_1, w_2) + \delta\check{R}_1(w_1, w_2) \tag{4.106}$$

$$= (w_1 + \delta_1)\check{R}_1(w_1, w_2) + w_2\check{R}_2(w_1, w_2). \tag{4.107}$$

The first inequality follows from the defining property that the weighted sum $w_1R_1 + w_2R_2$ is maximized at $(\check{R}_1(w_1, w_2), \check{R}_2(w_1, w_2))$. The second inequality follows from the assumptions $\delta > 0$ and $\check{R}_1(w_1+\delta, w_2) < \check{R}_1(w_1, w_2)$. This contradicts the fact that $(\check{R}_1(w_1+\delta, w_2), \check{R}_1(w_1+\delta, w_2))$ maximizes the weighted sum rate $(w_1 + \delta)R_1 + w_2R_2$. Hence, $\check{R}_1(w_1 + \delta, w_2) \geq \check{R}_1(w_1, w_2)$.

The remaining statement can be proved similarly. □

Using the property that $(\check{R}_1(w, 1 - w), \check{R}_2(w, 1 - w))$ spans through all Pareto optimal rate pairs when w varies from 0 to 1, we have the following

Theorem 4.14. *When $|\triangle R(w)|$ is minimized over $0 \leq w \leq 1$, say at $w = w_0$, then the rate pair $(\check{R}_1(w_0, 1 - w_0), \check{R}_2(w_0, 1 - w_0))$ is max-min fair.*

Proof. I consider three cases:

(i) Minimum of $|\triangle R(w)|$ equals 0. In this case, $\triangle R(w_0) = 0$. The rate pair $(\check{R}_1(w_0, 1 - w_0), \check{R}_2(w_0, 1 - w_0))$ is a Pareto optimal symmetric rate pair. Hence it is max-min fair.

(ii) $\triangle R(w) > 0$ for all $0 \le w \le 1$, i.e., $\check{R}_1(w_0, 1 - w_0) > \check{R}_2(w_0, 1 - w_0)$ for all w. Since $\triangle R(w)$ is non-decreasing, $\triangle R(w)$ is minimized at $w = 0$. I have shown in Prop. 4.13 that $\check{R}_2(w, 1 - w)$ is non-increasing as a function of w. Therefore $\check{R}_2(0, 1)$ is the maximal value of R_2 in the rate region. Since $(\check{R}_1(0, 1), \check{R}_2(0, 1))$ is Pareto optimal, we cannot further increase $\check{R}_1(0, 1)$ given $R_2 = \check{R}_2(0, 1)$. This verifies conditions (a) and (b) in the first paragraph of this section.

(iii) $\triangle R(w) < 0$ for all $0 \le w \le 1$. The proof is similar to part (ii) and is omitted. □

This leads to the following algorithm for computing (R_1^M, R_2^M). If $\triangle R(0) > 0$, then $(\check{R}_1(0, 1), \check{R}_2(0, 1))$ is the max-min fair rate pair. If $\triangle R(0) < 0$, then $(\check{R}_1(1, 0), \check{R}_2(1, 0))$ is the max-min fair rate pair. Otherwise, we search for the w between 0 and 1 such that $\triangle R(w) = 0$. The search is one-dimensional, and can be done by bisection search for instance.

For low SNR regime, we have the following closed form solution.

Theorem 4.15. *For $i \neq j$, let*

$$\overline{R}_i \equiv \frac{P_i}{\psi_i} + \left[1 - \frac{\chi_i}{\psi_i}\right]^+ \min\left\{\frac{P_i}{\chi_i}, \frac{P_j}{\varphi_j}\right\}, \tag{4.108}$$

where $[x]^+ = \max\{x, 0\}$. The max-min fair rate is given by

$$R = \frac{1}{\ln 2} \min\{\overline{R}_1, \overline{R}_2\}. \tag{4.109}$$

In addition, let R_{ii}^M and r_{ij}^M be the values of R_{ii} and r_{ij}, which achieve the max-min fairness, respectively. Then,

$$R_{ii}^M = \frac{1}{\psi_i}\left(\frac{P_i}{\ln 2} - \chi_i r_{ij}^M - \varphi_i r_{ji}^M\right) \tag{4.110}$$

and

$$r_{ij}^M = \begin{cases} r_{i,max}, & \text{if } \overline{R}_i > \overline{R}_j \text{ and } \psi_i > \chi_i \\ 0, & \text{otherwise.} \end{cases} \tag{4.111}$$

where $r_{i,max}$ is defined in (4.95).

Proof. By Theorem 4.11, the max-min fairness can be obtained by solving the following optimization problem.

$$\max R \tag{4.112}$$

subject to

$$R_{11} + r_{12} \geq R \tag{4.113}$$

$$R_{22} + r_{21} \geq R \tag{4.114}$$

$$\chi_1 r_{12} + \psi_1 R_{11} + \varphi_1 r_{21} \leq \frac{P_1}{\ln 2} \tag{4.115}$$

$$\chi_2 r_{21} + \psi_2 R_{22} + \varphi_2 r_{12} \leq \frac{P_2}{\ln 2} \tag{4.116}$$

$$R, R_{11}, R_{22}, r_{12}, r_{21} \geq 0. \tag{4.117}$$

Hence, if r_{12}^M and r_{21}^M are given,

$$R - r_{ij}^M \leq R_{ii} \leq \frac{1}{\psi_i}\left(\frac{P_i}{\ln 2} - \chi_i r_{ij}^M - \varphi_i r_{ji}^M\right) \tag{4.118}$$

which implies (4.110). It also implies that

$$R \leq R_{ii}^M + r_{ij}^M \leq \frac{1}{\psi_i}\left(\frac{P_i}{\ln 2}\right) + \left(1 - \frac{\chi_i}{\psi_i}\right) r_{ij}^M - \frac{\varphi_i}{\psi_i} r_{ji}^M, \quad i = 1, 2 \tag{4.119}$$

$$\chi_i r_{ij} + \varphi_i r_{ji} \leq \frac{P_i}{\ln 2}, \quad i = 1, 2. \tag{4.120}$$

But

$$\frac{1}{\psi_i}\left(\frac{P_i}{\ln 2}\right) + \left(1 - \frac{\chi_i}{\psi_i}\right) r_{ij}^{(S)} - \frac{\varphi_i}{\psi_i} r_{ji}^{(S)} \leq \frac{1}{\psi_i}\left(\frac{P_i}{\ln 2}\right) + \left(1 - \frac{\chi_i}{\psi_i}\right) r_{ij}^{(S)} \leq \frac{1}{\psi_i}\left(\frac{P_i}{\ln 2}\right) + \left[1 - \frac{\chi_i}{\psi_i}\right]^+ r_{i,max}. \tag{4.121}$$

Hence,

$$R \leq \frac{1}{\ln 2}\min\{\overline{R}_1, \overline{R}_2\} \tag{4.122}$$

where equality holds if R_{ii}^M and r_{ij}^M satisfy (4.110) and (4.111). □

From (4.111), it can be inferred that at most one of r_{12}^M and r_{21}^M is non-zero. That means, in order to achieve max-min fairness in low SNR regime, there is at most one data stream which requires relaying.

In the proof, it can be concluded that in low SNR regime, max-min fairness is equivalent to symmetric fairness. For $i = 1, 2$ in (4.120), R is equal to the rightmost side of the inequality. Therefore, when max-min fairness is achieved, R_1 and R_2 are equal.

4.7 Joint Utility Maximization

In this section, I propose an algorithm to extend the weighted sum rate maximization algorithm in Section 4.4 to solve a more general joint utility maximization problem. It turns out that this algorithm can be extended to a framework which can extend a certain class of weighted sum rate maximization algorithm into a joint utility maximization algorithm of the same rate region. A brief description of this framework can be found in Appendix B and a detailed description can be found in [75].

Let $\mathbf{R} = (R_1, R_2)$ and $\mathcal{C}$ be the achievable rate region of Scheme 1. The objective is to maximize a joint utility function $U(\mathbf{R})$ subject to $\mathbf{R} \in \mathcal{C}$. $U(\mathbf{R})$ is assumed to satisfy the following properties: (i) strictly concave, (ii) twice continuously differentiable, and (iii) increasing with any component of $\mathbf{R}$ with the other components fixed. Since the feasible region $\mathcal{C}$ is compact and convex, and U is strictly concave, there exists a unique solution. One example of this problem is the following *Harmonic Mean Fairness Problem.*

Example 4.1. *(Harmonic-Mean Fairness in Cooperative Transmission) In this example, the objective is to achieve harmonic-mean fairness [68]. Let R_i be the end-to-end rate from S_i to D_i, for $i = 1, 2$. The joint utility function is given by*

$$U_{hmf}(\mathbf{R}) \triangleq -\frac{1}{R_1} - \frac{1}{R_2}, \tag{4.123}$$

which is a strictly concave function of $\mathbf{R}$.

For two vectors $\mathbf{x}$ and $\mathbf{y}$, I will write $\mathbf{x} \succeq \mathbf{y}$ if each component of $\mathbf{x}$ is larger than or equal to the corresponding component in $\mathbf{y}$. With this notation, the third assumption of the utility function means that if $\mathbf{R}_1 \succeq \mathbf{R}_2$, then $U(\mathbf{R}_1) \geq U(\mathbf{R}_2)$.

4.7.1 Dual Decomposition

The dual problem is decomposed into two subproblems. The key of this decomposition is the introduction of an auxiliary rate vector $\tilde{\mathbf{R}} \triangleq (\tilde{R}_1, \tilde{R}_2)$ and reformulate the problem into:

$$\max_{\tilde{\mathbf{R}} \in \mathcal{B}} U(\tilde{\mathbf{R}}) \tag{4.124}$$

subject to

$$\mathbf{R} \succeq \tilde{\mathbf{R}} \tag{4.125}$$

$$\mathbf{R} \in \mathcal{C} \tag{4.126}$$

where $\mathcal{B}$ is a closed rectangular box in $\mathbb{R}_+^2$ of the form

$$\mathcal{B} \triangleq \{\tilde{\mathbf{R}} : \tilde{R}_n \in [0, b_n], n = 1, 2\} \tag{4.127}$$

that contains $\mathcal{C}$. Such a bounding box $\mathcal{B}$ exists because the achievable rate region is bounded. Using the property that $U(\tilde{\mathbf{R}})$ is monotonically increasing for every component of $\tilde{\mathbf{R}}$, we can see that the reformulated version is in fact equivalent to the original version.

Then, (B.6) is relaxed to form the partial Lagrangian:

$$L(\mathbf{R}, \tilde{\mathbf{R}}, \boldsymbol{\mu}) \triangleq U(\tilde{\mathbf{R}}) + \sum_{n=1}^{2} \mu_n (R_n - \tilde{R}_n), \tag{4.128}$$

where μ_n's are non-negative Lagrange multipliers. Denote the vector (μ_1, μ_2) by $\boldsymbol{\mu}$. The partial Lagrangian can be rearranged as

$$L(\mathbf{R}, \tilde{\mathbf{R}}, \boldsymbol{\mu}) = \left[U(\tilde{\mathbf{R}}) - \boldsymbol{\mu} \cdot \tilde{\mathbf{R}}\right] + \boldsymbol{\mu} \cdot \mathbf{R}. \tag{4.129}$$

Define the partial dual function $q(\boldsymbol{\mu})$ by

$$q(\boldsymbol{\mu}) \triangleq \max_{\mathbf{R} \in \mathcal{C}, \tilde{\mathbf{R}} \in \mathcal{B}} L(\mathbf{R}, \tilde{\mathbf{R}}, \boldsymbol{\mu}). \tag{4.130}$$

The dual problem is

$$\min_{\boldsymbol{\mu} \succeq \mathbf{0}} q(\boldsymbol{\mu}) \tag{4.131}$$

with the minimum taken over all nonnegative μ_n's. Computation of the partial dual function amounts to solving two independent optimization subproblems,

$$q(\boldsymbol{\mu}) = \max_{\tilde{\mathbf{R}} \in \mathcal{B}} \left\{ U(\tilde{\mathbf{R}}) - \boldsymbol{\mu} \cdot \tilde{\mathbf{R}} \right\} + \max_{\mathbf{R} \in \mathcal{C}} \boldsymbol{\mu} \cdot \mathbf{R}. \tag{4.132}$$

The first subproblem is the maximization of

$$U(\tilde{\mathbf{R}}) - \boldsymbol{\mu} \cdot \tilde{\mathbf{R}} \tag{4.133}$$

over all $\tilde{\mathbf{R}} \in \mathcal{B}$. Since U is strictly concave, a unique maximum exists. The optimal solution for a given $\boldsymbol{\mu}$ is denoted by $\tilde{\mathbf{R}}^*(\boldsymbol{\mu}) = (\tilde{R}_1^*(\boldsymbol{\mu}), \tilde{R}_2^*(\boldsymbol{\mu}))$. The variables $\tilde{R}_n$'s are optimized in a round robin manner by fixing other variables. Let $\tilde{R}_n(\boldsymbol{\mu}, t)$ be the value of $\tilde{R}_n$ at iteration t. $\tilde{R}_n(\boldsymbol{\mu}, t)$ is obtained by the following equation:

$$\tilde{R}_1(\boldsymbol{\mu}, t) = \arg \max_{0 \le \tilde{R}_1 \le b_1} U\big(\tilde{R}_1, \tilde{R}_2(\boldsymbol{\mu}, t)\big) - \mu_1 \tilde{R}_1 \tag{4.134}$$

$$\tilde{R}_2(\boldsymbol{\mu}, t) = \arg \max_{0 \le \tilde{R}_2 \le b_2} U\big(\tilde{R}_1(\boldsymbol{\mu}, t), \tilde{R}_2\big) - \mu_2 \tilde{R}_2 \tag{4.135}$$

The convergence of the above algorithm is shown below.

Proposition 4.16. *The Gauss-Seidel-type algorithm in (4.134)-(4.135) converges to the optimal* $\tilde{\mathbf{R}}$.

Proof. I apply the results in [33, Proposition 6]. The objective function of the first subproblem is differentiable and strictly convex. Hence, it is pseudoconvex.[2] In addition, within the feasible set of $\tilde{\mathbf{R}}$, the level sets of the objective function are compact. Thus, the objective function satisfies the assumption in [33, Proposition 6]. Finally, the feasible set is the Cartesian product of the compact sets $[0, b_n]$. Therefore, by [33, Proposition 6], the algorithm converges to the optimal solution. □

By taking the partial derivative of (4.133), the Gauss-Seidel update of $\tilde{R}_n$ is done by solving the following equation:

$$\frac{\partial U(\tilde{\mathbf{R}})}{\partial \tilde{R}_n} = \mu_n. \tag{4.136}$$

[2] A differentiable function f is pseudoconvex if it satisfies the property $\nabla f(x)(y - x) \ge 0$ implies $f(y) \ge f(x)$. Note that if a function g is differentiable and convex, it is pseudoconvex.

Since U is twice continuously differentiable and strictly concave, the left hand side of (4.136) is a non-increasing and continuous function of $\tilde{R}_n$. Hence, it can be efficiently solved by bisection method. However, the solution of (4.136) may not be within the interval $[0, b_n]$ and we need to project the solution to this interval. If the left hand side of (4.136) is negative for both $\tilde{R}_n = 0$ and $\tilde{R}_n = b_n$, the solution of (4.136) is negative and the optimal $\tilde{R}_n$ is 0. Similarly, if the left hand side of (4.136) is positive for both $\tilde{R}_n = 0$ and $\tilde{R}_n = b_n$, the optimal $\tilde{R}_n$ is b_n.

If the joint utility function can be expressed as

$$U(\tilde{\mathbf{R}}) = \sum_{n=1}^{2} U_n(\tilde{R}_n), \tag{4.137}$$

this subproblem can be solved in a simpler way. By taking the gradient of the objective function of this subproblem, $\tilde{\mathbf{R}}^*(\boldsymbol{\mu})$ can be obtained by solving the following system of equations:

$$U_n'(\tilde{R}_n) = \mu_n, \quad n = 1, 2. \tag{4.138}$$

Hence, we can solve for each R_n separately without the need of Gauss-Seidel iterations. In addition, U_n is concave and continuously differentiable for all n, then U_n' is a monotonic decreasing and continuous function and thus each $\tilde{R}_n$ can be solved by much faster numerical techniques such as the bisection method.

The second subproblem is the maximization of $\boldsymbol{\mu} \cdot \mathbf{R}$ over $\mathbf{R} \in \mathcal{C}$. This is the weighted sum rate maximization problem. Let $\mathbf{R}^*(\boldsymbol{\mu}) = (R_1^*(\boldsymbol{\mu}), R_2^*(\boldsymbol{\mu}))$ be the optimal solution for a given $\boldsymbol{\mu}$. It is well defined for all $\boldsymbol{\mu} \succeq \mathbf{0}$. It is because the weighted sum rate maximization problem has unique optimal solution.

4.7.2 The Iterative Numerical Algorithm

The proposed algorithm is a nonlinear Gauss-Seidel algorithm that maximizes the joint utility function by adjusting the dual variable $\boldsymbol{\mu}$ in a round-robin fashion. The algorithm computes recursively a sequence of dual variables $\boldsymbol{\mu}(t)$, $t \geq 1$, by

$$\mu_1(t+1) = \arg\min_{\xi \geq 0} q(\xi, \mu_2(t)) \tag{4.139}$$

$$\mu_2(t+1) = \arg\min_{\xi \geq 0} q(\mu_1(t+1), \xi) \tag{4.140}$$

In each iteration, one dual variable is optimized, say μ_n, of $\boldsymbol{\mu}$, while the other one is held fixed. Then, μ_n is replaced by the new value and continue with the other dual variable μ_{3-n}.

By construction, the value of the partial dual function $q(\boldsymbol{\mu})$ decreases as we run the algorithm. However, it does not imply that the value of the sum of utility functions is increasing. Now, define a new sequence of rate vector $\{\mathbf{R}_{\max}(t)\}$ recursively by $\mathbf{R}_{\max}(1) = \mathbf{R}^*(\boldsymbol{\mu}(1))$, and for $t > 1$, let $\mathbf{R}_{\max}(t)$ be

$$\begin{cases} \mathbf{R}^*(\boldsymbol{\mu}(t)) & \text{if } U(\mathbf{R}^*(\boldsymbol{\mu}(t))) > U(\mathbf{R}_{\max}(t-1)) \\ \mathbf{R}_{\max}(t-1) & \text{if } U(\mathbf{R}^*(\boldsymbol{\mu}(t))) \leq U(\mathbf{R}_{\max}(t-1)). \end{cases} \tag{4.141}$$

In words, $U(\mathbf{R}_{\max}(t))$ records the largest utility value up to iteration t.

The main result in this section is

Theorem 4.17. *The* arg min *in* (4.139)-(4.140) *exists and is finite. The sequence of rate vectors, $\{\mathbf{R}_{\max}(t)\}$, converges, and the sequence of utility values, $\{U(\mathbf{R}_{\max}(t))\}$, converges to the optimal value that maximizes $U(\mathbf{R})$ over all $\mathbf{R} \in \mathcal{C}$.*

In order to prove Theorem 4.17, the following properties about the partial dual function are needed.

Proposition 4.18. *$q(\boldsymbol{\mu})$ is convex, continuously differentiable and*

$$\frac{\partial q(\boldsymbol{\mu})}{\partial \mu_n} = -\tilde{R}_n^*(\boldsymbol{\mu}) + R_n^*(\boldsymbol{\mu}). \tag{4.142}$$

Proof. The fact that $q(\boldsymbol{\mu})$ is convex follows from a basic result in convex function theory that the point-wise maximum of affine functions is convex [11, Section 3.2.3]. The property that $q(\boldsymbol{\mu})$ is differentiable and that the partial derivative is given as in (4.142) are consequences of

Danskin's theorem. By appealing to Berge's maximal theorem, I can show that both $\tilde{R}_n^*(\boldsymbol{\mu})$ and $R_n^*(\boldsymbol{\mu})$ are continuous. Hence $-\tilde{R}_n^*(\boldsymbol{\mu}) + R_n^*(\boldsymbol{\mu})$ is continuous, implying that $q(\boldsymbol{\mu})$ is continuously differentiable. □

In the following proposition, $\tilde{\mathbf{R}}^*(\boldsymbol{\mu})$ and $\mathbf{R}^*(\boldsymbol{\mu})$ are investigated as functions of a single component, keeping the other components fixed. $\boldsymbol{\mu}_{-n}$ represents the vector $\boldsymbol{\mu}$ without the component μ_n, and write $\tilde{\mathbf{R}}^*(\boldsymbol{\mu}) = \tilde{\mathbf{R}}^*(\mu_n, \boldsymbol{\mu}_{-n})$ and $\mathbf{R}^*(\boldsymbol{\mu}) = \mathbf{R}^*(\mu_n, \boldsymbol{\mu}_{-n})$ in order to emphasize that μ_n is the variable, while $\boldsymbol{\mu}_{-n}$ is fixed.

Proposition 4.19. *For any $n = 1, 2$, given any fixed $\boldsymbol{\mu}_{-n}$, $R_n^*(\mu_n, \boldsymbol{\mu}_{-n})$ is an increasing function of μ_n, and $\tilde{R}_n^*(\mu_n, \boldsymbol{\mu}_{-n})$ is a decreasing function of μ_n.*

Proof. Suppose to the contrary that $R_n^*(\mu_n, \boldsymbol{\mu}_{-n})$ is not increasing. I can then find a $\delta > 0$ such that $R_n^*(\mu_n + \delta, \boldsymbol{\mu}_{-n}) < R_n^*(\mu_n, \boldsymbol{\mu}_{-n})$ for some value of μ_n. Define the vector

$$\boldsymbol{\mu}' \triangleq (\mu_n + \delta, \boldsymbol{\mu}_{-n}). \tag{4.143}$$

We have

$$\boldsymbol{\mu}' \cdot \mathbf{R}^*(\boldsymbol{\mu}') = (\mu_n + \delta) R_n^*(\boldsymbol{\mu}') + \sum_{m \neq n} \mu_m R_m^*(\boldsymbol{\mu}'), \tag{4.144}$$

$$= \boldsymbol{\mu} \cdot \mathbf{R}^*(\boldsymbol{\mu}') + \delta R_n^*(\boldsymbol{\mu}'). \tag{4.145}$$

By definition of $\mathbf{R}^*(\boldsymbol{\mu})$, the first term above is less than or equal to $\boldsymbol{\mu} \cdot \mathbf{R}^*(\boldsymbol{\mu})$. The second term is strictly less than $\delta R_n^*(\boldsymbol{\mu})$ by our hypothesis. Hence,

$$\boldsymbol{\mu}' \cdot \mathbf{R}^*(\boldsymbol{\mu}') < \boldsymbol{\mu} \cdot \mathbf{R}^*(\boldsymbol{\mu}) + \delta R_n^*(\boldsymbol{\mu}) \tag{4.146}$$

$$= \boldsymbol{\mu}' \cdot \mathbf{R}^*(\boldsymbol{\mu}). \tag{4.147}$$

This contradicts that $\boldsymbol{\mu}' \cdot \mathbf{R}$ is maximized at $\mathbf{R} = \mathbf{R}^*(\boldsymbol{\mu}')$. This proves the first part of the proposition.

For the second part, suppose to the contrary that $\tilde{R}_n^*(\mu_n, \boldsymbol{\mu}_{-n})$ is not a decreasing function of μ_n. We can then find a $\delta > 0$ such that $\tilde{R}_n^*(\mu_n + \delta, \boldsymbol{\mu}_{-n}) > \tilde{R}_n^*(\mu_n, \boldsymbol{\mu}_{-n})$ for some value of μ_n.

Similar to the first part of the proposition, we have

$$U(\tilde{\mathbf{R}}^*(\boldsymbol{\mu}')) - \boldsymbol{\mu}' \cdot \tilde{\mathbf{R}}^*(\boldsymbol{\mu}') \tag{4.148}$$

$$= U(\tilde{\mathbf{R}}^*(\boldsymbol{\mu}')) - \boldsymbol{\mu} \cdot \tilde{\mathbf{R}}^*(\boldsymbol{\mu}') - \delta \tilde{R}_n^*(\boldsymbol{\mu}'). \tag{4.149}$$

By definition of $\tilde{\mathbf{R}}^*(\boldsymbol{\mu})$,

$$U(\tilde{\mathbf{R}}^*(\boldsymbol{\mu}')) - \boldsymbol{\mu} \cdot \tilde{\mathbf{R}}^*(\boldsymbol{\mu}') \leq U(\tilde{\mathbf{R}}^*(\boldsymbol{\mu})) - \boldsymbol{\mu} \cdot \tilde{\mathbf{R}}^*(\boldsymbol{\mu}). \tag{4.150}$$

Combining with our hypothesis that $\tilde{R}_n^*(\boldsymbol{\mu}') > \tilde{R}_n^*(\boldsymbol{\mu})$, we obtain

$$U(\tilde{\mathbf{R}}^*(\boldsymbol{\mu}')) - \boldsymbol{\mu}' \cdot \tilde{\mathbf{R}}^*(\boldsymbol{\mu}') \tag{4.151}$$

$$< U(\tilde{\mathbf{R}}^*(\boldsymbol{\mu})) - \boldsymbol{\mu} \cdot \tilde{\mathbf{R}}^*(\boldsymbol{\mu}) - \delta \tilde{R}_n^*(\boldsymbol{\mu}) \tag{4.152}$$

$$= U(\tilde{\mathbf{R}}^*(\boldsymbol{\mu})) - \boldsymbol{\mu}' \cdot \tilde{\mathbf{R}}^*(\boldsymbol{\mu}). \tag{4.153}$$

This contradicts that $U(\tilde{\mathbf{R}}^*(\boldsymbol{\mu}')) - \boldsymbol{\mu}' \cdot \tilde{\mathbf{R}}^*(\boldsymbol{\mu}')$ is the maximal value of $U(\mathbf{R}) - \boldsymbol{\mu}' \cdot \mathbf{R}$ over all $\mathbf{R} \in \mathcal{C}$. □

By the last two propositions, the partial derivative of $q(\mu_n, \boldsymbol{\mu}_{-n})$ with respective to μ_n is an increasing function. The search for the arg min in (4.139)-(4.140) can be done by bisection search for instance.

Proposition 4.20. *Suppose that* $\{\boldsymbol{\mu}(t)\}$ *is a sequence of dual variables generated by the proposed algorithm in* (4.139)-(4.140). *Then* $\{\boldsymbol{\mu}(t)\}$ *is contained in a bounded set.*

Proof. By construction, the values of the partial dual function $\{q(\boldsymbol{\mu}(t))\}$ is a decreasing sequence. It suffices to show that there is a bounded set $\mathcal{A} \in \mathbb{R}_+^N$ so that $\mathbf{x} \notin \mathcal{A}$ implies that $q(\mathbf{x}) > q(\boldsymbol{\mu}(1))$.

Let $\mathbf{1}$ be the N-dimensional all-one vector, and ϵ be a positive real number, so that $\epsilon\mathbf{1} \in \mathcal{C}$.

For any $\boldsymbol{\mu} \succeq \mathbf{0}$, the optimal value of the first subproblem is the maximum of $U(\tilde{\mathbf{R}}) - \boldsymbol{\mu} \cdot \tilde{\mathbf{R}}$ taken over $\tilde{\mathbf{R}} \in \mathcal{B}$. In particular, it is bounded from below by $U(\epsilon\mathbf{1}) - \boldsymbol{\mu} \cdot (\epsilon\mathbf{1})$, i.e.,

$$\max_{\tilde{\mathbf{R}} \in \mathcal{B}} (U(\tilde{\mathbf{R}}) - \boldsymbol{\mu} \cdot \tilde{\mathbf{R}}) \geq U(\epsilon\mathbf{1}) - \epsilon\boldsymbol{\mu} \cdot \mathbf{1}. \tag{4.154}$$

For the second subproblem, I use the property that the optimal value is a homogeneous function of $\boldsymbol{\mu}$ of degree 1, i.e., for any $\lambda > 0$, we have

$$\max_{\mathbf{R}\in\mathcal{C}}(\lambda\boldsymbol{\mu})\cdot\mathbf{R} = \lambda\max_{\mathbf{R}\in\mathcal{C}}\boldsymbol{\mu}\cdot\mathbf{R}. \tag{4.155}$$

Let H_1 be the simplex

$$H_1 \triangleq \Big\{\boldsymbol{\mu}\succeq\mathbf{0} : \sum_{n=1}^{N}\mu_n = 1\Big\}. \tag{4.156}$$

Consider $\max_{\mathbf{R}\in\mathcal{C}}\boldsymbol{\mu}\cdot\mathbf{R}$ as a function of $\boldsymbol{\mu}$ with $\boldsymbol{\mu}$ restricted to H_1. Since H_1 is a compact set and the optimal value $\max_{\mathbf{R}\in\mathcal{C}}\boldsymbol{\mu}\cdot\mathbf{R}$ is a continuous function of $\boldsymbol{\mu}$ (by Berge's maximum theorem), we can find a point in H_1 such that the minimum is attained. Let the minimum value be denoted by m,

$$m = \min_{\boldsymbol{\mu}\in H_1}\max_{\mathbf{R}\in\mathcal{C}}\boldsymbol{\mu}\cdot\mathbf{R}. \tag{4.157}$$

For each $\boldsymbol{\mu}\succeq\mathbf{0}$, $\boldsymbol{\mu}\neq\mathbf{0}$, by the property that $\max_{\mathbf{R}}\boldsymbol{\mu}\cdot\mathbf{R}$ is homogeneous, we obtain

$$\max_{\mathbf{R}\in\mathcal{C}}\boldsymbol{\mu}\cdot\mathbf{R} \geq (\sum_{n}\mu_n)m \tag{4.158}$$

It follows that

$$q(\boldsymbol{\mu}) \geq \big[U(\epsilon\mathbf{1}) - \epsilon\boldsymbol{\mu}\cdot\mathbf{1}\big] + (\boldsymbol{\mu}\cdot\mathbf{1})m \tag{4.159}$$

$$= U(\epsilon\mathbf{1}) + (\boldsymbol{\mu}\cdot\mathbf{1})\big[-\epsilon + m\big]. \tag{4.160}$$

Note that m is a positive constant. (It is zero only if all points in $\mathcal{C}$ have some component identically zero. This degenerate case is not of practical interest.) By further decreasing ϵ, we can pick a sufficiently small ϵ such that $0 < -\epsilon + m$. Then, the value within the square bracket in (4.160) is positive. When $\boldsymbol{\mu}\cdot\mathbf{1}$ is sufficiently large, $q(\boldsymbol{\mu})$ is strictly larger than $q(\boldsymbol{\mu}(1))$. Therefore, by picking a sufficiently large constant λ, we have $q(\boldsymbol{\mu}) > q(\boldsymbol{\mu}(1))$ for all $\boldsymbol{\mu}\cdot\mathbf{1} > \lambda$. Now, I consider

$$\mathcal{A} = \Big\{\boldsymbol{\mu}\succeq\mathbf{0} : \sum_{n=1}^{N}\mu_n \leq \lambda\Big\}. \tag{4.161}$$

Then $\mathcal{A}$ has the desired property that $q(\mathbf{x}) > q(\boldsymbol{\mu}(1))$ for all $\mathbf{x}\notin\mathcal{A}$. □

Now, the main theorem can be proved.

Proof of Theorem 4.17. By Prop. 4.20, the minimum value in each iteration of the Gauss-Seidel algorithm is taken in a compact set. Since the minimum value of a continuous function in a compact domain is indeed attained by a point in the domain, the arg min in (4.139)-(4.140) exists. By Prop. 4.20 again, it is finite.

For the convergence, I apply the following result from [33]:

(Convergence of Gauss-Seidel Algorithm) Suppose that $\mathcal{X} \in \mathbb{R}^N$ is a non-empty closed and convex set, which is a Cartesian product $\prod_{i=1}^{N} \mathcal{X}_i$ with $\mathcal{X}_i \in \mathbb{R}$. Suppose that $f(x_1, \ldots, x_N)$ is a real-valued function such that (a) it is continuously differentiable and convex on $\mathcal{X}$ and (b) the set $\{\mathbf{x} \in \mathcal{X} : f(\mathbf{x}) \leq f(\mathbf{x}_0)\}$ is compact for any $\mathbf{x}_0 \in \mathcal{X}$, then the sequence generated by the Gauss-Seidel algorithm

$$\begin{aligned} x_n(t+1) = \arg\min_{\xi \in \mathcal{X}_n} f(&x_1(t+1), \ldots, x_{n-1}(t+1), \\ &\xi, x_{n+1}(t), \ldots, x_N(t)) \end{aligned} \tag{4.162}$$

has limit points and each limit point minimizes f.

We can check that the conditions (a) and (b) are satisfied using Prop. 4.18 and 4.20. By the above convergence result, I conclude that $\lim_{t\to\infty} q(\boldsymbol{\mu}(t))$ is the minimum value, say v^* of the dual problem. By the standard duality theory in convex analysis, $\sup_{t\to\infty} U(\mathbf{R}^*(\boldsymbol{\mu}(t)))$ is the maximum value, which equals v^*, of the primal problem. The sequence $U(\mathbf{R}_{\max}(t))$ is monotonically increasing, approaching the maximum value v^*,

$$\lim_{t\to\infty} q(\boldsymbol{\mu}(t)) = v^* = \lim_{t\to\infty} U(\mathbf{R}_{\max}(t)). \tag{4.163}$$

Since U is strictly concave, v^* is achieved by a unique point in $\mathcal{C}$. Let this maximizer be denoted by $\mathbf{R}^*$, i.e., $U(\mathbf{R}^*) = v^*$. I will show that the sequence $\{\mathbf{R}_{\max}(t)\}$ converges to $\mathbf{R}^*$ by contradiction.

Let $B_\epsilon(\mathbf{R}^*)$ be the sphere with center $\mathbf{R}^*$ and radius ϵ. If $\{\mathbf{R}_{\max}(t)\}$ does not converge to $\mathbf{R}^*$, then there exists an $\epsilon > 0$ such that $\|\mathbf{R}_{\max}(t) - \mathbf{R}^*\|_2 \geq \epsilon$ for infinitely many t. Let $t_1 < t_2 < t_3 < \ldots$ be an increasing sequence of integers such that $\|\mathbf{R}_{\max}(t_k) - \mathbf{R}^*\|_2 \geq \epsilon$, for all k. Since the sequence $\{\mathbf{R}_{\max}(t_k)\}$ is contained in a compact set $\mathcal{C}$, it contains a subsequence converging to a limit, say $\mathbf{R}_0$. As $U(\mathbf{R}_{\max}(t_k)) \to v^*$ as $k \to \infty$, by the continuity of U, we have $U(\mathbf{R}_0) = v^*$. Since the maximum is achieved by a unique point, it follows that $\mathbf{R}_0 = \mathbf{R}^*$. However, the metric $d(\mathbf{x}, \mathbf{R}^*)$ defined as $\|\mathbf{x} - \mathbf{R}^*\|_2$ is a continuous function of $\mathbf{x}$. It follows that $\|\mathbf{R}_0 - \mathbf{R}^*\|_2 \geq \epsilon$. This is a contradiction and completes the proof of the convergence of $\{\mathbf{R}_{\max}(t)\}$. □

4.7.3 Numerical Example

In this section, the harmonic mean fairness problem is used to illustrate the idea of the joint utility maximization algorithm.

By equating the derivative of $-\frac{1}{\tilde{R}_1} - \frac{1}{\tilde{R}_2} - \mu_1\tilde{R}_1 - \mu_2\tilde{R}_2$ to zero and then projecting the solution to the closed interval $[0, b_n]$, we have the following optimal choice function $\tilde{R}_n^*$,

$$\tilde{R}_n^*(\mu_n, \boldsymbol{\mu}_{-n}) = \min\left\{b_n, \frac{1}{\sqrt{\mu_n}}\right\}. \tag{4.164}$$

Here, b_1 and b_2 are chosen to be $R_1^*(1,0)$ and $R_2^*(0,1)$ respectively. By Proposition 4.3, $R_n^*(\mu_n, \boldsymbol{\mu}_{-n})$ is well-defined and can be computed by algorithm in Table 4.1. Hence, in each iteration of the Gauss-Seidel algorithm, we solve for μ_n in the following equation.

$$\tilde{R}_n^*(\mu_n, \boldsymbol{\mu}_{-n}) = R_n^*(\mu_n, \boldsymbol{\mu}_{-n}). \tag{4.165}$$

The power gain of each link is exponentially and independently distributed with mean 1. The transmission power of each source node is 10 and the noise power at the receiving nodes is assumed to be 1. The results for a sample run are illustrated in Figure 4.8 and 4.9. The sum of utility functions at each iteration is plotted in Figure 4.8. Unlike the previous example, $U(\mathbf{R}^*(t))$ is monotonically increasing. From Figure 4.9, we observe that the rates of the node pairs converge within 5 iterations.

4.8 Conclusion

Cooperation between source nodes are exploited for interference Gaussian channel in this chapter. I study two cooperative transmission schemes where the source nodes occupy disjoint bandwidth, and thus removing all interference. Simulation results show that in some

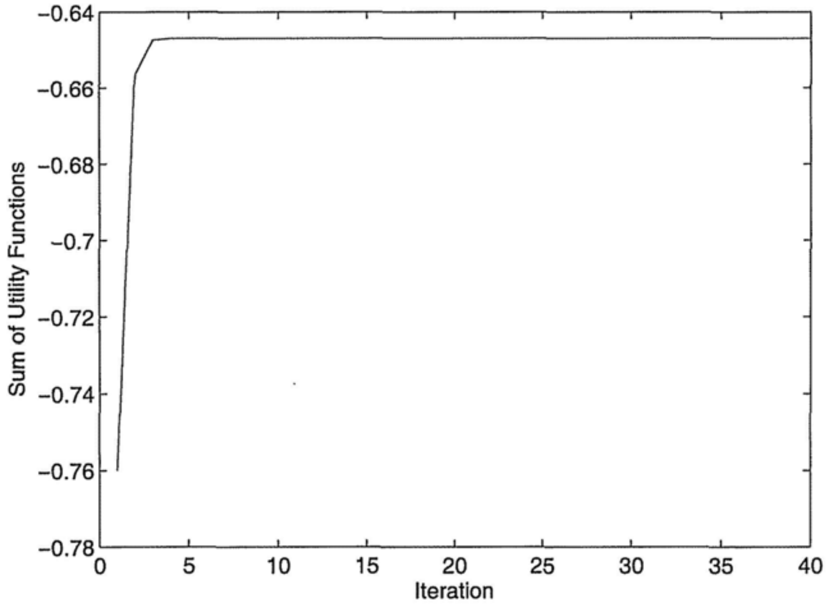

Figure 4.8: $U(\mathbf{R}^*(t))$, sum of utility functions as a function of iteration in cooperative transmission case.

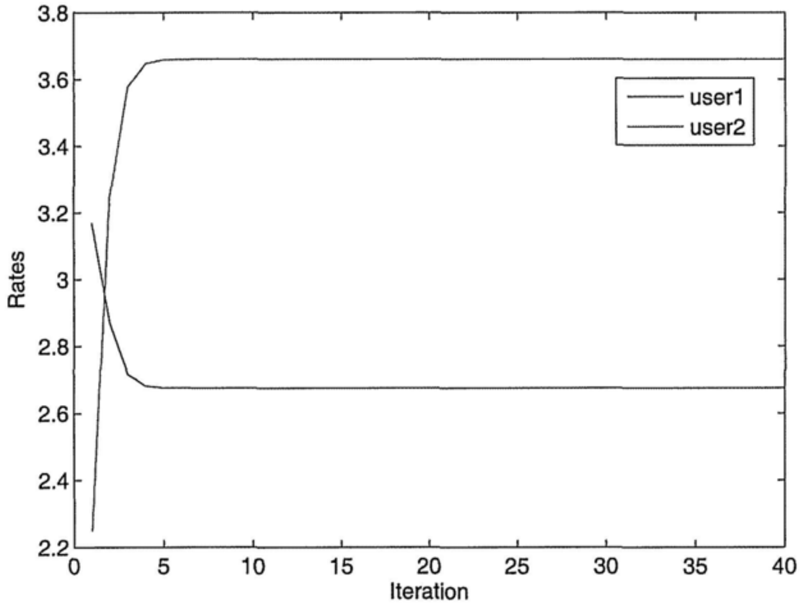

Figure 4.9: The rate of each user, $\mathbf{R}^*(\mu(t))$ in cooperative transmission case

channel realizations, the interference channel capacity region is a subset of the achievable rate regions of the two proposed schemes. Furthermore, one of the cooperative transmission schemes can achieve twice the order of diversity than a simplified Han-Koybayashi scheme. I devise an efficient iterative algorithm that computes the maximum weighted sum rate. Furthermore, the achievable rate region in the low SNR regime is characterized and the aforementioned weighted sum rate maximization problem becomes much simpler. The proposed algorithm can be applied to achieve max-min fairness. Finally, I demonstrate how to extend the proposed algorithm to solve a more general joint utility maximization problem.

Chapter 5

Pricing Game for Distributed Cooperative Transmission

In this chapter, two pricing games for two cooperative transmission schemes are devised. The objective is to minimize the total transmission power. Based on the game formulations, the adaptive and distributed implementations for the two cooperative transmission schemes are derived.

In the first game, the outgoing links of a node are considered to be a degraded Gaussian broadcast channel, where all Pareto-optimal rates can be achieved by superposition coding, which is the extension of Scheme 1 described in Chapter 4. In [9], it is shown that the achievable rate region by allocating orthogonal sub-channels to the outgoing links is enclosed by the capacity region of a broadcast channel. However, implementing superposition codes is more complex. A natural question is: if the source nodes divide their own channels into orthogonal sub-channels for their outgoing links instead of multiplexing the data streams by superposition codes, how much extra power do we need to trade for the reduction of complexity? This is the motivation of the second game where the nodes perform time division multiplex (TDM) over their outgoing links.

With the best response update algorithms for the two proposed games, we are able to compare the performance of the two corresponding cooperative transmission schemes through simulations. Both cooperative transmission schemes are able to achieve full order of spatial diversity. Also, we observe the rapid convergence of both games.

This chapter is organized as follows. The system model is provided in Section 6.1. In Section 5.2, we describe the cooperative transmission scheme based on superposition codes. The TDM-based cooperative transmission scheme is detailed in Section 5.3. In Section 5.4, we compare the performance of our proposed algorithms via simulations.

Part of the contents in this chapter can also be found in [73, 74]

5.1 System Model

Consider a wireless network with M distinct node pairs. The source and the destination nodes of the i-th node pair are denoted by S_i and D_i respectively. The link from S_i to S_j and the link from S_i to D_j are denoted by (i, S_j) and (i, D_j) respectively.

Let Z_{i,S_j} and Z_{i,D_j} be the power gains of (i, S_j) and (i, D_j) respectively. Notice that since the power gains are real numbers, the probability that two links have equal power gains is zero. The nodes know the power gains of their incoming and outgoing links only. The noise of the links is assumed to be independent additive white Gaussian noise (AWGN) with one-sided power spectral density N_0.

Every source node transmits over its assigned channel which is orthogonal to other channels. For simplicity, each channel is assumed to have an equal bandwidth W.[1] Without loss of generality, it is assumed that channel i is assigned to S_i.

Apart from their own messages, the source nodes can overhear and relay others' messages from other channels. Each source node S_i divides its own data stream into M sub-streams. The first one is transmitted directly to D_i. The remaining $M-1$ sub-streams are forwarded by other source nodes. These $M-1$ sub-streams are decoded, buffered, re-encoded and then transmitted by the relaying source nodes. Hence, besides its own M sub-streams of data, each source node, in general, forwards $M-1$ data sub-streams originated from other nodes. In this case, each channel is regarded as a degraded Gaussian broadcast channel with

[1] The results in this chapter can be easily generalized to the uneven bandwidth case

$2M - 1$ receivers. As mentioned in Section 2.1, all Pareto optimal rates can be achieved by superposition codes on the sender side and successive interference cancelation on the receiver side. Alternatively, we may trade the throughput for simplicity in implementation by multiplexing the outgoing links with TDM. Details and comparisons of these two methods are provided in succeeding sections. In both cases, the destination nodes must listen to all channels to recover their messages.

Let P_i be the transmission power of S_i. Our objective is to minimize the total transmission power, which is the sum of P_i's of both cooperative transmission schemes.

5.2 Superposition Code Based Scheme

The first scheme is the *superposition code based scheme* which is a natural extension of the two-user scheme described in Chapter 4.

5.2.1 Minimum Sum Transmit Power Problem

Let R_{i,S_j} and R_{i,D_j} be the rates[2] of (i, S_j) and (i, D_j) respectively. The rate vector in channel i is denoted as $(\mathbf{R}_i^{(S)}, \mathbf{R}_i^{(D)})$ where $\mathbf{R}_i^{(S)} = (R_{i,S_1}, R_{i,S_2}, \dots, R_{i,S_{i-1}}, R_{i,S_{i+1}}, \dots, R_{i,S_M})$ and $\mathbf{R}_i^{(D)} = (R_{i,D_1}, R_{i,D_2}, \dots, R_{i,D_M})$.

Let $\mathcal{C}_i(P_i)$ be the capacity region of channel i, which means $(\mathbf{R}_i^{(S)}, \mathbf{R}_i^{(D)}) \in \mathcal{C}_i(P_i)$. Let $\Gamma_i = \frac{P_i}{N_0 W}$ be the transmit signal-to-noise ratio (SNR) of S_i. The minimum required Γ_i is given by this lemma, which is derived by induction from the degraded Gaussian broadcast channel capacity region.

Lemma 5.1. *Let $\mathcal{V} = \{S_1, S_2, \dots, S_{i-1}, S_{i+1}, \dots, S_M, D_1, D_2, \dots, D_M\}$. For each source node S_i, we define a function $\pi_i : \{1, 2, \dots, 2M-1\} \to \mathcal{V}$ such that $Z_{i,\pi_i(1)} < Z_{i,\pi_i(2)} < \dots < Z_{i,\pi_i(2M-1)}$. Given $(\mathbf{R}_i^{(S)}, \mathbf{R}_i^{(D)}) \in \mathcal{C}_i(P_i)$, the minimum required Γ_i is*

$$\Gamma_i = \sum_{j=1}^{2M-1} \frac{1}{Z_{i,\pi_i(j)}} \exp\left(\sum_{k=j+1}^{2M-1} R_{i,\pi_i(k)}\right) \left[\exp\left(R_{i,\pi_i(j)}\right) - 1\right] \tag{5.1}$$

[2]In fact, the word "rates" means "spectral efficiency", which has the unit nats/s/Hz This terminology is used for convenience of presentation in this chapter

Let R_i be the minimum required rate for the node pair S_i and D_i. The rate constraint is

$$R_{i,D_i} + \sum_{j=1,j\neq i}^{M} R_{i,S_j} \geq R_i. \tag{5.2}$$

The minimum sum power problem, which is equivalent to the minimum sum transmit SNR problem below, is considered.

Problem 5.1.

$$\min_{\mathbf{R}_i \geq 0} \sum_{i=1}^{M} \Gamma_i \tag{5.3}$$

subject to

$$R_{i,D_i} + \sum_{j=1,j\neq i}^{M} R_{i,S_j} \geq R_i, \quad i = 1, 2, \ldots, M \tag{5.4}$$

$$R_{i,D_j} = R_{j,S_i}, \quad i \neq j \tag{5.5}$$

The final constraint is the flow conservation constraint at the relaying source node S_i. By Lemma 5.1, Γ_i is strictly convex and continuous. Also, the constraint set is convex and compact. Thus, Problem 5.1 is a convex optimization problem [11] and the optimal solution exists.

5.2.2 Pricing Game Formulation

Let $\mathbf{R}_i$ be $(R_{i,D_i}, \mathbf{R}_i^{(S)})$ and $\mathbf{R}_{-i} = (\mathbf{R}_1, \mathbf{R}_2, \ldots, \mathbf{R}_{i-1}, \mathbf{R}_{i+1}, \ldots, \mathbf{R}_M)$. Problem 5.1 can be solved by playing a best-reply potential game (see Chapter 2) with the potential function

$$\Phi(\mathbf{R}_1, \mathbf{R}_2, \ldots, \mathbf{R}_M) = -\sum_{i=1}^{M} \Gamma_i. \tag{5.6}$$

The main idea of solving Problem 5.1 as a best-reply potential game is that $\mathbf{R}_i$'s are optimized in a round robin manner. By removing the terms in (5.6) which are independent of $\mathbf{R}_i$ and substituting $R_{k,D_i} = R_{i,S_k}$ in Γ_k for $k \neq i$, we can optimize $\mathbf{R}_i$ for fixed $\mathbf{R}_{-i}$ from the following problem.

Problem 5.2.

$$\max_{\mathbf{R}_i} u_i(\mathbf{R}_i, \mathbf{R}_{-i}) = -\Gamma_i - \sum_{j=1, j\neq i}^{M} K_{i,j} \exp\left(R_{i,S_j}\right) + \sum_{j=1, j\neq i}^{M} \frac{1}{Z_{j,D_i}} \exp\left(\sum_{k=\pi_j^{-1}(D_i)+1}^{2M-1} R_{j,\pi_j(k)}\right) \tag{5.7}$$

subject to

$$R_{i,D_i} + \sum_{j=1, j\neq i}^{M} R_{i,S_j} \geq R_i \tag{5.8}$$

$$0 \leq R_{i,D_i}, R_{i,S_j} \leq R_i, \quad \forall j \neq i \tag{5.9}$$

where for $j \neq i$,

$$K_{i,j} = \sum_{k=1}^{\pi_j^{-1}(D_i)-1} \frac{1}{Z_{j,\pi_j(k)}} \exp\left(\sum_{k'=k+1, k'\neq \pi_j^{-1}(D_i)}^{2M-1} R_{j,\pi_j(k')}\right) \left[\exp\left(R_{j,\pi_j(k)}\right) - 1\right]$$
$$+ \frac{1}{Z_{j,D_i}} \exp\left(\sum_{k=\pi_j^{-1}(D_i)+1}^{2M-1} R_{j,\pi_j(k)}\right) \tag{5.10}$$

In (5.7), the j-th term of the second summation can be interpreted as the price S_i needs to 'pay' to S_j for a relay rate R_{i,S_j}. That is, S_i 'pays' an amount of $\overline{K}_{i,j} = K_{i,j}\exp\left(R_{i,S_j}\right)$ "virtual money" to S_j of relaying the message at the rate $R_{i,S_j} = \ln\left(\frac{\overline{K}_{i,j}}{K_{i,j}}\right)$ to D_i. It is called virtual money because this pricing interpretation is only used to derive the distributed implementation. In the actual implementation described later, no forms of payments will be made. The last summation in (5.7) is independent of $\mathbf{R}_i$ and it is introduced for the convenience of setting up the best-reply potential game.

More specifically, consider the following game $G = [\mathcal{P}, \{\mathcal{R}_i\}, \{u_i\}]$. The player set $\mathcal{P}$ is the set of source nodes, $\{S_1, S_2, \ldots, S_M\}$. The strategy set of S_i, which is denoted by $\mathcal{R}_i$, is the feasible set of $\mathbf{R}_i$ given by (5.8) and (5.9) for $i = 1, 2, \ldots, M$ and the rate vector $\mathbf{R}_i$ is adjusted by S_i. The utility function of the player S_i is $u_i(\mathbf{R}_i, \mathbf{R}_{-i})$. The solution of Problem 5.2, which maximizes the utility function u_i, is the best response update of the player S_i. It is denoted by $B_i(\mathbf{R}_{-i})$. Since u_i is a strictly concave function of $\mathbf{R}_i$, the best response update is unique. Thus,

$$B_i(\mathbf{R}_{-i}) = \arg\max_{\mathbf{R}_i \in \mathcal{R}_i} u_i(\mathbf{R}_i, \mathbf{R}_{-i}) \tag{5.11}$$

is a well-defined function of $\mathbf{R}_{-i}$. Furthermore, as u_i is a continuous function of $\mathbf{R}_i$ and $\mathbf{R}_{-i}$ and the best response update is unique, by Maximum Theorem (see Section 2.3), B_i is a continuous function of $\mathbf{R}_{-i}$.

Notice that for fixed $\mathbf{R}_{-i}$, $u_i(\mathbf{R}_i, \mathbf{R}_{-i})$ is formed by deleting some of the terms in (5.3) which are independent of $\mathbf{R}_i$. Hence, a best response update by S_i is a minimization of total transmit SNR by choosing $\mathbf{R}_i$ for fixed $\mathbf{R}_{-i}$. Thus, the optimal solution of Problem 5.1 is a pure strategy Nash Equilibrium.

The following proposition shows how Problem 5.1 can be solved by playing the best-reply potential game.

Proposition 5.2. *Any admissible sequential best-reply path of the pricing game G converges to the unique pure strategy Nash Equilibrium.*

Proof. By Theorem 2.1 and [47, Theorem 2], this proposition is true if the strategy sets are compact, $B_i(\mathbf{R}_{-i})$ is continuous and there is a unique pure strategy Nash Equilibrium. The strategy sets are obviously compact and as mentioned above, $B_i(\mathbf{R}_{-i})$ is continuous. Thus, we only need to prove the existence and uniqueness of the pure strategy Nash Equilibrium.

As mentioned above, the optimal solution of Problem 5.1 is a pure strategy Nash Equilibrium of the game G. Now, we prove the uniqueness part. It is stated in [88, Theorem 3] that in a potential game, if the strategy sets are compact and Φ is continuously differentiable function in the interior of the strategy set and strictly concave in the strategy set, the Nash Equilibrium is unique. Φ satisfies all these conditions and hence the Nash Equilibrium, which is a pure strategy Nash Equilibrium and is the optimal solution of Problem 5.1, is unique. □

By playing the game G, we can reach the optimal solution of Problem 5.1 at Nash Equilibrium. One way to play this game is outlined in Table 5.1. Proposition 5.2, in fact, guarantees that a large number of ways of playing the game G can converge to optimal solution of Problem 5.1. The only requirement is that the sequential best-reply update path is admissible.

Table 5.1: Pricing Game for Power Minimization in Superposition Code Based Scheme

$i \leftarrow 1$
while Nash Equilibrium has not been reached **do**
  for all $j \neq i$ **do**
    if $K_{i,j}^{(3)}$ is changed **then**
      D_i notifies S_i the new value of $K_{i,j}^{(3)}$.
    end if
  end for
  S_i maximizes $u_i(\mathbf{R}_i, \mathbf{R}_{-i})$ by optimizing $\mathbf{R}_i$ with algorithm in Table 5 2
  $i \leftarrow i + 1$
  $i \leftarrow 1$ **if** $i = M + 1$
end while

Therefore, it is very flexible to choose the way how the game is played. For example, if after the best response update of a player S_i, S_j's transmission power increases drastically, S_j does not need to wait for its iteration to perform the best response update to reduce its transmission power. The best choice of the flow of the game is beyond the scope of the thesis. In this chapter, only the flow specified in Table 5.1 is considered.

5.2.3 Distributed Implementation

Now, I describe how the source nodes play the game in a distributed manner. In this game, the source node S_i is responsible to allocate the rates of direct and relay paths and the corresponding destination node D_i is responsible to inform S_i any changes of the pricing information $K_{i,j}$'s which cannot be detected by S_i.

Firstly, we observe that the pricing information $K_{i,j}$ can be expressed as follows.

$$K_{i,j} = \underbrace{\exp\left(\sum_{k=\pi_j^{-1}(D_i)+1}^{2M-1} R_{j,\pi_j(k)}\right)}_{K_{i,j}^{(1)}} \left\{ \underbrace{\sum_{k=1}^{\pi_j^{-1}(D_i)-1} \frac{1}{Z_{j,\pi_j(k)}} \exp\left(\sum_{k'=k+1}^{\pi_j^{-1}(D_i)-1} R_{j,\pi_j(k')}\right) \left[\exp\left(R_{j,\pi_j(k)}\right) - 1\right]}_{K_{i,j}^{(2)}} + \frac{1}{Z_{j,D_i}} \right\} \tag{5.12}$$

In superposition code, each frame header consists of the rate of each receiver so that the receivers know which codebook is adopted. Therefore, $K_{i,j}^{(1)}$ can be computed by the rates of the corresponding links specified in the frame header. Therefore, D_i only needs to notify S_i the changes of $K_{i,j}^{(2)}$ and Z_{j,D_i}.

The destination node D_i measures the received power of the preamble of the frames from S_j and then infers the power gain of the link (j, D_i), which is Z_{j,D_i}. Next, by proving in a similar way as Lemma 5.1, it can be showed that $K_{i,j}^{(2)}$ is the total transmitted SNR of the messages over the links $(j, \pi_j(k))$, where $\pi_j^{-1}(D_i) + 1 \leq k \leq 2M - 1$. $Z_{j,D_i} K_{i,j}^{(2)}$ is in fact the received interference-to-noise ratio at D_i after its successive interference cancelation. That is, after canceling the signals for the nodes $\pi_j(k)$, $\pi_j^{-1}(D_i) < k \leq 2M - 1$, from the received signals from S_j, the total power of the remaining interference to D_i is $Z_{j,D_i} K_{i,j}^{(2)} N_0 W$.

After the successive interference cancelation described above, D_i measures the power of the remaining signal and we let this signal power be $\tilde{P}_{i,j}$. D_i can compute $K_{i,j}^{(2)}$ by

$$K_{i,j}^{(2)} = \frac{1}{Z_{j,D_i}} \left[\frac{\tilde{P}_{i,j}}{\exp\left(R_{j,D_i}\right) N_0 W} - 1 \right]. \tag{5.13}$$

By the above observations, the source nodes can allocate the rates adaptively when the power gains of some links or some of the rate requirements, R_j's, change. As mentioned above, S_i only needs to concern about whether $K_{i,j}^{(3)} = K_{i,j}^{(2)} + \frac{1}{Z_{j,D_i}}$ or the power gains of its outgoing links have changed. When $K_{i,j}^{(3)}$ changes, D_i sends a control message about the new value of $K_{i,j}^{(3)}$ to S_i before the best response update. If the power gain of a link changes, the corresponding receiving node notifies the sending node about the new value by a control channel message. When the power gains or the rate requirements change, the nodes perform their best response update based on the new pricing and channel state information. In this case, we have a new admissible best-reply path and by Proposition 5.2, we can reach the new Nash Equilibrium. In Section 5.4, it will be shown that the convergence rate to the Nash Equilibrium is high enough so that the nodes can adapt to the changes of the power gains and rate requirements in a distributive manner.

On the other hand, in a centralized implementation, a centralized node has to be informed when the power gains or rate requirements change. Then, it performs a centralized optimization for Problem 5.1 and notifies every source node S_i the optimal $\mathbf{R}_i$. If the channel conditions fluctuate frequently, the communication overhead will be much higher.

5.2.4 Best Response Update

Problem 5.2 is solved by considering its dual and solve it with the algorithm in Table 5.2. Relax (5.8) to form the following partial Lagrangian.

$$L_i(R_{i,D_i}, \mathbf{R}_i^{(S)}, \lambda) = -u_i(\mathbf{R}_i, \mathbf{R}_{-i}) - \lambda\left(R_{i,D_i} + \sum_{j=1, j\neq i}^{M} R_{i,S_j} - R_i\right) \tag{5.14}$$

where $\lambda \geq 0$. The corresponding partial dual function is

$$q_i(\lambda) = \min_{0 \leq R_{i,D_i}, R_{i,S_j} \leq R_i} L_i(R_{i,D_i}, \mathbf{R}_i^{(S)}, \lambda) \tag{5.15}$$

and the dual problem is

Problem 5.3. $\max_{\lambda \geq 0} q_i(\lambda)$.

$q_i(\lambda)$ is computed by optimizing R_{i,D_i} and R_{i,S_j}'s in a round-robin way. Each variable is optimized by equating the partial derivative of L_i with respect to it to zero and then projecting it to the interval $[0, R_i]$. Thus, the optimal R_{i,D_i} and R_{i,S_j} are

$$R_{i,D_i} = \left[\ln\left(\frac{\lambda}{\Upsilon_{i,i}^{(D)}(\mathbf{R}_i^{(S)})}\right)\right]_i^+ \tag{5.16}$$

$$R_{i,S_j} = \left[\ln\left(\frac{\lambda}{\Upsilon_{i,j}^{(S)}(R_{i,D_i}, \mathbf{R}_{i,-j}^{(S)})}\right)\right]_i^+ . \tag{5.17}$$

where $[x]_i^+ = \min\{\max\{x,0\}, R_i\}$ and

$$\Upsilon_{i,i}^{(D)}(\mathbf{R}_i^{(S)}) = \sum_{j=1}^{\pi_i^{-1}(D_i)-1} \frac{1}{Z_{i,\pi_i(j)}} \exp\left(\sum_{k=j+1,k\neq\pi_i^{-1}(D_i)}^{2M-1} R_{i,\pi_i(k)}\right)\left[\exp\left(R_{i,\pi_i(j)}\right)-1\right] + \frac{1}{Z_{i,D_i}} \exp\left(\sum_{k=\pi_i^{-1}(D_i)+1}^{2M-1} R_{i,\pi_i(k)}\right) \tag{5.18}$$

$$\Upsilon_{i,j}^{(S)}(R_{i,D_i}, \mathbf{R}_{i,-j}^{(S)}) = \sum_{k=1}^{\pi_i^{-1}(S_j)-1} \frac{1}{Z_{i,\pi_i(k)}} \exp\left(\sum_{k'=k+1,k\neq\pi_i^{-1}(S_j)}^{2M-1} R_{i,\pi(k')}\right)\left[\exp\left(R_{i,\pi_i(k)}\right)-1\right] + \frac{1}{Z_{i,S_j}} \exp\left(\sum_{k=\pi_i^{-1}(S_j)+1}^{2M-1} R_{i,\pi_i(k)}\right) + K_{i,j}. \tag{5.19}$$

The following proposition shows that the inner while-loop of the algorithm in Table 5.2 indeed computes $q_i(\lambda)$ for a given λ.

Proposition 5.3. *Given the value of λ. The inner while-loop of the algorithm in Table 5.2 converges to $q_i(\lambda)$.*

Proof. This algorithm is a Gauss-Seidel algorithm. By [33, Proposition 6], it minimizes the partial Lagrangian $L_i(R_{i,D_i}, \mathbf{R}_i^{(S)}, \lambda)$ for a fixed λ if L_i is pseudoconvex[3] on the feasible set of $(R_{i,D_i}, \mathbf{R}_i^{(S)})$ and the corresponding level sets are compact.

By Lemma 5.1, Γ_i is a differentiable and strictly convex function of $(R_{i,D_i}, \mathbf{R}_i^{(S)})$. Therefore, $L_i(R_{i,D_i}, \mathbf{R}_i^{(S)}, \lambda)$ is a pseudoconvex function of $(R_{i,D_i}, \mathbf{R}_i^{(S)})$. Also, since $L_i(R_{i,D_i}, \mathbf{R}_i^{(S)}, \lambda)$ is differentiable, it is lower semicontinuous. By [84, Theorem 7.1], the level sets of $L_i(R_{i,D_i}, \mathbf{R}_i^{(S)}, \lambda)$ are closed. Since we have a bounded feasible set, the level sets are bounded. Therefore, the level sets are compact and result follows. □

We move on to solve Problem 5.3. Since the partial Lagrangian $L_i(R_{i,D_i}, \mathbf{R}_i^{(S)}, \lambda)$ is a strictly convex function of $\mathbf{R}_i$, the optimal R_{i,D_i}, denoted by $R^*_{i,D_i}(\lambda)$, and the optimal R_{i,S_j}, denoted by $R^*_{i,S_j}(\lambda)$, are unique for any fixed λ. By the Envelope Theorem [28], $q_i(\lambda)$ is

[3] A differentiable function f if it satisfies the property $\nabla f(x)(y-x) \geq 0$ implies $f(y) \geq f(x)$ Note that if a function g is differentiable and convex, it is pseudoconvex

Table 5.2: Best Response Update of S_i

while λ has not converged **do**
  $\lambda \leftarrow (\lambda_{\min} + \lambda_{\max})/2$
  Initialize R_{i,D_i} and $\mathbf{R}_i^{(S)}$ with feasible values
  while R_{i,D_i} or $\mathbf{R}_i^{(S)}$ has not converged **do**
    $R_{i,D_i} \leftarrow \left[\ln\left(\frac{\lambda}{\Upsilon_{i,i}^{(D)}(\mathbf{R}_i^{(S)})}\right)\right]_i^+$.
    for all $j \neq i$ and $1 \leq j \leq M$ **do**
      $R_{i,S_j} \leftarrow \left[\ln\left(\frac{\lambda}{\Upsilon_{i,j}^{(S)}(R_{i,D_i}, \mathbf{R}_{i,-j}^{(S)})}\right)\right]_i^+$.
    end for
  end while
  if $R_{i,D_i} + \sum_{j=1, j\neq i}^{M} R_{i,S_j} > R_i$ **then**
    $\lambda_{\max} \leftarrow \lambda$
  else
    $\lambda_{\min} \leftarrow \lambda$
  end if
end while

differentiable and

$$\frac{dq_i}{d\lambda} = R_i - R_{i,D_i}^*(\lambda) - \sum_{j=1, j\neq i}^{M} R_{i,S_j}^*(\lambda). \tag{5.20}$$

Due to the concavity of $q_i(\lambda)$, it can be maximized by solving $\frac{dq_i}{d\lambda} = 0$. Let λ^* be the optimal λ. The following proposition implies that λ^* can be obtained by bisection method, which is the outer loop of the algorithm in Table 5.2.

Proposition 5.4. *$\frac{dq_i}{d\lambda}$ is a decreasing function of λ and $\lambda_{\min} \leq \lambda^* \leq \lambda_{\max}$, where*

$$\lambda_{min} = \min\left\{\frac{1}{Z_{i,D_i}}, \frac{1}{Z_{i,S_j}} + K_{j,i}\right\}, \tag{5.21}$$

$$\lambda_{max} = \max\left\{\Upsilon_{i,i}^{(D)}(\tilde{\mathbf{R}}_i^{(S)})\exp(R_i), \Upsilon_{i,j}^{(S)}(R_i, \tilde{\mathbf{R}}_{i,-j}^{(S)})\exp(R_i)\right\}, \tag{5.22}$$

and $\tilde{\mathbf{R}}_i^{(S)}$ and $\tilde{\mathbf{R}}_{i,-j}^{(S)}$ are $\mathbf{R}_i^{(S)}$ and $\mathbf{R}_{i,-j}^{(S)}$ with all the components equal to R_i respectively.

Proof. The first part of the proposition is the direct consequence of the concavity of $q_i(\lambda)$. Now, we prove the second part. It can be verified that if $\lambda = \lambda_{\min}$, R^*_{i,D_i} and all R^*_{i,S_j} are equal to zero. If $\lambda = \lambda_{\max}$, R^*_{i,D_i} and all R^*_{i,S_j} are equal to R_i. Hence, $\lambda_{\min} \leq \lambda^* \leq \lambda_{\max}$. □

5.2.5 Algorithm for Low Rate Regime

At the low rate regime, i.e. all R_i's are sufficiently small, we can minimize the total transmission power by the following algorithm. To begin with, the total transmission power can be approximated as below in this regime.

Proposition 5.5. *If all R_i's are sufficiently small, Γ_i can be approximated as*

$$\Gamma_i \approx \sum_{i=1}^{M} \sum_{j=1,j\neq i}^{M} \left(\frac{1}{Z_{i,S_j}} - \frac{1}{Z_{i,D_i}} + \frac{1}{Z_{j,D_i}} \right) R_{i,S_j} + \sum_{i=1}^{M} \frac{1}{Z_{i,D_i}} R_i. \tag{5.23}$$

Proof. See Appendix C.3. □

In this case, the sum power minimization problem becomes

$$\min_{\mathbf{R}_i^{(S)},\mathbf{R}_i^{(D)}} \sum_{i=1}^{M} \sum_{j=1,j\neq i}^{M} \frac{1}{Z_{i,S_j}} R_{i,S_j} + \sum_{i=1}^{M} \sum_{j=1}^{M} \frac{1}{Z_{i,D_j}} R_{i,D_j} \tag{5.24}$$

subject to

$$R_{i,D_i} + \sum_{j=1,j\neq i}^{M} R_{i,S_j} \geq R_i, \quad i = 1, 2, \ldots, M \tag{5.25}$$

$$R_{i,S_j}, R_{i,D_i}, R_{i,D_j} \geq 0, \quad i = 1, 2, \ldots, M, j \neq i. \tag{5.26}$$

Firstly, it is noted that at the optimal point, equality holds for (5.25). Therefore, we can replace R_{i,D_i} by $R_i - \sum_{j=1,j\neq i}^{M} R_{i,S_j}$ and replace (5.25) by

$$\sum_{j=1,j\neq i}^{M} R_{i,S_j} \leq R_i. \tag{5.27}$$

In addition, the objective function becomes

$$\min \sum_{i=1}^{M} \sum_{j=1,j\neq i}^{M} \left(\frac{1}{Z_{i,S_j}} - \frac{1}{Z_{i,D_i}} + \frac{1}{Z_{j,D_i}} \right) R_{i,S_j} + \frac{1}{Z_{i,D_i}} R_i. \tag{5.28}$$

Now, the optimization problem can be divided into M independent subproblems. Each subproblem corresponds to the rate allocation of a transmitting node. The subproblem for S_i is

Problem 5.4.

$$\min_{R_{i,S_j}\geq 0} \sum_{j=1,j\neq i}^{M} \left(\frac{1}{Z_{i,S_j}} - \frac{1}{Z_{i,D_i}} + \frac{1}{Z_{j,D_i}}\right) R_{i,S_j} + \frac{1}{Z_{i,D_i}} R_i \tag{5.29}$$

subject to

$$\sum_{j=1,j\neq i}^{M} R_{i,S_j} \leq R_i. \tag{5.30}$$

If

$$\frac{1}{Z_{i,D_i}} < \frac{1}{Z_{i,S_j}} + \frac{1}{Z_{j,D_i}}, \quad j = 1, 2, \ldots, M, \quad j \neq i, \tag{5.31}$$

$R_{i,S_j} = 0$ for all $j \neq i$ which means $R_{i,D_i} = R$. Otherwise, $R_{i,S_j} = R$ if

$$j = \arg\min_{j\neq i} \left(\frac{1}{Z_{i,S_j}} - \frac{1}{Z_{i,D_i}} + \frac{1}{Z_{j,D_i}}\right). \tag{5.32}$$

Hence, at low rate regime, each transmitting node can independently perform the above rate allocation.

5.3 TDM-Based Scheme

Although the transmitting nodes can have a larger capacity region by using superposition codes over the outgoing links, it requires more complicated coding and modulation techniques [100, Discussion 6.1]. In the cooperative transmission scheme in this section, the nodes time-share their outgoing links. Apart from the rates of the outgoing links, we need to optimize the time allocated to each link.

5.3.1 Minimum Sum Transmit Power Problem

Consider the transmission of the outgoing links of S_i which time-share the same channel. The whole transmission time of S_i is divided into time frames of fixed durations. The duration of each time frame is normalized to be one unit which comprises of $2M-1$ time slots with variable durations. Each time slot corresponds to the transmission of one outgoing link. Let t_{i,S_j} be the duration for the time slot for transmission from S_i to S_j and t_{i,D_j} be the duration for the time slot for transmission from S_i to D_j. Therefore, t_{i,S_j} and t_{i,D_j} satisfy the following constraints.

$$\sum_{j=1,j\neq i}^{M} t_{i,S_j} + \sum_{j=1}^{M} t_{i,D_j} = 1, \quad i = 1,2,\ldots,M. \tag{5.33}$$

In each time frame, let x_{i,S_j} be the amount of data (in nats/Hz) transmitted from S_i to S_j and let x_{i,D_j} be the amount of data transmitted from S_i to D_j. The rate requirement becomes

$$x_{i,D_i} + \sum_{j=1,j\neq i}^{M} x_{i,S_j} \geq R_i, \quad i = 1,2,\ldots,M. \tag{5.34}$$

The transmitted SNR from S_i to S_j and the one from S_i to D_j are

$$\Gamma_{i,S_j} = \frac{1}{Z_{i,S_j}}\left[\exp\left(\frac{x_{i,S_j}}{t_{i,S_j}}\right) - 1\right] \quad \text{and} \tag{5.35}$$

$$\Gamma_{i,D_j} = \frac{1}{Z_{i,D_j}}\left[\exp\left(\frac{x_{i,D_j}}{t_{i,D_j}}\right) - 1\right] \tag{5.36}$$

respectively. Thus, the transmitted SNR of S_i is given by

$$\Gamma_i^{(T)} = \sum_{j=1,j\neq i}^{M} \frac{t_{i,S_j}}{Z_{i,S_j}}\left[\exp\left(\frac{x_{i,S_j}}{t_{i,S_j}}\right) - 1\right] + \sum_{j=1}^{M} \frac{t_{i,D_j}}{Z_{i,D_j}}\left[\exp\left(\frac{x_{i,D_j}}{t_{i,D_j}}\right) - 1\right]. \tag{5.37}$$

The minimum sum power problem is equivalent to the following optimization problem.

Problem 5.5.

$$\min_{x_{i,S_j}, x_{i,D_j}, t_{i,S_j}, t_{i,D_j} \geq 0} \sum_{i=1}^{M} \Gamma_i^{(T)} \tag{5.38}$$

subject to

$$\sum_{j=1,j\neq i}^{M} t_{i,S_j} + \sum_{j=1}^{M} t_{i,D_j} = 1, \quad i = 1,2,\ldots,M \tag{5.39}$$

$$x_{i,D_i} + \sum_{j=1,j\neq i}^{M} x_{i,S_j} \geq R_i, \quad i = 1,2,\ldots,M \tag{5.40}$$

$$x_{i,S_j} = x_{j,D_i}, \quad \forall i \neq j. \tag{5.41}$$

(5.41) is the flow conservation constraint for relaying messages from S_i to D_i through S_j. The objective function is a strictly convex function and the feasible set is convex. Therefore, it is also a convex optimization problem.

5.3.2 Pricing Interpretation for Distributed Implementation

Let $\mathbf{X}_i = (x_{i,D_i}, x_{i,S_1}, x_{i,S_2}, \ldots, x_{i,S_M})$, $\mathbf{T}_i = (t_{i,D_i}, t_{i,S_1}, t_{i,S_2}, \ldots, t_{i,S_M})$ and $\mathbf{Y}_i = (\mathbf{X}_i, \mathbf{T}_i)$. Also, we let $\mathbf{Y}_{-i} = (\mathbf{Y}_1, \mathbf{Y}_2, \ldots, \mathbf{Y}_{i-1}, \mathbf{Y}_{i+1}, \ldots, \mathbf{Y}_M)$. Similar to the superposition code based scheme, Problem 5.5 can be solved by playing a best-reply potential game with the potential function

$$\Phi^{(T)}(\mathbf{Y}_1, \mathbf{Y}_2, \ldots, \mathbf{Y}_M) = -\sum_{i=1}^{M} \Gamma_i^{(T)}. \tag{5.42}$$

In this case, Problem 5.5 can be solved by considering a best-reply potential game, in which $\mathbf{Y}_i$'s are optimized in a round-robin manner. By removing the terms independent of $\mathbf{Y}_i$ in (5.37) and substituting $x_{j,D_i} = x_{i,S_j}$ in (5.37), for fixed $\mathbf{Y}_{-i}$, $\mathbf{Y}_i$ can be optimized by solving the following problem.

Problem 5.6.

$$\begin{aligned}\max_{\mathbf{Y}_i} u_i^{(T)}(\mathbf{Y}_i, \mathbf{Y}_{-i}) = &-\frac{t_{i,D_i}}{Z_{i,D_i}}\left[\exp\left(\frac{x_{i,D_i}}{t_{i,D_i}}\right) - 1\right] - \sum_{j=1,j\neq i}^{M} \frac{t_{i,S_j}}{Z_{i,S_j}}\left[\exp\left(\frac{x_{i,S_j}}{t_{i,S_j}}\right) - 1\right] \\ &- \sum_{j=1,j\neq i}^{M} \frac{t_{j,D_i}}{Z_{j,D_i}}\left[\exp\left(\frac{x_{i,S_j}}{t_{j,D_i}}\right) - 1\right]\end{aligned} \tag{5.43}$$

subject to

$$\sum_{j=1,j\neq i}^{M} t_{i,S_j} + \sum_{j=1}^{M} t_{i,D_j} = 1 \tag{5.44}$$

$$x_{i,D_i} + \sum_{j=1,j\neq i}^{M} x_{i,S_j} \geq R_i \tag{5.45}$$

$$0 < \epsilon \leq t_{i,S_j}, t_{i,D_j} \leq 1, \quad j = 1, 2, \ldots, M \tag{5.46}$$

$$0 \leq x_{i,S_j}, x_{i,D_i} \leq R_i, \quad j = 1, 2, \ldots, M \tag{5.47}$$

where ϵ *in (5.46) is a sufficiently small positive constant.*

Problem 5.6 can be interpreted in a similar way as in Section 5.2.2. That is, the sum of the first two terms is the total transmitted SNR of S_i and the last term can be regarded as the total amount of virtual money for relaying the message. This is the motivation to formulate a similar game as in Section 5.2.2.

Consider the game $G_T = [\mathcal{P}_T, \{\mathcal{R}_i^{(T)}\}, \{u_i^{(T)}\}]$. The player set $\mathcal{P}_T$ is the set of source nodes $\{S_1, S_2, \ldots, S_M\}$. The strategy set of S_i, which is denoted by $\mathcal{R}_i^{(T)}$, is the feasible set of $\mathbf{Y}_i$, which is given by (5.44)-(5.47). The utility function of S_i is $u_i^{(T)}(\mathbf{Y}_i, \mathbf{Y}_{-i})$. The best response update of S_i is

$$B_i^{(T)}(\mathbf{Y}_{-i}) = \arg \max_{\mathbf{Y}_i \in \mathcal{R}_i^{(T)}} u_i^{(T)}(\mathbf{Y}_i, \mathbf{Y}_{-i}). \tag{5.48}$$

By following the derivations in Proposition 5.2, it can be shown that if a sequential best-reply path is admissible, the game converges to the Nash Equilibrium, which is exactly the optimal solution of Problem 5.5. But in this thesis, I only focus on the case that the game is played as outlined in Table 5.3.

Instead of estimating and computing the values of $K_{i,j}$, in this game, S_i only needs to know the values of t_{j,D_i} and Z_{j,D_i}. The values of Z_{j,D_i} can be informed by D_i whenever the value of Z_{j,D_i} changes. S_i can know the values of t_{j,D_i} from the headers of the MAC layer frames from S_j. In many network standards, the header of a MAC layer frame contains the identifier of the intended receiver, such as the MAC address of the network interface card. Therefore, S_i can measure t_{j,D_i} by counting the number of MAC layer frames from S_j to D_i

Table 5 3 Pricing Game for Power Minimization in TDM-Based Scheme

$i \leftarrow 1$
while Nash Equilibrium has not been reached **do**
  S_i maximizes $u_i^{(T)}(\mathbf{Y}_i, \mathbf{Y}_{-i})$ by optimizing $\mathbf{Y}_i$
  $i \leftarrow i + 1$
  $i \leftarrow 1$ **if** $i = M + 1$
end while

in one TDM time frame With this information, the source nodes can play the game with the flow specified in Table 5 3

Notice that (5 46) and (5 47) are not optional constraints They ensure the compactness of the strategy set $\mathcal{R}_i^{(T)}$, which is one of the sufficient conditions for the convergence of the best response updates in Proposition 5 2 It is noted that the introduction of ϵ slightly reduces the feasible set Suppose ϵ is replaced by 0 The objective function is no longer continuous at the point $t_{i,S_j} = x_{i,S_j} = 0$ for some j The best response update $B_i^{(T)}(\mathbf{Y}_{-i})$ is not a continuous function so it is invalid to imply that the best response update is a continuous function, which is one of the sufficient conditions in Proposition 5 2 But in practice, it is difficult to make a time slot arbitrarily small Even if it is possible to have such a small time slot, when x_{i,S_j} is positive, the transmission power is very high x_{i,S_j} must be reduced to 0 Thus, if t_{i,S_j} (resp t_{i,D_j}) is equal to ϵ, there will be no transmission over the link (i, S_j) (resp (i, D_j)) After introducing ϵ, the problem of the discontinuity of the best response update is alleviated and a similar result as in Proposition 5 2 can be proved

5.3.3 Best Response Update

Problem 5 6 is solved by optimizing $\mathbf{X}_i$ and $\mathbf{T}_i$ in round robin manner It can be proved in a similar way as Proposition 5 3 that this round robin optimization indeed solves Problem 5 6

Rate Allocation

For fixed $\mathbf{T}_i$, the rate allocation problem is

Problem 5.7.

$$\max_{0\le \mathbf{X}_i \le R_i} -\frac{t_{i,D_i}}{Z_{i,D_i}}\left[\exp\left(\frac{x_{i,D_i}}{t_{i,D_i}}\right)-1\right]-\sum_{j=1,j\neq i}^{M}\frac{t_{i,S_j}}{Z_{i,S_j}}\left[\exp\left(\frac{x_{i,S_j}}{t_{i,S_j}}\right)-1\right]-\sum_{j=1,j\neq i}^{M}\frac{t_{j,D_i}}{Z_{j,D_i}}\exp\left(\frac{x_{i,S_j}}{t_{j,D_i}}\right) \tag{5.49}$$

subject to

$$x_{i,D_i}+\sum_{j=1,j\neq i}^{M}x_{i,S_j}\ge R_i \tag{5.50}$$

The following partial Lagrangian is formed by relaxing (5.50).

$$\begin{aligned} L_i^{(X)}(\mathbf{X}_i,\lambda_i) &= \frac{t_{i,D_i}}{Z_{i,D_i}}\left[\exp\left(\frac{x_{i,D_i}}{t_{i,D_i}}\right)-1\right]+\sum_{j=1,j\neq i}^{M}\frac{t_{i,S_j}}{Z_{i,S_j}}\left[\exp\left(\frac{x_{i,S_j}}{t_{i,S_j}}\right)-1\right] \\ &+\sum_{j=1,j\neq i}^{M}\frac{t_{j,D_i}}{Z_{j,D_i}}\exp\left(\frac{x_{i,S_j}}{t_{j,D_i}}\right)-\lambda_i\left(x_{i,D_i}+\sum_{j=1,j\neq i}^{M}x_{i,S_j}-R_i\right). \end{aligned} \tag{5.51}$$

The corresponding partial dual function is

$$q_i^{(X)}(\lambda_i)=\min_{0\le x_{i,S_j},x_{i,D_i}\le R_i} L_i^{(X)}(\mathbf{X}_i,\lambda_i) \tag{5.52}$$

and the dual problem is

Problem 5.8. $\max_{\lambda_i\ge 0} q_i^{(X)}(\lambda_i)$.

To compute the partial dual function, the partial derivatives of $L_i^{(X)}$ with respect to each component of $\mathbf{X}_i$ are set to zero as follows.

$$\frac{\partial L_i^{(X)}}{\partial x_{i,D_i}}=\frac{1}{Z_{i,D_i}}\exp\left(\frac{x_{i,D_i}}{t_{i,D_i}}\right)-\lambda_i=0 \tag{5.53}$$

$$\frac{\partial L_i^{(X)}}{\partial x_{i,S_j}}=\frac{1}{Z_{i,S_j}}\exp\left(\frac{x_{i,S_j}}{t_{i,S_j}}\right)+\frac{1}{Z_{j,D_i}}\exp\left(\frac{x_{i,S_j}}{t_{j,D_i}}\right)-\lambda_i=0 \tag{5.54}$$

Clearly, the optimal x_{i,D_i} is $[t_{i,D_i}\ln(\lambda_i Z_{i,D_i})]_i^+$. $\frac{\partial L_i^{(X)}}{\partial x_{i,S_j}}$ is an increasing and continuous function of x_{i,S_j} so we can solve (5.54) by bisection method over the interval $[0,R_i]$. The

Table 5.4: Rate Allocation Algorithm of S_i

while λ_i has not converged **do**
  $\lambda_i \leftarrow (\lambda_i^{\min} + \lambda_i^{\max})/2$
  while $\mathbf{X}_i$ has not converged **do**
    $x_{i,D_i} \leftarrow [t_{i,D_i} \ln(\lambda_i Z_{i,D_i})]_i^+$
    for all $j \neq i$ and $1 \leq j \leq M$ **do**
      Solve for x_{i,S_j} in (5.54) by bisection search
    end for
  end while
  if $x_{i,D_i} + \sum_{j=1,j\neq i}^{M} x_{i,S_j} > R$ **then**
    $\lambda_i^{\max} \leftarrow \lambda_i$
  else
    $\lambda_i^{\min} \leftarrow \lambda_i$
  end if
end while

solution of (5.54) can be outside this interval. Thus, before the bisection search, $x_{i,S_j} = 0$ should be substituted in (5.54) to check if $\frac{\partial L_i^{(X)}}{\partial x_{i,S_j}}$ is non-negative. If it is non-negative, the optimal x_{i,S_j} must be 0. Similarly, it can be determined if the optimal x_{i,S_j} is equal to R_i.

The next question is how to solve the dual problem. Let $x^*_{i,D_i}(\lambda_i)$ and $x^*_{i,S_j}(\lambda_i)$ be the optimal x_{i,D_i} and x_{i,S_j} for a given λ_i respectively. Since $L_i^{(X)}$ is a strictly convex function of $\mathbf{X}_i$, x^*_{i,D_i} and x^*_{i,S_j} are well-defined functions of λ_i. By the Envelope Theorem, $q_i^{(X)}$ is differentiable and

$$\frac{dq_i^{(X)}}{d\lambda_i} = R_i - x^*_{i,D_i}(\lambda_i) - \sum_{j=1,j\neq i}^{M} x^*_{i,S_j}(\lambda_i). \tag{5.55}$$

By the concavity of $q_i^{(X)}(\lambda_i)$, it can be maximized by solving $\frac{dq_i^{(X)}}{d\lambda_i} = 0$ and the solution is denoted by λ_i^*. From (5.53) and (5.54), $\frac{dq_i^{(X)}}{d\lambda_i}$ is a decreasing and continuous function of λ_i.

If

$$\lambda_i = \lambda_i^{(\min)} = \min\left\{ \frac{1}{Z_{i,D_i}^{(D)}} \exp\left(\frac{R_i}{Mt_{i,D_i}}\right), \frac{1}{Z_{i,S_j}} \exp\left(\frac{R_i}{Mt_{i,S_j}}\right) + \frac{1}{Z_{j,D_i}} \exp\left(\frac{R_i}{Mt_{j,D_i}}\right) \right\}, \tag{5.56}$$

from (5.53) and (5.54), $x_{i,D_i}^*(\lambda_i)$ and $x_{i,S_j}^*(\lambda_i)$ are upper bounded by $\frac{R_i}{M}$ so $\frac{dq_i^{(X)}}{d\lambda_i}$ is non-negative and $\lambda_i^* \geq \lambda_i^{(\min)}$. On the other hand, if

$$\lambda_i = \lambda_i^{(\max)} = \min\left\{ \frac{1}{Z_{i,D_i}} \exp\left(\frac{R_i}{t_{i,D_i}}\right), \frac{1}{Z_{i,S_j}} \exp\left(\frac{R_i}{t_{i,S_j}}\right) + \frac{1}{Z_{j,D_i}} \exp\left(\frac{R_i}{t_{i,D_j}}\right) \right\}, \tag{5.57}$$

at least one of $x_{i,D_i}^*(\lambda_i)$ and $x_{i,S_j}^*(\lambda_i)$ is equal to R_i. Thus, $\frac{dq_i^{(X)}}{d\lambda_i}$ is non-positive and $\lambda_i^* \leq \lambda_i^{(\max)}$. Also, it implies that we can perform bisection search over the interval $\left[\lambda_i^{(\min)}, \lambda_i^{(\max)}\right]$. The rate allocation algorithm is summarized in Table 5.4. The outer loop is the bisection search for the optimal λ_i and the inner loop is the Gauss-Seidel update of $\mathbf{X}_i$.

Scheduling

For fixed $\mathbf{X}_i$, the scheduling problem is

Problem 5.9.

$$\max_{0<\epsilon\leq \mathbf{T}_i \leq 1} -\sum_{j=1}^{M} \frac{t_{i,D_j}}{Z_{i,D_j}} \left[\exp\left(\frac{x_{i,D_j}}{t_{i,D_j}}\right) - 1\right] - \sum_{j=1, j\neq i}^{M} \frac{t_{i,S_j}}{Z_{i,S_j}} \left[\exp\left(\frac{x_{i,S_j}}{t_{i,S_j}}\right) - 1\right] \tag{5.58}$$

subject to

$$\sum_{j=1, j\neq i}^{M} t_{i,S_j} + \sum_{j=1}^{M} t_{i,D_j} = 1 \tag{5.59}$$

To find the optimal schedule, the following partial Lagrangian is formed by relaxing (5.44).

$$\begin{aligned} L_i^{(T)}(\mathbf{T}_i, \mu_i) = &\sum_{j=1}^{M} \frac{t_{i,D_j}}{Z_{i,D_j}} \left[\exp\left(\frac{x_{i,D_j}}{t_{i,D_j}}\right) - 1\right] + \sum_{j=1, j\neq i}^{M} \frac{t_{i,S_j}}{Z_{i,S_j}} \left[\exp\left(\frac{x_{i,S_j}}{t_{i,S_j}}\right) - 1\right] \\ &- \mu_i \left(\sum_{j=1, j\neq i}^{M} t_{i,S_j} + \sum_{j=1}^{M} t_{i,D_j} - 1\right). \end{aligned} \tag{5.60}$$

The corresponding partial dual function is

$$q_i^{(T)}(\mu_i) = \min_{\epsilon \le t_{i,D_j}, t_{i,S_j} \le 1} L_i^{(T)}(\mathbf{T}_i, \mu_i) \tag{5.61}$$

and the corresponding dual problem is

Problem 5.10. $\max_{\mu_i} q_i^{(T)}(\mu_i)$.

For a given μ_i, $q_i^{(T)}(\mu_i)$ is computed by solving the following equations.

$$\frac{\partial L_i^{(T)}}{\partial t_{i,S_j}} = \frac{1}{Z_{i,S_j}}\left[\exp\left(\frac{x_{i,S_j}}{t_{i,S_j}}\right) - 1\right] - \frac{x_{i,S_j}}{Z_{i,S_j} t_{i,S_j}} \exp\left(\frac{x_{i,S_j}}{t_{i,S_j}}\right) - \mu_i = 0 \tag{5.62}$$

$$\frac{\partial L_i^{(T)}}{\partial t_{i,D_j}} = \frac{1}{Z_{i,D_j}}\left[\exp\left(\frac{x_{i,D_j}}{t_{i,D_j}}\right) - 1\right] - \frac{x_{i,D_j}}{Z_{i,D_j} t_{i,D_j}} \exp\left(\frac{x_{i,D_j}}{t_{i,D_j}}\right) - \mu_i = 0 \tag{5.63}$$

Both $\frac{\partial L_i^{(T)}}{\partial t_{i,S_j}}$ and $\frac{\partial L_i^{(T)}}{\partial t_{i,D_j}}$ are increasing and continuous functions of t_{i,S_j} and t_{i,D_j} respectively. Thus, the optimal t_{i,S_j}'s and t_{i,D_j}'s can be obtained by performing bisection search for the corresponding equations in the closed interval of $[\epsilon, 1]$.

For a given μ_i, let $t^*_{i,S_j}(\mu_i)$ and $t^*_{i,D_j}(\mu_i)$ be the optimal t_{i,S_j} and t_{i,D_j} respectively. Since $L_i^{(T)}$ is a strictly convex function of t_{i,S_j} and t_{i,D_j}, $t_{i,j}^{(S)*}(\mu_i)$ and $t^*_{i,D_j}(\mu_i)$ are well-defined functions of μ_i. By Envelope Theorem,

$$\frac{dq_i^{(T)}}{d\mu_i} = 1 - \sum_{j=1, j\neq i}^{M} t^*_{i,S_j}(\mu_i) - \sum_{j=1}^{M} t^*_{i,D_j}(\mu_i). \tag{5.64}$$

μ_i is optimized by solving $\frac{dq_i^{(T)}}{d\mu_i} = 0$. In (5.62) and (5.63), t^*_{i,S_j} and t^*_{i,D_j} are increasing and continuous functions of μ_i. If

$$\mu_i = \max\left\{\frac{1 - x_{i,S_j}}{Z_{i,S_j}} \exp\left(x_{i,S_j}\right) - \frac{1}{Z_{i,S_j}}, \frac{1 - x_{i,D_j}}{Z_{i,D_j}} \exp\left(x_{i,D_j}\right) - \frac{1}{Z_{i,D_j}}\right\}, \tag{5.65}$$

at least one of t^*_{i,S_j} and t^*_{i,D_j} is equal to 1 so $\frac{dq_i^{(T)}}{d\mu_i}$ is negative. On the other hand, if

$$\mu_i = \min\left\{\frac{1 - 2Mx_{i,S_j}}{Z_{i,S_j}} \exp\left(2Mx_{i,S_j}\right) - \frac{1}{Z_{i,S_j}}, \frac{1 - 2Mx_{i,D_j}}{Z_{i,D_j}} \exp\left(2Mx_{i,D_j}\right) - \frac{1}{Z_{i,D_j}}\right\}, \tag{5.66}$$

all t^*_{i,S_j} and all t^*_{i,D_j} are not greater than $\frac{1}{2M}$ so $\frac{dq_i^{(T)}}{d\mu_i}$ is positive. Thus, there exists a μ_i^* such that $g_i^{(T)}(\mu_i^*) = 0$ and it can be found by bisection method. The algorithm in Table 5.5 is

Table 5.5: Scheduling Algorithm of S_i

while μ_i has not converged **do**
  $\mu_i \leftarrow (\mu_i^{\min} + \mu_i^{\max})/2$
  while $\mathbf{T}_i$ has not converged **do**
    for $j = 1$ to M **do**
      if $i \neq j$ **then**
        Solve for t_{i,S_j} in (5.62) by bisection search
      end if
      Solve for t_{i,D_j} in (5.63) by bisection search
    end for
  end while
  if $\sum_{j=1,j\neq i}^{M} t_{i,S_j} + \sum_{j=1}^{M} t_{i,D_j} > 1$ **then**
    $\mu_i^{\max} \leftarrow \mu_i$
  else
    $\mu_i^{\min} \leftarrow \mu_i$
  end if
end while

the scheduling algorithm. The outer loop is the bisection search of μ_i and the inner loop is the optimization of $\mathbf{T}_i$ for a fixed μ_i.

5.4 Performance Evaluation

The performance of our cooperative transmission schemes is evaluated through simulations. Their abilities to combat against fading, their implementation costs and their ability to adapt to power gain changes are compared in these simulations.

In these simulations, it is assumed that the distances between the nodes are one unit. The power gains of the links are exponentially distributed with mean 1 which is the Rayleigh fading case. The channel is assumed to change slowly that throughout the games, the power gains remain unchanged. This is not an overly stringent assumption because both games will be shown to have high enough convergence rate in Section 5.4.2. For simplicity, R_i's are set to be the same. The algorithms for both cooperative transmission schemes are terminated when the total transmission power is improved by less than 10^{-4} units. In the TDM-based scheme, ϵ is chosen to be 10^{-6}.

5.4.1 Outage Performance

Outage probability is defined as the probability that for a given set of R_i's, the minimum required normalized SNR is greater than a given normalized SNR. The outage probability of direct transmission is plotted so that we can investigate the order of diversity, which will be detailed below, of our schemes. The common required rates are chosen to be 1, 2 and 4.

To compare the outage performance of the two cooperative transmission schemes with different number of nodes and different common rate requirements, the total transmitted SNR is normalized. Let R be the common required rate, i.e. $R_1 = R_2 = \ldots = R_M = R$. The normalized total SNR $\overline{\Gamma}_T$ is given by

$$\overline{\Gamma}_T = \frac{\sum_{i=1}^{M} \Gamma_i}{M\left(e^R - 1\right)}. \tag{5.67}$$

The denominator is the total transmitted SNR for no fading (i.e. a pure AWGN channel with power gain 1). Hence, $\overline{\Gamma}_T$ can be regarded as the additional total SNR, in dB, to support the same rate requirement R in the presence of fading.

With this normalized SNR, the *diversity order* is defined as follows. Let $\Phi_D(\overline{\Gamma}_T), \Phi_{BC}(\overline{\Gamma}_T)$ and $\Phi_{TDM}(\overline{\Gamma}_T)$ be the outage probability of direct transmission, superposition code based scheme and TDM based scheme when the normalized total transmitted SNR is $\overline{\Gamma}_T$. If there exists a real number r such that

$$\lim_{\overline{\Gamma}_T \to \infty} \frac{\Phi_{BC}(\overline{\Gamma}_T)}{\Phi_D(\overline{\Gamma}_T)} \overline{\Gamma}_T^k = r, \tag{5.68}$$

the diversity order of superposition code based scheme is said to be k. The diversity order of TDM-based scheme can be defined similarly. If the diversity order of a particular scheme is k, in the log-log plot of the outage probability, the slope of the curve of that scheme is k

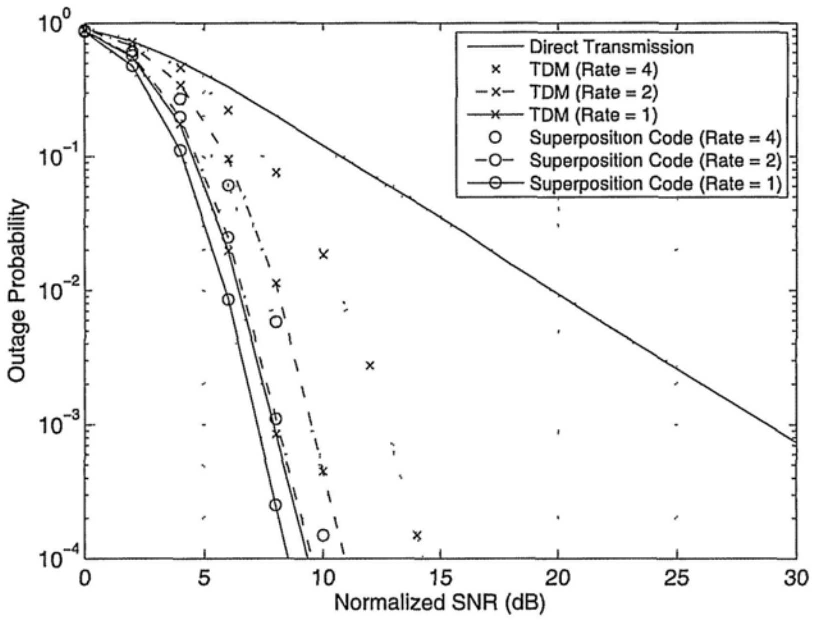

Figure 5.1: Comparison between direct transmission and the 4-user cooperative transmission schemes

times of the slope of the curve of direct transmission scheme for sufficiently high normalized total SNR.

The simulation results for $M = 4$ are shown in Figure 5.1. The slope of the curves for both cooperative transmission schemes is 4 times of the one for direct transmission asymptotically. That is, the diversity order of both schemes is 4. When R increases, the required normalized total SNR of the TDM-based scheme increases more rapidly than the one for the superposition code based scheme. It is because when the total transmission power of a node is increased, the difference between the broadcast channel capacity region and the TDM rate region increases. At high rate regime, the advantage of superposition coding is more apparent.

Figure 5.2 illustrates the simulation results for $M = 6$. The diversity order for both cooperative transmission schemes is about 6. Similar to the 4-user case, when the required rate increases, the difference of outage probabilities between the two cooperative transmission schemes increases.

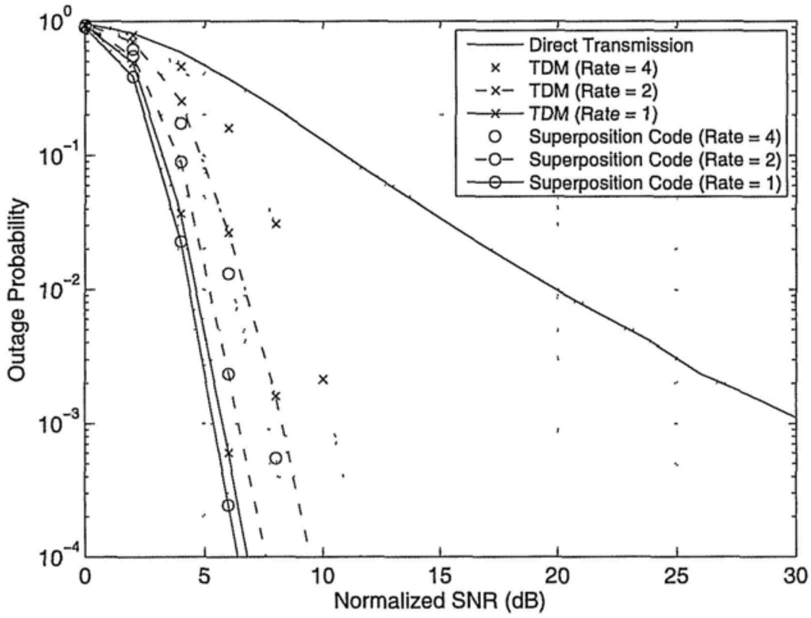

Figure 5.2: Comparison between direct transmission and the 6-user cooperative transmission schemes

5.4.2 Implementation Cost

In this part, the implementation costs of these two algorithms are compared. Here, implementation cost means the overhead communication cost of the algorithms which is the amount of information needed to be exchanged before reaching the Nash Equilibrium point. The communication cost can be inferred from the number of rounds in both games. In each round of both games, each node needs to send two real numbers to each other node. In each round of the superposition code based scheme, S_i needs to send the value of $\bar{K}_{i,j}$, which is the amount of virtual payment to S_j, and the value of $K_{i,j}$, which is the pricing information that S_j needs for its rate allocation, to S_j. Similarly, in each round of the TDM-based scheme, S_i needs to send the amount of payment and pricing information to S_j. Hence, the communication cost for both schemes are proportional to the number of rounds in both games by the same factor.

Two ways of initializing the algorithms are considered. In the first way, all nodes perform direct transmissions only. In the second way, the nodes allocate equal amount of rate and time to the outgoing links. Due to the page limitations, only the results of $M = 4$ are provided. Similar results can be obtained for other values of M.

The distributions of the number of best response updates required by each node, which can infer the communication costs of both schemes, are obtained. The cases that R is 2 and 4 are considered. The results are shown in Figures 5.3 to 5.6. If initially, the nodes perform direct transmissions only, the number of best response updates is less than 10. But with equal rate and time allocation initially, the TDM-based scheme requires much more best response updates. Such difference is more apparent for higher R. It is because when we use a relay path in the TDM-based scheme, the relay node has smaller portion of time for its own message. It has to increase the transmission power for its own message exponentially. If it has to relay the messages from more nodes, on average, each link has even smaller portion of time and the transmission power grows with the rate more rapidly. Hence, in the TDM-based scheme, typically, the nodes transmit their messages over the direct path together with a small number of relay paths. If initially, we allocate equal amount of rate and time to every

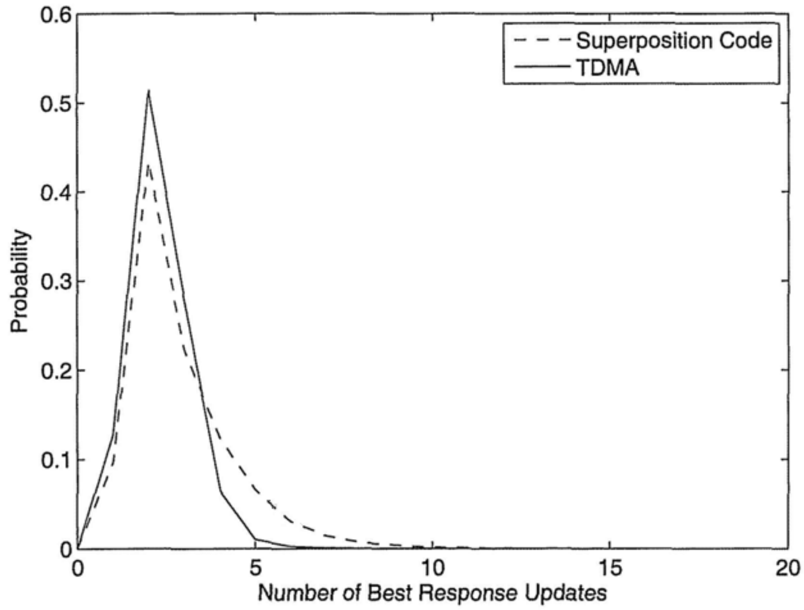

Figure 5.3: Distribution of number of best response updates of the 4-user cooperative transmission schemes (Rate = 2). The initial point for both schemes is direct transmission.

link, the nodes need more best response updates.

5.4.3 Adaptation to Power Gain Variations

Finally, the adaptability of our algorithms to the variations of power gains is investigated. The case with four transmit-receive pairs is considered. The common rate R is chosen to be 2. Similar results can be obtained when there are different number of users and common rate. The source nodes initially allocate equal amount of rates for all outgoing paths. After every 5 iterations of the algorithms, a new independent set of power gains are generated even if the Nash Equilibrium of the games have not been reached. Here, in one iteration of an algorithm, every source node has performed the best response update once.

The total transmission SNR is plotted in Figure 5.7. The horizontal dotted lines denote the optimal total transmission SNR of the corresponding cooperative transmission scheme

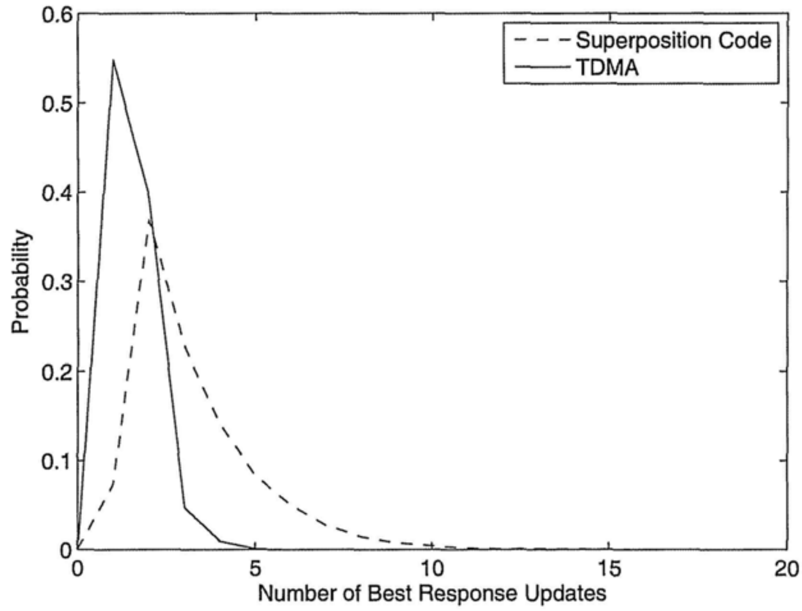

Figure 5.4: Distribution of number of best response updates of the 4-user cooperative transmission schemes (Rate = 4). The initial point for both schemes is direct transmission.

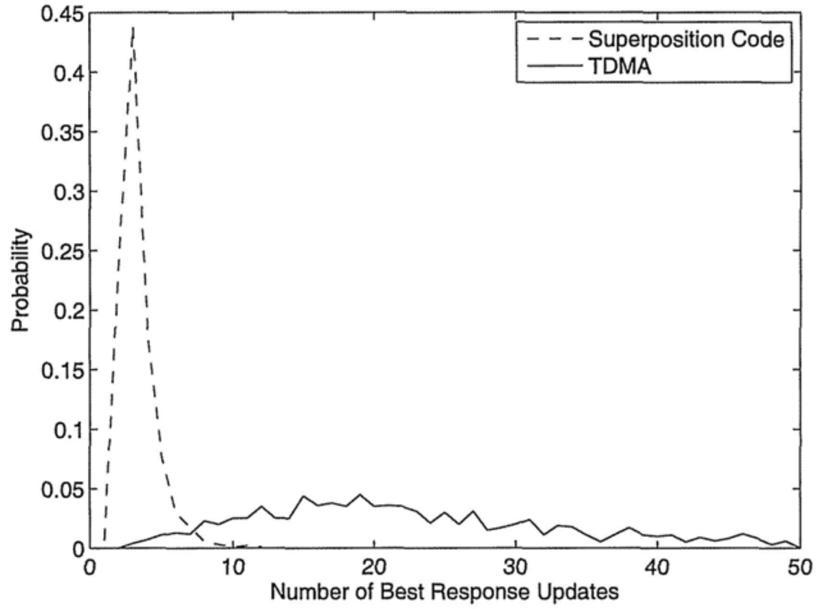

Figure 5.5: Distribution of number of best response updates of the 4-user cooperative transmission schemes (Rate = 2). Initially, all links are allocated equal amount of rate and time.

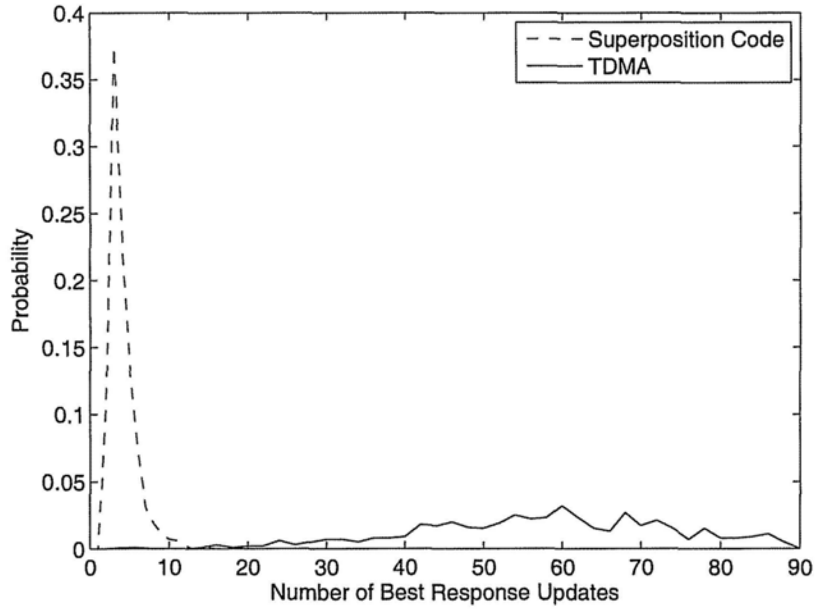

Figure 5.6: Distribution of number of best response updates of the 4-user cooperative transmission schemes (Rate = 4). Initially, all links are allocated equal amount of rate and time.

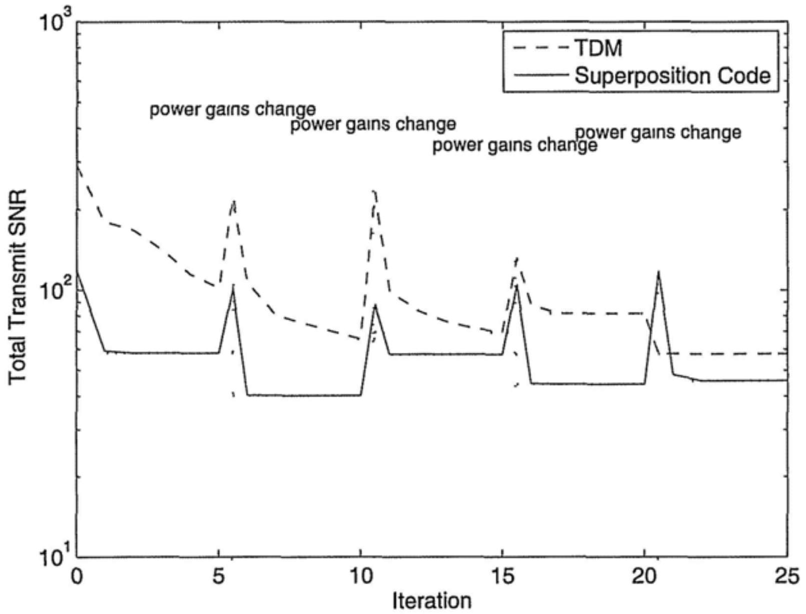

Figure 5.7: Total transmission SNR of the two 4-user cooperative transmission schemes with respect to the power gain changes.

under the corresponding set of power gains. In this figure, the total transmission SNR of both cooperative transmission schemes have been improved significantly after the first few iterations of their algorithms. The difference between the total transmission SNR and the optimal point has been reduced by more than half. This is very crucial in the adaptation to channel variations. Even if the Nash Equilibrium has not been reached before some instances of power gain changes, the source nodes can reduce the total transmission power significantly within a short time period.

5.5 Conclusion

In this chapter, I propose two pricing games for two cooperative transmission schemes. In the first game, the nodes multiplex the messages of their outgoing links by superposition codes while in the second game, they time-share their outgoing links. At Nash Equilibria of both games, the total transmission power is minimized. Based on both games, I propose the adaptive and distributed implementations. Simulation results show the significant improvement in power consumption by our schemes and their high convergence rates. When the demanded rate is increased, the advantage of the superposition code based scheme over the TDM-based scheme in terms of outage performance becomes more apparent.

Chapter 6

Partner Selection in Cooperative Transmission

In Chapter 5, the TDM-based cooperative transmission scheme is proposed to alleviate the coding complexity in the superposition code based scheme. In a broadcast channel, if the number of users is small (for example, fewer than 4 users), the coding complexity of the superposition code is still acceptable in practical systems. Therefore, an alternative method to alleviate the coding complexity in the superposition code based scheme is to perform partner selection among a set of point-to-point links.

The main difference from the previous two schemes is that the source node only relays the message for its assigned partnered source node. Hence, one advantage of partner selection scheme over the previous two schemes is that instead of listening to all the spectrum consumed by the source nodes, each source node only needs to listen to the channel of the partnered source node. Each source node multiplexes the messages of its three outgoing links, namely, the links to its destination node, the partnered source node and the partnered destination node, through broadcast channel coding. The partnered node decodes the incoming message, re-encodes and forwards it to the destination node.

In this chapter, I consider the partner selection problem for cooperative transmission among a set of point-to-point links. The objective is to minimize the sum power. By obtaining the minimum sum power, we can compare the outage performance of the partner selection scheme with the previous two cooperative transmission scheme.

The partner selection problem is solved in two steps. Firstly, the minimum sum power for each pair of nodes if they cooperate is computed. In some cases, it has closed-form optimal solutions. For other cases, we propose a simple iterative algorithm which converges rapidly. With these results, we can perform the partner selection. For a large number of nodes, in order to reduce the overhead of obtaining the first step results, I propose the *grouping algorithm* which is shown to be near-optimal and has fast convergence by simulations.

This chapter is organized as follows. The system model is provided in Section 6.1. Based on this model, we propose our partner selection algorithm in Section 6.2. Then, we evaluate the performance of the algorithm in Section 6.3.

Part of the contents can be found in [72].

6.1 System Model and Problem Formulation

Consider a wireless network with N distinct source-destination pairs. The source and destination node of the i-th node pair are denoted by S_i and D_i respectively. This node pair has a minimum required rate of R_i.

Each source node is assigned an orthogonal channel. For notational simplicity, it is assumed that each channel has an equal bandwidth of W. The results can be easily generalized to the unequal bandwidth case. The noise of each channel is independent and Gaussian distributed with N_0W as the variance. The power gains of the links from S_i to S_j and D_j are denoted by $Z_{i,j}^{(S)}$ and $Z_{i,j}^{(D)}$ respectively. The slow fading case is considered so the power gains are assumed to have negligible changes throughout the whole transmission.

Each node pair can be partnered with another one so that these two source nodes can overhear and relay the message for one other. The two partnered source nodes adopt the two-user cooperative transmission scheme described in Chapter 4. Each source node divides its data stream into two sub-streams. One sub-stream is directly transmitted to its destination node. Another one is relayed by the partnered source node. Thus, in its allocated channel, each source node has three outgoing sub-streams, namely, its two sub-streams and the one originated from the partnered source node. Hence, it is a three-user Gaussian broadcast channel which has the coding and decoding schemes described in Section 2.1.2.

The objective is to minimize the total transmission power. Let $P_{i,j}$, $i \neq j$, be the minimum sum power of S_i and S_j if they cooperate and $P_{i,i}$ be the power for direct transmission from S_i to D_i. The sum power of S_i and S_j is minimized by a rate allocation of their sub-streams. It will be detailed in Section 6.2.1. The partner selection problem is formulated below.

Problem 6.1.

$$\min_{x_{i,j}\in\{0,1\}} \sum_{i=1}^{N}\sum_{j\geq i}^{N} P_{i,j}x_{i,j} \tag{6.1}$$

subject to

$$\sum_{j=1}^{N} x_{i,j} = 1, \quad i = 1, 2, \quad , N \tag{6.2}$$

$$x_{i,j} = x_{j,i}, \quad i < j. \tag{6.3}$$

$x_{i,j}$ is the decision variable of partnering S_i and S_j. If S_i partners with S_j, $x_{i,j} = 1$. Otherwise, $x_{i,j} = 0$. $x_{i,i} = 1$ means S_i has no partners.

6.2 Proposed Solution

Problem 6.1 is solved with two steps. Firstly, the values of $P_{i,j}$'s are computed. With these values, the partner selection can be performed.

6.2.1 Computation of $P_{i,j}$

For the convenience in presentation, the computation of $P_{1,2}$ is considered. Let $R_i^{(R)}$ and $R_i^{(D)}$ be the rates of the sub-streams on the relay and direct path from S_i respectively. The rates of all outgoing links from S_i are represented compactly as $(R_i^{(D)}, R_i^{(R)}, R_{3-i}^{(R)})$. The capacity region of these links is denoted by $\mathcal{C}_i(P_i)$ where P_i is the transmission power of S_i. $P_{1,2}$ is computed by solving the following optimization problem.

Problem 6.2.

$$\min_{R_i^{(R)}, R_i^{(D)} \geq 0} \Gamma_1 + \Gamma_2 \tag{6.4}$$

subject to

$$(R_i^{(D)}, R_i^{(R)}, R_{3-i}^{(R)}) \in \mathcal{C}_i(P_i), \quad i = 1, 2 \tag{6.5}$$

$$R_i^{(D)} + R_i^{(R)} \geq R_i, \quad i = 1, 2, \tag{6.6}$$

where Γ_i is the transmit signal-to-noise ratio (SNR) which is given by

$$\Gamma_i = \frac{P_i}{N_0 W}. \tag{6.7}$$

$P_{1,2}$ is equal to the minimum $N_0 W(\Gamma_1 + \Gamma_2)$.

Constraint (6.5), which is a set statement, can be replaced by the following close-form expression which is a direct consequence of Lemma 5.1. Let

$$\delta_{i,j}^{(S)} = \begin{cases} 1 & \text{, if } Z_{i,3-i}^{(S)} > Z_{i,j}^{(D)} \\ 0 & \text{, otherwise.} \end{cases} \tag{6.8}$$

$$\delta_{i,3-i}^{(D)} = \begin{cases} 1 & \text{, if } Z_{i,i}^{(D)} > Z_{i,3-i}^{(D)} \\ 0 & \text{, otherwise.} \end{cases} \tag{6.9}$$

Given $R_i^{(R)}$ and $R_i^{(D)}$, the minimum required Γ_i is

$$\begin{aligned} \Gamma_i = & \frac{1}{Z_{i,i}^{(D)}} \exp\left[\left(1 - \delta_{i,i}^{(S)}\right) R_i^{(R)} + \delta_{i,3-i}^{(D)} R_{3-i}^{(R)}\right] \left(e^{R_i^{(D)}} - 1\right) \\ & + \frac{1}{Z_{i,3-i}^{(S)}} \exp\left(\delta_{i,i}^{(S)} R_i^{(D)} + \delta_{i,3-i}^{(S)} R_{3-i}^{(R)}\right) \left(e^{R_i^{(R)}} - 1\right) \\ & + \frac{1}{Z_{i,3-i}^{(D)}} \exp\left[\left(1 - \delta_{i,3-i}^{(S)}\right) R_i^{(R)} + \left(1 - \delta_{i,3-i}^{(D)}\right) R_i^{(D)}\right] \\ & \left(e^{R_{3-i}^{(R)}} - 1\right). \end{aligned} \tag{6.10}$$

Let $R_i^{(R)*}$ be the optimal $R_i^{(R)}$. It is noted that at the optimal point, equality holds for (6.6). Hence, we only need the characterization of the optimal rate allocation in terms of $R_i^{(R)*}$. Firstly, the close-form solutions for the following five cases, which are proved in Appendix D, are provided.

Case I: $Z_{1,1}^{(D)} > Z_{1,2}^{(S)}$, $Z_{2,2}^{(D)} > Z_{2,1}^{(S)}$

$R_1^{(R)*} = R_2^{(R)*} = 0.$

Case II: $Z_{\imath,3-\imath}^{(D)} > Z_{\imath,\imath}^{(D)} > Z_{\imath,3-\imath}^{(S)}$, $Z_{3-\imath,3-\imath}^{(D)} < Z_{3-\imath,\imath}^{(S)}$

$$R_\imath^{(R)*} = 0. \tag{6.11}$$

$$R_{3-\imath}^{(R)*} = \mathcal{P}_{3-\imath}\left[\frac{R_{3-\imath} - R_\imath}{2} + \frac{1}{2}\ln\left(\frac{1}{Z_{3-\imath,3-\imath}^{(D)}} - \frac{1}{Z_{3-\imath,\imath}^{(S)}}\right) - \frac{1}{2}\ln\left(\frac{1}{Z_{\imath,3-\imath}^{(D)}}\right)\right]. \tag{6.12}$$

and $\mathcal{P}_\imath[x]$ is the projection of x into the interval $[0, R_\imath]$.

Case III: $Z_{\imath,\imath}^{(D)} > \max\left\{Z_{\imath,3-\imath}^{(S)}, Z_{\imath,3-\imath}^{(D)}\right\}$, $Z_{3-\imath,3-\imath}^{(D)} < Z_{3-\imath,\imath}^{(S)}$

$$R_\imath^{(R)*} = 0. \tag{6.13}$$

$$R_{3-\imath}^{(R)*} = \mathcal{P}_{3-\imath}\left[\frac{R_{3-\imath}}{2} + \frac{1}{2}\ln\left(\frac{1}{Z_{3-\imath,3-\imath}^{(D)}} - \frac{1}{Z_{3-\imath,\imath}^{(S)}}\right) - \frac{1}{2}\ln\left(\frac{e^{R_\imath} - 1}{Z_{\imath,\imath}^{(D)}} + \frac{1}{Z_{\imath,3-\imath}^{(D)}}\right)\right]. \tag{6.14}$$

Case IV: $Z_{1,2}^{(S)} > Z_{1,1}^{(D)} > Z_{1,2}^{(D)}$, $Z_{2,1}^{(S)} > Z_{2,2}^{(D)} > Z_{2,1}^{(D)}$

Let

$$R_\imath^{(R)**} = \mathcal{P}_\imath\left[\frac{R_\imath}{2} + \frac{1}{2}\ln\left(\frac{1}{Z_{3-\imath,3-\imath}^{(D)}} - \frac{1}{Z_{3-\imath,\imath}^{(S)}}\right) - \frac{1}{2}\ln\left(\frac{e^{R_{3-\imath}} - 1}{Z_{\imath,\imath}^{(D)}} + \frac{1}{Z_{\imath,3-\imath}^{(D)}}\right)\right]. \tag{6.15}$$

and let $\Gamma_T^*(r_1^{(R)}, r_1^{(D)}, r_2^{(R)}, r_2^{(D)})$ be the value of Γ_T if $(R_1^{(R)}, R_1^{(D)}, R_2^{(R)}, R_2^{(D)}) = (r_1^{(R)}, r_1^{(D)}, r_2^{(R)}, r_2^{(D)})$. If $\Gamma_T^*(R_1^{(R)**}, R_1 - R_1^{(R)**}, 0, R_2) < \Gamma_T^*(0, R_1, R_2^{(R)**}, R_2 - R_2^{(R)**})$, $R_1^{(R)*} = R_1^{(R)**}$ and $R_2^{(R)*} = 0$. Otherwise, $R_1^{(R)*} = 0$ and $R_2^{(R)*} = R_2^{(R)**}$.

Case V: $Z_{1,2}^{(D)} > Z_{1,2}^{(S)} > Z_{1,1}^{(D)}$, $Z_{2,1}^{(D)} > Z_{2,1}^{(S)} > Z_{2,2}^{(D)}$

$$R_\imath^{(R)*} = \mathcal{P}_\imath\left[\frac{R_\imath - R_{3-\imath}}{2} + \frac{1}{2}\ln\left(\frac{1}{Z_{\imath,\imath}^{(D)}} - \frac{1}{Z_{\imath,3-\imath}^{(S)}}\right) - \frac{1}{2}\ln\left(\frac{1}{Z_{3-\imath,\imath}^{(D)}}\right)\right]. \tag{6.16}$$

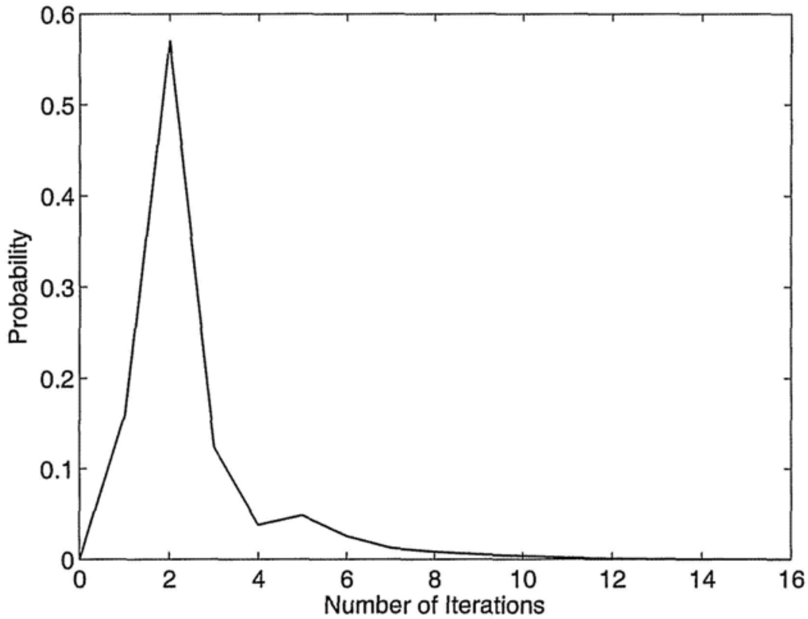

Figure 6.1: Distribution for number of iterations required for the round robin optimization ($R_1 = R_2 = 2$).

The remaining cases can be solved by a round robin optimization of $R_i^{(R)}$ and $R_{3-i}^{(R)}$. The following proposition justifies that this approach converges to the optimal solution.

Proposition 6.1. *The optimal $R_i^{(R)}$ and $R_{3-i}^{(R)}$ can be obtained by the above algorithm.*

Proof. This round robin algorithm is a Gauss-Seidel type algorithm [10]. To ensure the convergence to the optimal solution, it is required that [10, p. 219] the objective function is convex and differentiable and the feasible set is compact. By (6.10), the objective function is strictly convex and differentiable. The feasible set is $[0, R_1] \times [0, R_2]$ which is obviously compact. Therefore, for any feasible initial point, the algorithm converges to the optimal solution. □

Since Γ_i's are convex functions of $R_i^{(R)}$, the optimal $R_i^{(R)}$ (for fixed $R_{3-i}^{(R)}$) is obtained by solving $\frac{\partial \Gamma_1 + \Gamma_2}{\partial R_i^{(R)}} = 0$ and then projecting the solution to the interval $[0, R_i]$. Due to the strict convexity of Γ_i's, $\frac{\partial \Gamma_1 + \Gamma_2}{\partial R_i^{(R)}}$ is an increasing function of $R_i^{(R)}$, the solution can be efficiently solved by bisection method.

For $R_1 = R_2 = 2$, the distribution of the required number of iterations of this approach, which is obtained by 100000 simulation trials, is plotted in Fig. 6.1. All power gains are independent exponential random variables with mean 1. $R_1^{(R)}$ and $R_2^{(R)}$ are 0 initially. In most cases, we need at most 6 iterations to converge. For other values of R_1, R_2, similar distributions can be observed. Thus, $P_{i,j}$ can be efficiently computed.

For sufficiently small R_i's, the optimal rate allocation is as below.

Proposition 6.2. *For sufficiently small R_1 and R_2,*

$$R_i^{(R)*} = \begin{cases} 0, & \text{if } \frac{1}{Z_{i,i}^{(D)}} < \frac{1}{Z_{i,3-i}^{(S)}} + \frac{1}{Z_{i,3-i}^{(D)}}, \\ R_i, & \text{otherwise.} \end{cases} \tag{6.17}$$

Proof. Please refer to the derivations in Section 5.2.5. □

Here, the messages are transmitted either through the direct or relay path.

Table 6.1: Grouping Algorithm

All nodes partner with themselves and form groups of $M \leq N$ nodes randomly.
while total transmission power has not converged **do**
  Apply Gabow's algorithm in each group independently.
  Select two nodes in each group which either are partnered or both have no partners Choose those which have the maximum sum power
  The nodes selected in group i are moved to group $i + 1 \bmod \lceil \frac{N}{M} \rceil$
end while

6.2.2 Partner Selection

After computing the $P_{i,j}$'s, Problem 6.1 is solved by considering the following weighted undirected graph, $G = (V, E)$. A vertex $v \in V = \{1, 2, \ldots, N\}$ corresponds to node pair v. The weight of an edge (i, j), $i \neq j$, is $(2\bar{P} - P_{i,j})$ and the weight of (i, i) is $(\bar{P} - P_{i,i})$ where $\bar{P}$ is chosen so that the edge weights are positive.

A subset of edges are selected such that the sum of their weights is maximized and no two selected edges share a common vertex. Hence, for every vertex in V, exactly one incoming edge is selected. Thus, the total weight of the selected edges is $(N\bar{P} - \sum_{(i,j)\in\mathcal{S}} P_{i,j})$ where $\mathcal{S}$ is the set of selected edges. Maximizing this quantity is equivalent to minimizing the sum of $P_{i,j}$'s, which is exactly Problem 6.1.

This graph problem is known as the *maximum weighted matching problem* and can be solved by Gabow's algorithm (Due to the page limit, please refer to [30] for details.) with time complexity $O(N^3)$. However, it is still impractical for networks with large N due to the communication overhead of exchanging the channel state information (CSI) of $(2N - 1)$ links and the computational overhead of the $\frac{N(N+1)}{2}$ values of $P_{i,j}$'s. To solve this problem, I propose an iterative algorithm called *grouping algorithm* in Table 6.1

Now, the nodes exchange the CSI only when they join a new group or new members join their group. Therefore, the total amount of CSI exchange is reduced. In each iteration, the sum power of each group is solved independently so the solution remains feasible but the sum power is non-increasing. Since it is lower bounded by the optimal solution, the algorithm converges.

6.3 Performance Evaluation

The performance of the grouping algorithm is evaluated through simulations. More precisely, its outage performance, which infers the probability distribution of the required sum power, and its convergence speed are investigated.

The distance between every pair of nodes is assumed to be equal. The power gains of the links are independent exponential random variables with mean 1, which is the Rayleigh fading case. For simplicity, all R_i's are equal to a common required rate R. The outage probability is defined in the same way as in Chapter 5. Firstly, the normalized total SNR is defined as below.

Definition 6.1. *If Γ_T is the total SNR of a particular implementation, the corresponding* normalized total SNR, *$\tilde{\Gamma}_T$, is*

$$\tilde{\Gamma}_T = \frac{\Gamma_T}{N\left(e^R - 1\right)}. \tag{6.18}$$

The denominator in (6.18) is the required SNR in the absence of fading (i.e. the required SNR in a Gaussian channel for the common required rate R). (6.18) can be interpreted as the additional amount of dB of SNR to combat against fading.

Definition 6.2. *Let $\tilde{\Gamma}_T^*$ be the minimum required normalized total SNR for an implementation. Given a normalized total SNR of a network $\tilde{\Gamma}_T$, the* outage probability *is defined as $Pr\{\tilde{\Gamma}_T^* > \tilde{\Gamma}_T\}$.*

The outage performance between the grouping algorithm, the direct transmission scheme and the optimal partner selection are compared. The cases that R is equal to 2 and 4, N is equal to 12 and 18 and M is equal to 2 and 3 are considered. The results are plotted in Fig. 6.2-6.5. In these figures, the grouping algorithm and the optimal partner selection outperforms the direct transmission scheme. The curves for the grouping algorithm deviate

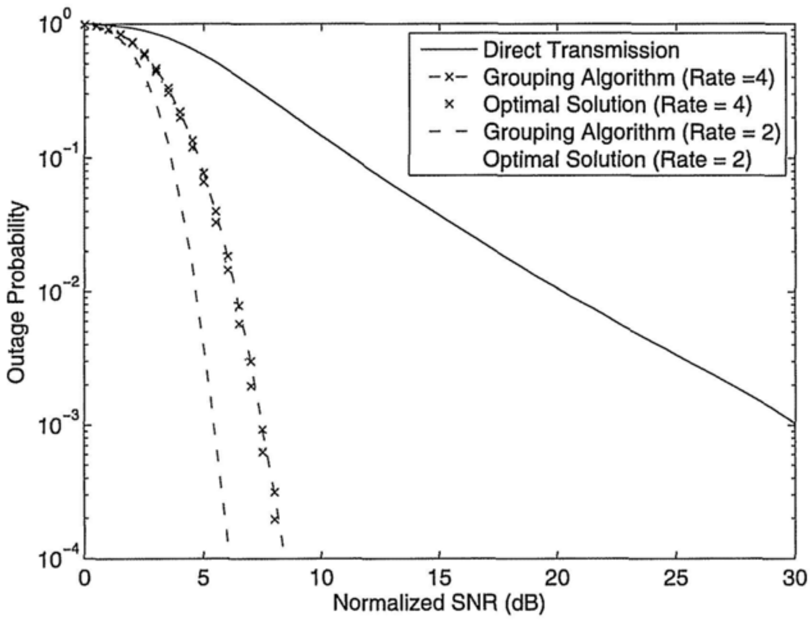

Figure 6.2: Outage performance for $N = 12$ and $M = 2$.

by at most 1 dB from the optimal partner selection. Hence, the reduction of communication overhead by grouping algorithm does not cost too much transmission power.

The diversity order, which is another commonly considered performance measure for cooperative transmission, is considered. Let $P_C'(\gamma)$ and $P_D'(\gamma)$ be the slopes of the curves of our cooperative transmission scheme and direct transmission scheme at the normalized SNR γ respectively. The diversity order Δ is defined as

$$\Delta = \lim_{\gamma \to \infty} \frac{P_C'(\gamma)}{P_D'(\gamma)}. \tag{6.19}$$

In all cases, the figures show that both the diversity order achieved by the optimal partner selection scheme and the grouping algorithm are approximately equal to the number of nodes. Therefore, both schemes achieve full order of diversity.

Finally, the convergence rate of the grouping algorithm is investigated through 100000 Monte Carlo simulation trials. The case of $N = 18$, M is equal to 2 and $R = 2$ is considered. The distributions of the number of iterations of the grouping algorithm is plotted in Fig. 6.6.

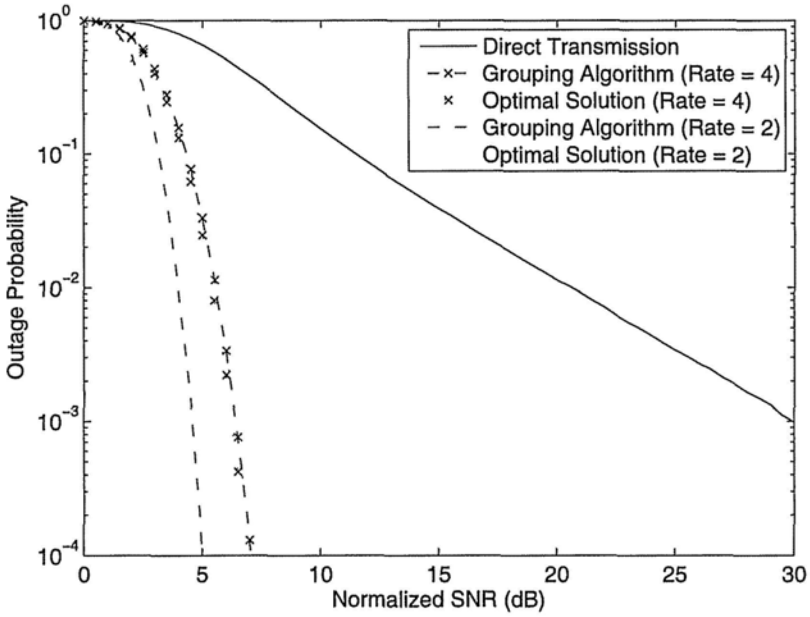

Figure 6.3: Outage performance for $N = 18$ and $M = 2$.

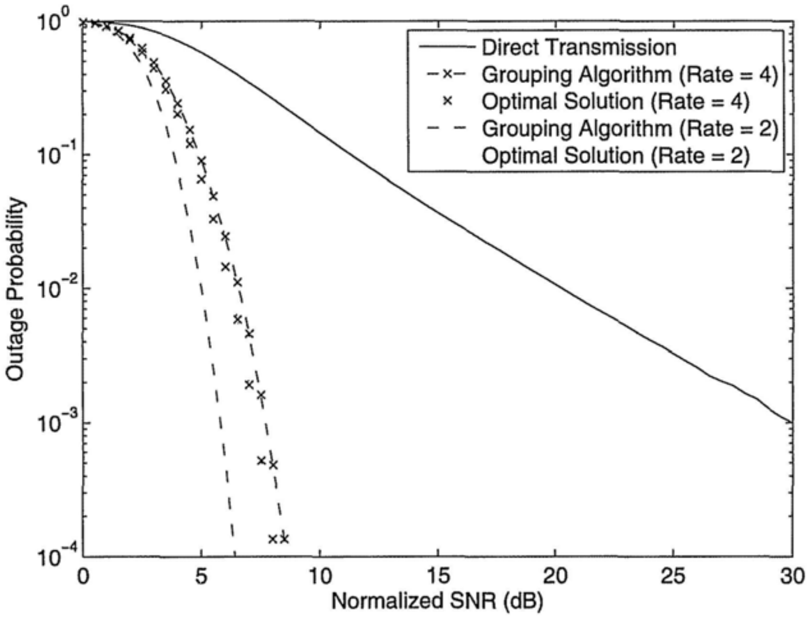

Figure 6.4: Outage performance for $N = 12$ and $M = 3$.

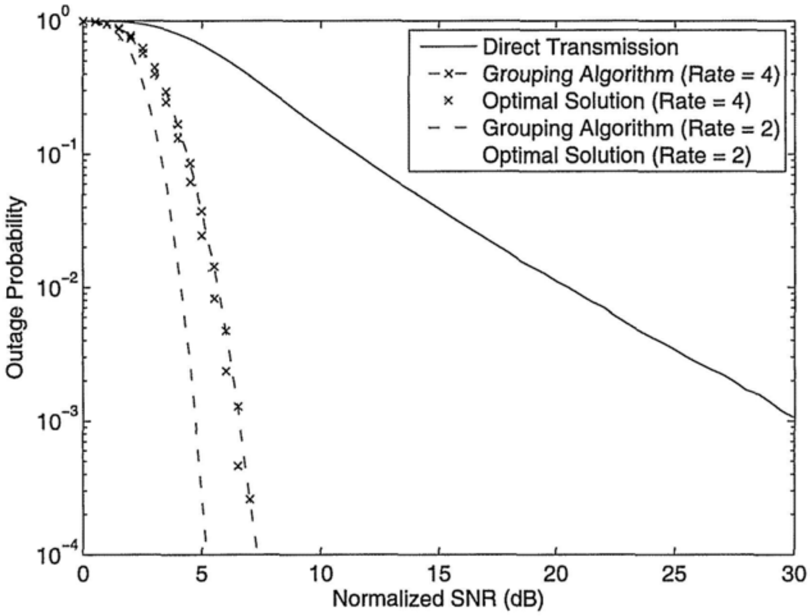

Figure 6.5: Outage performance for $N = 18$ and $M = 3$.

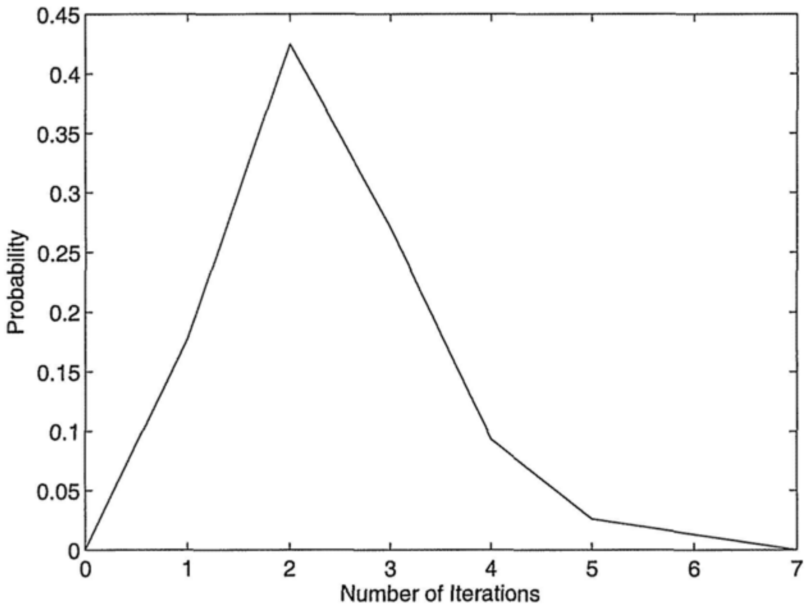

Figure 6.6: Distribution of number of iterations for $N = 18$ and $M = 2$.

In most cases, it requires fewer than 7 iterations. Similar observations can be found for other values of M, N and R. Thus, the grouping algorithm converges rapidly.

6.4 Conclusion

The partner selection problem for cooperative transmission is considered in this chapter. It provides an alternative way to alleviate the broadcast channel coding complexity for large number of users. One advantage of this approach over the schemes proposed in the previous chapters is that the source nodes only need to listen to the channel assigned to its partnered source node instead of all the channels of other nodes.

The sum power minimization problem is studied in this chapter. Firstly, a simple optimal rate allocation scheme for two cooperating node pairs is proposed. Closed-form results are obtained for some cases. For the remaining cases, a simple iterative algorithm, which has high convergence rate, is proposed. With the optimal rate allocation between each pair of nodes, we can select the partners for each node pair. Although Gabow's algorithm can provide the optimal partner selection, the communication overhead among the nodes is high for large number of nodes. Therefore, the grouping algorithm is proposed to reduce the communication overhead. Simulation results show that the grouping algorithm is near-optimal and has high convergence rate. Also, both the optimal partner selection scheme and the grouping algorithm can achieve full order of diversity.

Chapter 7

Conclusion and Future Works

In this chapter, I summarize the contributions of this thesis in Section 7.1. Then, I propose some future research topics. Motivations and comparisons with previous research works are discussed in those sections.

7.1 Summary of Contributions

In this thesis, I propose some cooperative transmission codes and protocols for distributed resource allocation. In Chapter 4, I propose two cooperative transmission schemes for the case of two source-destination pairs. Both schemes have larger achievable rate regions than direct transmission schemes. Although one of these two schemes has a slightly larger achievable rate region, the other one is considered in the remaining parts of the thesis because it allows simpler code implementation and resource allocations. Simulation results show that the latter code can achieve twice the diversity order of the interference channel code proposed by [26]. The achievable rate region is simulated by the proposed weighted sum rate maximization algorithm. Simplified implementations at low SNR regime are proposed as well. I also illustrate how this algorithm can be extended to solve the max-min fairness problem and the joint utility maximization problem.

In Chapter 5, this cooperative transmission code is extended for more than two source-destination pairs. A pricing game is proposed to have a distributed implementation for sum power minimization. Simulation results show that this cooperative transmission scheme can achieve full diversity order. It is noted that the building blocks of this cooperative transmission code are superposition codes which are too complex to implement for larger number of source nodes. Therefore, I propose another TDM-based cooperative transmission scheme to alleviate this complexity. Another pricing game is proposed for distributed sum power minimization. Simulation results show that this scheme also achieves full diversity order. In addition, simulation results also show that the sum power minimization protocols for the two cooperative transmission schemes can adapt to channel fluctuations.

Another way to alleviate the coding complexities of cooperative transmission code is to perform partner selection, which is discussed in Chapter 6. One advantage of this approach over the other two schemes is that the source nodes only need to listen to the channel assigned to their partnered nodes instead of all the channels in the network. For each pair of partnered source node, I propose a simple rate allocation algorithm. In some cases, there are closed-form solutions for the optimal rate allocation. For the remaining cases, a simple iterative algorithm, which has high convergence rate, is proposed to provide the optimal rate allocation. Then, I map the partner selection problem to the maximum weighted matching problem in graph theory which can be solved with polynomial-time algorithms. For large number of source nodes, I propose the grouping algorithm to reduce the computational and communication overhead. Simulation results show that the grouping algorithm deviates from the optimal solution by at most 1 dB. Both partner selection schemes can also achieve the full diversity order.

7.2 Cooperative Transmission for MIMO Systems

As mentioned in Chapter 1, one application of cooperative transmission is to further increase the diversity order of MIMO systems. Here are some questions to be answered:

1. How should the nodes, which are equipped with multiple antennas, cooperate with one another?

2. How does the diversity order scale with the number of relay paths and the number of antennas on each node?

One straightforward scheme is to extend the proposed schemes in this thesis is by replacing the scalar broadcast channel codes by the MIMO broadcast channel codes. The capacity region of MIMO broadcast channel is studied in [105]. However, the authors have shown by an example that in general, in order to achieve a Pareto-optimal rate vector in a MIMO broadcast channel, we need to perform time-sharing among a set of dirty-paper codes. This has two implications to the implementation.

The most obvious implication is the complexity of the implementation of the cooperative transmission code. Due to the time-sharing of dirty-paper codes, the transmitting node and the receiving nodes have to synchronize at the carrier level so that the receiver can switch and decode with the correct codebook. However, this is difficult to be implemented in practice. One possible way is to adopt sub-optimal but practical MIMO broadcast channel encoder and decoder designs such as zero-forcing beamforming and minimum mean square error (MMSE) decoder.

Even if we can implement the MIMO broadcast channel code, another problem is related to resource allocation. Now, the weighted sum rate maximization problem is no longer guaranteed to have unique optimal solution. The derivation of the weighted sum rate maximization algorithm in Section 4.4 heavily relies on the factor that the weighted sum rate maximization problem for scalar Gaussian broadcast channel has unique optimal solution. This affects the convergence of the algorithm. Therefore, new resource allocation algorithms are needed for this cooperative transmission scheme. One way to solve the problem of the lack of strict convexity is the *proximal optimization technique* [10, p.233]. Briefly speaking, we introduce a strictly convex penalty function to the objective function. The new optimization problem has the same optimal solution but we can have better convergence results. This technique has been widely used in some resource allocation problems in wireless multi-hop networks such as [62, 109].

Besides the problems due to the need of time-sharing of dirty-paper codes, another problem related to resource allocation is that the amount of channel state information in MIMO is much more than the single antenna case. It is no longer practical to assume perfect knowledge of the channel condition when there are a large number of antennas and users. One way to solve the resource allocation problems is to use stochastic optimization formulations [102] and solve the corresponding stochastic optimization problems. For example, we can consider problems such as maximization of weighted sum of ergodic rates or we can have constraints of outage probabilities in the optimization model.

7.3 Presence of Selfish and Malicious Users

7.3.1 Selfish Behaviors

One of the assumptions of the works in this thesis is that the source nodes are willing to forward information for other source nodes. This may be true in some applications. However, in general, the users of a wireless network are selfish in nature. They are unwilling to use extra amount of power to forward other users' messages unless they can receive some incentives for that.

The pricing game in Chapter 5 may help solving this problem if the users are not malicious. That is, if a user helps forwarding the message for another user, he or she can receive the corresponding amount of payment as an incentive. The whole pricing mechanism can be hardwired in the mobile terminals as mentioned in [43].

Nevertheless, in general, it is not possible to guarantee that all mobile terminals are hardwired with the whole pricing mechanism. Malicious users can use other types of mobile terminals which provide fake pricing information to other users. Consider the example in Fig. 7.1. Suppose S_{N+1} is a malicious and selfish user. On one hand, since it is very close to its intended destination node D_{N+1}, it is optional for S_{N+1} to have messages relayed by other source nodes. But on the other hand, other source nodes will be benefited if S_{N+1} relays the messages originated from them. However, if S_{N+1} announces an unreasonably high price to

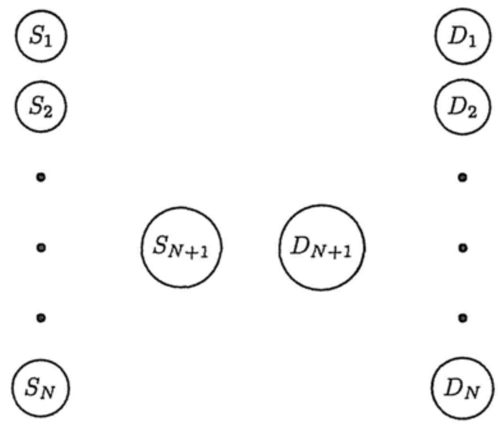

Figure 7.1: An example network for illustrating effects of malicious users.

other source nodes, none of them will request it to relay their messages. Another possibility is that after S_{N+1} receives the payment, it relays the message at a much lower rate than the optimal rate. Therefore, we need to modify the protocols proposed in Chapter 5 so that the nodes announce the correct prices and relay the messages at the optimal rate.

In some network protocols, such as BitTorrent [16], 'tit for tat' strategy is used to punish the malicious and selfish behaviors. In the previous example, if the nodes are fast moving nodes, it will be probable that for certain portion of time, S_{N+1} needs some other nodes to forward the information. Then, it will be suffered from the retaliation of other nodes.

Alternatively, better pricing can help solving the problem. If S_{N+1} announces an unreasonably high price, although none of the other nodes will request it to relay the information, it earns nothing from other nodes either. This is a loss-loss situation. Therefore, a natural question is how we can efficiently estimate the reasonable price so that S_{N+1} is willing to forward the information at a correct rate and other nodes are willing to pay for it. Some game-theoretic mechanism design techniques, such as Vickrey-Clarke-Groves (VCG) auction [78], may be helpful to answer this question.

7.3.2 Attacks from Malicious Users

Besides the problem of willingness of relaying information, malicious users can introduce different types of attacks to the cooperative communication system. For example, instead of relaying the original message, the malicious nodes can replace it with some other information so that the destination node receives the wrong message.

Here is a rough outline of the solution. Firstly, before we transmit the message through the cooperative transmission scheme, we encode it with another *error-correction code*. Although ordinary error-correction code can help correcting or detecting the incorrect message portion, it is possible that too many redundant bits are introduced. A malicious relay node can only change the information that it is asked to forward. Therefore, the error is localized to this portion of message. It is possible to design a code with higher coding rate by using this property. This is one of the possible research direction.

Besides error-correction code, on the system level, trust-based management systems [97] can also be introduced to avoid certain malicious relay nodes. The aforementioned error-correction does not only correct the wrong message portions, it also helps identifying some malicious nodes. After identifying a malicious node, we can report it to the trust-based management system, which is a mechanism to inform other nodes so that they can avoid the attacks from the malicious node. There are several issues to be addressed in the design of this mechanism. Firstly, the mechanism has to be robust against any malicious attacks. That is, malicious nodes cannot jeopardize the system by reporting fake information. Moreover, the mechanism has to be efficient so that the source nodes can be notified with the minimum latency.

7.4 Transmitter Co-opetition

Co-opetition [12] is a business neologism to describe the cooperations between the competing parties. That is, the competing parties work together for parts of their business where they do not believe they have competitive advantage, and where they believe they can share common costs. In *transmitter co-opetition*, part of the messages are transmitted by cooperative transmission but we allow the transmissions of the remaining part of the messages interfere with one another.

The motivation of combining these two different types of transmissions, instead of pure cooperative transmission, is the following. Similar to the case of multiple-input multiple-output (MIMO) systems, one cost of diversity gain by cooperative transmission is the reduction in degree of freedom [114]. Firstly, a relay node has to sacrifice part of its transmission time or bandwidth to overhear the messages of other nodes. As mentioned in Section 3.1, in practical systems, we can only implement half-duplex wireless nodes. Hence, in order to forward a message, we have to reduce the degree of freedom of the relay nodes.

Moreover, after being decoded by the relay node, the message has to be re-encoded and retransmitted at another set of time slots or frequency bands. Thus, compared with direct transmission, this relayed message has lower spectral efficiency. If the power gain of the direct link is high, the benefit from the spatial diversity may not compensate the loss from the reduction of degree of freedom. If the direct links become much higher power gains than the relay paths, instead of having any cooperations, it may be wiser to allow the sources to transmit in a non-cooperative manner.

Transmitter co-opetition schemes in multiple access (MAC) channel and interference channel are briefly introduced in the following parts of this section.

7.4.1 Multiple Access Channel

A MAC channel is a information-theoretic generalization of uplink transmission systems. Its channel capacity is derived in [1]. The Pareto-optimal set of rate vectors can be achieved by successive interference cancelation. The problem of this non-cooperative approach is that when a user suffers from deep fading, it does not only reduce the received power at the base station, but also affects the decoding order of this user. That is, his/her message is decoded under the interference of greater number of users. As a result, a direct uplink can convey a very small amount of information for this user. Relaying from other users is needed.

On the other hand, if the users have high power gain for their direct links to the base station but thy are separated from one another with large distance, the users, who relay the messages for the others, sacrifice their transmission time/bandwidth to overhear and forward a very small amount of information. In this situation, non-cooperative approach is preferable.

This is the motivation to derive transmission schemes which performs well in a more general channel conditions. One possible transmission scheme is as follows. Suppose there are N users in the system. The whole transmission time is divided into time frames of equal durations. Each time frame is divided into $N+1$ time slots of variable lengths. Each user divides his/her message into $N+1$ partitions. In the i-th time slot, $1 \leq i \leq N$, user i transmits his/her first N parts of the messages to the base station and other users by using superposition coding. Also, he or she forwards part of the buffered messages from other users to the base station. In the $(N+1)$-th time slot, all the users transmit the remaining partition of their own message and the remaining buffered messages from other users to the base station. The encoding and decoding are done in the same way as in MAC channels.

In this scheme, if the duration of the last time slot tends to zero, it approximates the special case in Chapter 5 where the receivers are collocated. On the other hand, if the durations of the first N time slots, it becomes closer and closer to a non-cooperative MAC channel. By optimizing the durations of all the time slots, a good balance of diversity and degree of freedom can probably be achieved in a practical fashion.

7.4.2 Interference Channel

Similar to the MAC channel case, we also have the question on how to strike the balance between diversity gain and multiplexing gain in interference channels by transmitter co-opetition. One of the motivation behind this question is as follows. Consider the case of two source-destination pairs. The two source nodes are assumed to be close to one another so that the power gains of the links between these two nodes are very high. Firstly, consider the strong interference channel case [34], where the interfering links have higher power gains than the direct links. The rates achieved by interference channel coding are limited by the direct link capacities. However, in cooperative transmission, the messages can be forwarded through the interfering links. Therefore, cooperative transmission may be more preferable.

However, if the direct links have much higher power gains than the interfering links, the rates of the relay paths are limited by the weak interfering link capacities. Furthermore, the relaying node needs to sacrifice their transmission time to listen to another source node before forwarding this marginal amount of information. On the other hand, since the interfering links have much smaller power gains, if the source nodes use interference channel coding, they do not interfere one another so much. They can make full use of their transmission time to transmit over their direct links which have high power gains. Therefore, in this case, interference channel coding may be more preferable. Hence, a natural question is: for a general wireless network, how should we transmit the messages?

The cooperative transmission scheme proposed in Chapter 4 can be extended and it can be compared with existing interference channel transmission schemes and some outer bounds. Unfortunately, the capacity region of a general Gaussian interference channel remains an open problem even for the case with two sources and two destinations. The capacity region is known only for the very strong interference channel case [14] and the strong interference channel case. The largest achievable rate region is the one proposed in [34]. A simpler achievable rate region is the one studied in [26]. The authors proved that if the rate vector (R_1, R_2) is in the capacity region of an interference channel, their scheme can achieve the rates $(R_1 - 1, R_2 - 1)$. The later will be adopted as a simple approximation of the capacity region of an interference channel.

Networks with two sources and two destinations can be the starting point of the whole studies. One possible approach is as follows. The whole transmission time is divided into time windows of equal duration. Each time window comprises of three time slots which have variable lengths. Each source node divides its message into 4 parts which will be explained shortly. In time slot i, $1 \leq i \leq 2$, only S_i transmits 3 of the 4 parts of its message. The first part, denoted by $M_{i,1}$, is transmitted directly to its destination node. The second part, denoted by $M_{i,2}$, is relayed by another source node in time slot $3-i$. The third part, denoted by $M_{i,3}$, is relayed by another source node in the third time slot. At the same time, it forwards the second part of another source node's message to the destination node. The source node multiplexes the messages over the three outgoing links by superposition coding.

In the third time slot, both source nodes transmit and they do not listen to each other.

The messages $M_{1,3}$ and $M_{2,3}$ are forwarded to the destination node. The remaining part of the messages are transmitted directly to the destination node. These messages are multiplexed in the following way. The fourth part of the messages is divided into two messages which are called *public messages* and *private messages.* The signals of these three messages are superimposed and transmitted simultaneously. Firstly, the public messages are decoded by treating all the interferences as noise. The public messages are encoded to the rate that both destination nodes can decode the two public messages.

After decoding the public messages, D_i can cancel the corresponding signals and move on to decode $M_{i,3}$. Since both source nodes have the knowledge of the messages $M_{1,3}$ and $M_{2,3}$, they can jointly encode the messages and this reduces to a MISO broadcast channel [105]. Although each source node has its own power constraint instead of the total power constraints of the antennas in usual MISO broadcast channels, the authors in [105] mention that the same coding technique can be applied and the capacity region can be characterized in a similar manner.

Since both source nodes jointly encode $M_{1,3}$ and $M_{2,3}$, they have the knowledge of the received signal of these messages at their destination nodes. Therefore, they can encode their public messages by dirty paper coding to remove the interference from the signals of $M_{1,3}$ and $M_{2,3}$. The destination nodes decode the public messages by treating the interference of the signal for another node's public message as noise.

The cooperative transmission scheme in Chapter 4 and the coding scheme for the interference channel in [26] become special cases of the above transmitter co-opetition scheme. Therefore, we can achieve a full order of diversity by this scheme. However, the tradeoff between the diversity order and the multiplexing gain is unknown. This is one research question to be answered.

After that, we can consider the partial cooperations among more interfering links. One practical approach is to consider partner selection protocols based on the above partial cooperative transmission scheme. Besides the aforementioned DF-based approach, other relaying approach or interference channel coding strategies can be considered and compared.

7.5 Multi-way Communication

In the works previously discussed in this thesis, the source nodes and the destination nodes are considered to be two disjoint sets of nodes. One research direction is to consider the case that these two sets coincide. That is, a coalition of nodes would like to exchange the independent messages with one another. This can also be viewed as a generalization of a two-way channel [91] with more than two terminals. Applications include the exchanging control messages in ad hoc networks and mesh networks, portable game consoles with multiple players and video conferencing. This explains why it has drawn a lot of attention in both industry and academic communities [19, 37, 81, 82, 113].

Information exchange via pure relay nodes has been widely studied [20, 37, 81, 82, 94, 108, 113]. Typically, there are two source nodes which would like exchange their messages through a number of pure relay nodes. The relay nodes have their messages to be exchanged with other nodes. Apart from broadcasting their own messages to others, the nodes help relaying the messages as well. In the research works referenced at the beginning of this paragraph, network coding [2] is applied and the authors have shown the improvements of achievable rates. Hence, instead of pure DF approach, the nodes encode and forward their received messages and their own messages.

The aim for this new research direction is to derive resource allocation algorithms for cooperative information exchanges among a set of nodes. This can be done by translating the problem to a multi-source multicast problem. Suppose there are N nodes. We can apply the directed acyclic graph (DAG) shown in Figure 7.2. Each node i in the original network is represented by three nodes S_i, R_i and D_i in the DAG, which represent the source state, relay state and destination state of node i respectively. The edges from S_i to R_i and the ones from R_i to D_i have infinite capacities. The edges from S_i to R_j and the ones from R_i to D_j, $i \neq j$, is equal to the capacity of the link from node i to node j in the original network.

In the network layer, the nodes can perform randomized linear network coding [38] for their received coded messages and their own messages. In the physical layer, the nodes can multiplex the outgoing links by superposition coding or other types of multiplexing. Due to

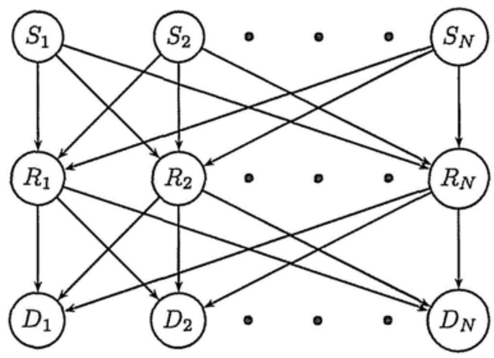

Figure 7.2: A single-source multicast model for information exchange problem among N nodes.

the half-duplex nature of the nodes, the portion of transmission time for each source node has to be scheduled. With the resource allocation algorithms designed based on the above model, we can have more ideas of the potential benefit of node cooperations in both physical layer and network layer over non-cooperative approaches.

Appendix A

Proofs of Theorems in Chapter 4

A.1 Proof of Proposition 4.2

Proof. (i) $\Rightarrow$ (ii): Immediate.

(ii) $\Rightarrow$ (iii): Assume that (R_1, R_2) is not Pareto optimal. There is an achievable rate pair (R_1', R_2') such that $R_1' \geq R_1$ and $R_2' \geq R_2$ and $(R_1, R_2) \neq (R_1', R_2')$. If $R_1' > R_1$ and $R_2' > R_2$, then (R_1, R_2) cannot be a boundary point since the rate region is comprehensive. Now consider the case $R_1' > R_1$ but $R_2 = R_2'$. Suppose that $R_2^{\max}$ is achieved by $(\bar{R}_1, R_2^{\max})$. If $\bar{R}_1 > R_1$, then (R_1, R_2) cannot be a boundary point by comprehensiveness. If $\bar{R}_1 \leq R_1$, we can find a convex combination

$$(R_1'', R_2'') = \alpha(\bar{R}_1, R_2^{\max}) + (1 - \alpha)(R_1', R_2') \tag{A.1}$$

with $R_1'' > R_1$ and $R_2'' > R_2$. Since (R_1'', R_2'') lies within the rate region by convexity, (R_1, R_2) cannot be boundary point by comprehensiveness again. The case that $R_2' > R_2$ but $R_1 = R_1'$ can be treated in the same way.

(iii) $\Rightarrow$ (i): Consider the set S in $\mathbb{R}_+^2$ consisting of the rate pair that Pareto dominates (R_1, R_2). S is a closed and convex cone. Since the interior of the rate region and the interior of S are disjoint, it follows from basic separation theorem in convex analysis [11, Sec. 2.5.1] that there is a hyperplane in $\mathbb{R}^2$ that separates them. The normal vector of the hyperplane is the required w_1 and w_2. $\square$

A.2 Proof of Lemma 4.8

Let $\mathcal{I}_i$ be the set of channels of user i. For $i = 1, 2$ and the $k \in \mathcal{I}_i$, let

$$u_{i1}^{(k)}(\boldsymbol{\mu}, z) = \frac{B^{(k)}\mu_i}{n_{i1}^{(k)} + z} - \lambda_i(\boldsymbol{\mu}), \tag{A.2}$$

$$u_{i2}^{(k)}(\boldsymbol{\mu}, z) = \frac{B^{(k)}w_i}{n_{i2}^{(k)} + z} - \lambda_i(\boldsymbol{\mu}), \tag{A.3}$$

$$u_{i3}^{(k)}(\boldsymbol{\mu}, z) = \frac{B^{(k)}(w_{3-i} - \mu_{3-i})}{n_{i3}^{(k)} + z} - \lambda_i(\boldsymbol{\mu}), \tag{A.4}$$

where $\lambda_i(\boldsymbol{\mu})$ is the solution of

$$\sum_{k\in\mathcal{I}_i} \left[\max\left\{\left(\frac{B^{(k)}\mu_i}{\lambda_i(\boldsymbol{\mu})} - n_{i1}^{(k)}\right), \left(\frac{B^{(k)}w_i}{\lambda_i(\boldsymbol{\mu})} - n_{i2}^{(k)}\right), \left(\frac{B^{(k)}(w_{3-i} - \mu_{3-i})}{\lambda_i(\boldsymbol{\mu})} - n_{i3}^{(k)}\right)\right\}\right]^+ = P_i \tag{A.5}$$

with $[x]^+ = \max\{x, 0\}$.

Furthermore, for $i = 1, 2$ and $k \in \mathcal{I}_i$, let

$$u_i^{(k)*}(\boldsymbol{\mu}) = \left[\max_{1\le j\le 3}\left\{u_{ij}^{(k)}(\boldsymbol{\mu}, z)\right\}\right]^+, \tag{A.6}$$

$$\mathcal{A}_{ij}^{(k)}(\boldsymbol{\mu}) = \left\{z \ge 0 | u_{ij}^{(k)}(\boldsymbol{\mu}, z) = u_i^{(k)*}(\boldsymbol{\mu})\right\}, \quad 1 \le j \le 3. \tag{A.7}$$

The optimal r_{ij} and R_{ij}, $i \ne j$, for a given $\boldsymbol{\mu}$ are given by [101]

$$r_{ij}^*(\boldsymbol{\mu}) = \sum_{k\in\mathcal{I}_i} \int_{\mathcal{A}_{i1}^{(k)}(\boldsymbol{\mu})} \frac{B^{(k)}}{n_{i1}^{(k)} + z} dz, \tag{A.8}$$

$$R_{ij}^*(\boldsymbol{\mu}) = \sum_{k\in\mathcal{I}_i} \int_{\mathcal{A}_{i3}^{(k)}(\boldsymbol{\mu})} \frac{B^{(k)}}{n_{i3}^{(k)} + z} dz. \tag{A.9}$$

As shown in [101], the set $\mathcal{A}_{ij}^{(k)}$ is an interval for all i, j, k. The continuity of r_{ij}^* and R_{ij}^* follows from (A.2)-(A.4), (A.6) and (A.7) together with the fact that an integral over an interval $[x, y]$ is a continuous function of x and y. This proves the first part of Lemma 4.8.

Let $\boldsymbol{\mu}^{(1)} = (\mu_1^{(1)}, \mu_2^{(1)})$ and $\boldsymbol{\mu}^{(2)} = (\mu_1^{(2)}, \mu_2^{(2)})$ such that $\mu_1^{(2)} = a\mu_1^{(1)}$, $a > 1$, $\mu_2^{(2)} = \mu_2^{(1)}$. Firstly, we would like to show that $\lambda_1(\boldsymbol{\mu}^{(2)}) \le a\lambda_1(\boldsymbol{\mu}^{(1)})$. To begin with, we observe that for all $k \in \mathcal{I}_1$,

$$\frac{B^{(k)}\mu_1^{(2)}}{a\lambda_1(\boldsymbol{\mu}^{(1)})} - n_{11}^{(k)} = \frac{B^{(k)}\mu_1^{(1)}}{\lambda_1(\boldsymbol{\mu}^{(1)})} - n_{11}^{(k)} \tag{A.10}$$

$$\frac{B^{(k)}w_1}{a\lambda_1(\boldsymbol{\mu}^{(1)})} - n_{12}^{(k)} < \frac{B^{(k)}w_1}{\lambda_1(\boldsymbol{\mu}^{(1)})} - n_{12}^{(k)} \tag{A.11}$$

$$\frac{B^{(k)}(w_2 - \mu_2^{(2)})}{a\lambda_1(\boldsymbol{\mu}^{(1)})} - n_{13}^{(k)} < \frac{B^{(k)}(w_2 - \mu_2^{(1)})}{\lambda_1(\boldsymbol{\mu}^{(1)})} - n_{13}^{(k)}. \tag{A.12}$$

Hence, from (A.5),

$$\sum_{k\in\mathcal{I}_1}\left[\max\left\{\left(\frac{B^{(k)}\mu_1^{(2)}}{a\lambda_1(\boldsymbol{\mu}^{(1)})} - n_{11}^{(k)}\right), \left(\frac{B^{(k)}w_1}{a\lambda_1(\boldsymbol{\mu}^{(1)})} - n_{12}^{(k)}\right), \left(\frac{B^{(k)}(w_2 - \mu_2^{(2)})}{2a\lambda_1(\boldsymbol{\mu}^{(1)})} - n_{13}^{(k)}\right)\right\}\right]^+ \leq P_1. \tag{A.13}$$

Since the left hand side of the (A.5) is a monotonic decreasing function of λ_1 and $\lambda_1(\boldsymbol{\mu}^{(2)})$ must satisfy (A.5), $\lambda_1(\boldsymbol{\mu}^{(2)}) \leq a\lambda_1(\boldsymbol{\mu}^{(1)})$. Similarly, we can also prove that $\lambda_1(\boldsymbol{\mu}^{(2)}) \geq \lambda_1(\boldsymbol{\mu}^{(1)})$. Thus, for all $k \in \mathcal{I}_1$ and z,

$$\frac{B^{(k)}\mu_1^{(2)}}{n_{11}^{(k)} + z} - \lambda_1(\boldsymbol{\mu}^{(2)}) = \frac{B^{(k)}a\mu_1^{(1)}}{n_{11}^{(k)} + z} - \lambda_1(\boldsymbol{\mu}^{(2)}) \tag{A.14}$$

$$\geq \frac{B^{(k)}a\mu_1^{(1)}}{n_{11}^{(k)} + z} - a\lambda_1(\boldsymbol{\mu}^{(1)}) \tag{A.15}$$

which implies that for all $k \in \mathcal{I}_1$ and z,

$$u_{11}^{(k)}(\boldsymbol{\mu}^{(2)}, z) \geq a u_{11}^{(k)}(\boldsymbol{\mu}^{(1)}, z). \tag{A.16}$$

As $\lambda_1(\boldsymbol{\mu}^{(2)}) \geq \lambda_1(\boldsymbol{\mu}^{(1)})$, it is trivial to see that for all $k \in \mathcal{I}_1$ and z,

$$u_{1j}^{(k)}(\boldsymbol{\mu}^{(2)}, z) \leq u_{1j}^{(k)}(\boldsymbol{\mu}^{(1)}, z), \quad j = 2, 3. \tag{A.17}$$

Hence, for all $k \in \mathcal{I}_1$,

$$\mathcal{A}_{11}^{(k)}(\boldsymbol{\mu}^{(1)}) \subseteq \mathcal{A}_{11}^{(k)}(\boldsymbol{\mu}^{(2)}). \tag{A.18}$$

Since the right hand side of (A.8) is the sum of integrals of non-negative function,

$$r_{12}^*(\boldsymbol{\mu}^{(2)}) \geq r_{12}^*(\boldsymbol{\mu}^{(2)}). \tag{A.19}$$

Suppose $z' \in \mathcal{A}_{13}^{(k)}(\boldsymbol{\mu}^{(2)})$, $k \in \mathcal{I}_1$.

$$u_{13}^{(k)}(\boldsymbol{\mu}^{(2)}, z') \geq u_{12}^{(k)}(\boldsymbol{\mu}^{(2)}, z') \tag{A.20}$$

$$\frac{B^{(k)}(w_2 - \mu_2^{(2)})}{n_{13}^{(k)} + z'} \geq \frac{B^{(k)}w_1}{n_{12}^{(k)} + z'} \tag{A.21}$$

$$\frac{B^{(k)}(w_2 - \mu_2^{(1)})}{n_{13}^{(k)} + z'} - \lambda_1(\boldsymbol{\mu}^{(1)}) = \frac{B^{(k)}(w_2 - \mu_2^{(2))}}{n_{13}^{(k)} + z'} - \lambda_1(\boldsymbol{\mu}^{(1)}) \geq \frac{B^{(k)}w_1}{n_{12}^{(k)} + z'} - \lambda_1(\boldsymbol{\mu}^{(1)}) \tag{A.22}$$

$$\therefore u_{13}^{(k)}(\boldsymbol{\mu}^{(1)}, z') \geq u_{12}^{(k)}(\boldsymbol{\mu}^{(1)}, z'). \tag{A.23}$$

Also, $u_{13}^{(k)}(\boldsymbol{\mu}^{(1)}, z') \geq u_{13}^{(k)}(\boldsymbol{\mu}^{(2)}, z') \geq u_{11}^{(k)}(\boldsymbol{\mu}^{(2)}, z') \geq u_{11}^{(k)}(\boldsymbol{\mu}^{(1)}, z')$. Therefore, $z' \in \mathcal{A}_{13}^{(k)}(\boldsymbol{\mu}^{(2)})$. This implies that

$$\mathcal{A}_{13}^{(k)}(\boldsymbol{\mu}^{(2)}) \subseteq \mathcal{A}_{13}^{(k)}(\boldsymbol{\mu}^{(1)}). \tag{A.24}$$

By the same argument as above,

$$R_{12}^*(\boldsymbol{\mu}^{(2)}) \leq R_{12}^*(\boldsymbol{\mu}^{(1)}). \tag{A.25}$$

The relationship of r_{12}^*, R_{12}^* and μ_2 can be proved in the same manner as above.

A.3 Proof of Lemma 4.9

This can be proved by mathematical induction on t. Initially, $\mu_1 = w_1$ and $\mu_2 = 0$. This implies that $R_{21}^*(\boldsymbol{\mu}(0)) = 0$ as $w_1 - \mu_1 = 0$ and $r_{21}^*(\boldsymbol{\mu}(0)) = 0$. Hence,

$$r_{12}^*(\boldsymbol{\mu}(0)) \geq 0 = R_{21}^*(\boldsymbol{\mu}(0)) \tag{A.26}$$

$$R_{12}^*(\boldsymbol{\mu}(0)) \geq 0 = r_{21}^*(\boldsymbol{\mu}(0)) \tag{A.27}$$

which means it is true for $t = 0$.

Suppose it is true for $t \geq t'$ for some $t' \geq 0$, i.e.

$$r_{12}^*(\boldsymbol{\mu}(t)) \geq R_{21}^*(\boldsymbol{\mu}(t)), \tag{A.28}$$

$$r_{21}^*(\boldsymbol{\mu}(t)) \leq R_{12}^*(\boldsymbol{\mu}(t)), \tag{A.29}$$

$\mu_1(t)$ is decreasing and $\mu_2(t)$ is increasing for $t \geq t'$.

Consider the iteration $t'+1$. Firstly, we update μ_1. According to the induction hypothesis, $r^*_{12}(\boldsymbol{\mu}(t')) \geq R^*_{21}(\boldsymbol{\mu}(t'))$. Therefore, we have to decrease μ_1 so that $r^*_{12}(\boldsymbol{\mu}(t'+1)) = R^*_{21}(\boldsymbol{\mu}(t'+1))$ according to Lemma 4.8.

According to Lemma 4.8 and the induction hypothesis, as μ_1 decreases,

$$r^*_{21}(\boldsymbol{\mu}(t'+1)) \leq r^*_{21}(\boldsymbol{\mu}(t')) \leq R^*_{12}(\boldsymbol{\mu}(t')) \leq R^*_{12}(\boldsymbol{\mu}(t'+1)). \tag{A.30}$$

Therefore, after updating μ_1,

$$r^*_{12}(\boldsymbol{\mu}(t'+1)) = R^*_{21}(\boldsymbol{\mu}(t'+1)) \tag{A.31}$$

$$r^*_{21}(\boldsymbol{\mu}(t'+1)) \leq R^*_{12}(\boldsymbol{\mu}(t'+1)) \tag{A.32}$$

$$\mu_1(t'+1) < \mu_1(t'). \tag{A.33}$$

Similarly, we can also prove that after updating μ_2,

$$r^*_{12}(\boldsymbol{\mu}(t'+1)) \geq R^*_{21}(\boldsymbol{\mu}(t'+1)) \tag{A.34}$$

$$r^*_{21}(\boldsymbol{\mu}(t'+1)) = R^*_{12}(\boldsymbol{\mu}(t'+1)) \tag{A.35}$$

$$\mu_2(t'+1) > \mu_2(t'). \tag{A.36}$$

Hence, the proposition is also true for $t = t'+1$. By mathematical induction, the proposition is true for all t.

A.4 Proof of Theorem 4.12

Firstly, optimal solution of a linear programming problem is a corner point of the feasible region [22]. Also, the origin is not possible to be the optimal solution because both α_1 and α_2 are positive. Hence, there are at most 3 candidate solutions.

Since the objective function is equivalent to maximize $\frac{\alpha_1}{\alpha_2} r_{12} + r_{21}$ subject to the same set of constraints, if $\frac{\alpha_1}{\alpha_2} \leq \underline{m}$, from (4.92) and (4.93),

$$\frac{\alpha_1}{\alpha_2} r_{12} + r_{21} \leq \min\left\{\frac{BP_1}{(\ln 2)\varphi_1}, \frac{BP_2}{(\ln 2)\chi_2}\right\} = r_{2,max}, \quad \forall r_{12}, r_{21}. \tag{A.37}$$

Equality holds if $r_{12} = 0$ and $r_{21} = r_{2,max}$. By similar argument, we can infer that if $\frac{\alpha_1}{\alpha_2} \geq \overline{m}$, $r_{12}^* = r_{1,max}$ and $r_{21}^* = 0$.

Now, consider $\underline{m} \leq \frac{\alpha_1}{\alpha_2} \leq \overline{m}$ and assume that $\chi_1\chi_2 - \varphi_1\varphi_2 > 0$. This implies that

$$\underline{m} = \frac{\varphi_2}{\chi_2} \leq \frac{\alpha_1}{\alpha_2} \leq \frac{\chi_1}{\varphi_1} = \overline{m}. \tag{A.38}$$

Suppose both $\tilde{r}_{12}$ and $\tilde{r}_{21}$ are nonnegative. By (4.99),

$$P_1\chi_2 - P_2\varphi_1 \geq 0 \tag{A.39}$$

$$\frac{P_2}{\chi_2} \leq \frac{P_1}{\varphi_1} \tag{A.40}$$

which means

$$r_{2,max} = \frac{BP_2}{(\ln 2)\chi_2}. \tag{A.41}$$

By (A.38),

$$\frac{\alpha_1}{\alpha_2}\chi_2(P_1\chi_2 - P_2\varphi_1) \geq \frac{\varphi_2}{\chi_2}\chi_2(P_1\chi_2 - P_2\varphi_1) \tag{A.42}$$

$$\frac{\alpha_1}{\alpha_2}(P_1\chi_2^2) - \frac{\alpha_1}{\alpha_2}(P_2\chi_2\varphi_1) - P_1\varphi_2\chi_2 + P_2\varphi_1\varphi_2 \geq \frac{\varphi_2}{\chi_2}\chi_2(P_1\chi_2 - P_2\varphi_1) - P_1\varphi_2\chi_2 + P_2\varphi_1\varphi_2 = 0 \tag{A.43}$$

$$\frac{\alpha_1}{\alpha_2}(P_1\chi_2) - \frac{\alpha_1}{\alpha_2}(P_2\varphi_1) - P_1\varphi_2 + \frac{P_2}{\chi_2}\varphi_1\varphi_2 + P_2\chi_1 - P_2\chi_1 \geq 0 \tag{A.44}$$

$$\frac{\alpha_1}{\alpha_2}\left(\frac{\chi_2}{\chi_1\chi_2 - \varphi_1\varphi_2}\right)P_1 - \frac{\alpha_1}{\alpha_2}\left(\frac{\varphi_1}{\chi_1\chi_2 - \varphi_1\varphi_2}\right)P_2 + \left(\frac{\chi_1}{\chi_1\chi_2 - \varphi_1\varphi_2}\right)P_2 - \left(\frac{\varphi_2}{\chi_1\chi_2\varphi_1\varphi_2}\right)P_1 - \frac{P_2}{\chi_2} \geq 0 \tag{A.45}$$

$$\frac{\alpha_1}{\alpha_2}\tilde{r}_{12} + \tilde{r}_{21} - r_{2,max} \geq 0 \tag{A.46}$$

$$\alpha_1\tilde{r}_{12} + \alpha_2\tilde{r}_{21} \geq \alpha_2 r_{2,max}. \tag{A.47}$$

Similarly, we can also prove that $\alpha_1\tilde{r}_{12} + \alpha_2\tilde{r}_{21} \geq \alpha_1 r_{1,max}$. Furthermore, it can be easily shown that equality holds for (4.92) and (4.93) if $r_{12} = \tilde{r}_{12}$ and $r_{21} = \tilde{r}_{21}$. Hence, if $\tilde{r}_{12}$ and $\tilde{r}_{21}$ are non-negative, it is one of the corner point of the feasible region. Therefore, if $\tilde{r}_{12}$ and $\tilde{r}_{21}$ are non-negative and $\chi_1\chi_2 - \varphi_1\varphi_2 > 0$, $r_{12}^* = \tilde{r}_{12}$ and $r_{21}^* = \tilde{r}_{21}$.

Suppose $\chi_1\chi_2 - \varphi_1\varphi_2 > 0$ and $\tilde{r}_{12} < 0$. By (4.99),

$$P_1\chi_2 - P_2\varphi_1 < 0 \tag{A.48}$$

$$\frac{P_1}{\varphi_1} < \frac{P_2}{\chi_2}. \tag{A.49}$$

Together with (A.38), we have

$$\frac{\varphi_2}{\chi_2}r_{12} + r_{21} \leq \frac{\chi_1}{\varphi_1}r_{12} + r_{21} \leq \frac{BP_1}{(\ln 2)\varphi_1} \leq \frac{BP_2}{(\ln 2)\chi_2}. \tag{A.50}$$

It means that (4.93) is redundant. Also, by (A.38),

$$\frac{\alpha_1}{\alpha_2}r_{12} + r_{21} \leq \frac{\chi_1}{\varphi_1}r_{12} + r_{21} \leq \frac{BP_1}{(\ln 2)\varphi_1}. \tag{A.51}$$

If $r_{12} = 0$ and $r_{21} = r_{2,max}$,

$$\frac{\alpha_1}{\alpha_2}r_{12} + r_{21} = \frac{BP_1}{(\ln 2)\varphi_1}. \tag{A.52}$$

This proves for the result for $\chi_1\chi_2 - \varphi_1\varphi_2 > 0$ and $\tilde{r}_{12} < 0$. Similar approach can prove for the result for $\chi_1\chi_2 - \varphi_1\varphi_2 > 0$ and $\tilde{r}_{12} < 0$. Since if $r_{12} = \tilde{r}_{12}$ and $r_{21} = \tilde{r}_{21}$, equality holds for (4.92) and (4.93), it is impossible to have $\tilde{r}_{12}, \tilde{r}_{21} < 0$. Hence, we have finished the proof for $\underline{m} \leq \frac{\alpha_1}{\alpha_2} \leq \overline{m}$ and $\chi_1\chi_2 - \varphi_1\varphi_2 > 0$. With similar approach, we can prove the result for $\underline{m} \leq \frac{\alpha_1}{\alpha_2} \leq \overline{m}$ and $\chi_1\chi_2 - \varphi_1\varphi_2 < 0$.

Appendix B

Joint Utility Maximization Framework

Consider a generic communication network consisting of N source-destination pairs. Let R_n be the rate of the n-th source-destination pair, $n = 1, 2, \ldots, N$, and $\mathbf{R}$ be the N-dimensional vector, $(R_1, R_2, \ldots, R_N)$. The set of all feasible rate vector in this network is denoted by $\mathcal{C} \subseteq \mathbb{R}_+^N$. We assume that $\mathcal{C}$ satisfies the following properties: (i) compact[1] and convex, and (ii) every Pareto-optimal rate vector is an extreme point.

The definitions of Pareto-optimal rate vectors and extreme points are provided below.

Definition B.1. *A point* $\mathbf{r} \in \mathcal{C}$ *is called* Pareto-optimal *if no component of* $\mathbf{r}$ *can be increased with the other components remaining fixed while remaining in* $\mathcal{C}$.

Definition B.2. *A point* $\mathbf{x} \in \mathcal{C}$ *is an* extreme point *of a convex set* $\mathcal{C}$ *if there is no way to express* $\mathbf{x}$ *as* $\alpha\mathbf{y} + (1-\alpha)\mathbf{z}$ *such that* $\mathbf{y}, \mathbf{z} \in \mathcal{C}$, $\mathbf{y} \neq \mathbf{x} \neq \mathbf{z}$, *and* $0 < \alpha < 1$.

The objective is to maximize a joint utility function $U(\mathbf{R})$ subject to $\mathbf{R} \in \mathcal{C}$. We assume that $U(\mathbf{R})$ satisfies the following properties: (i) strictly concave, (ii) twice continuously differentiable, and (iii) increasing with any component of $\mathbf{R}$ with the other components fixed. Since the feasible region $\mathcal{C}$ is assumed to be compact and convex, and U is strictly concave, there exists a unique solution.

For two vectors $\mathbf{x}$ and $\mathbf{y}$, we will write $\mathbf{x} \succeq \mathbf{y}$ if each component of $\mathbf{x}$ is larger than or equal to the corresponding component in $\mathbf{y}$. With this notation, the third assumption of the

[1] A set in an Euclidean space is compact if and only if it is bounded and closed.

utility function means that if $\mathbf{R}_1 \succeq \mathbf{R}_2$, then $U(\mathbf{R}_1) \geq U(\mathbf{R}_2)$.

We also assume that an algorithm which maximizes the weighted sum rate $\sum_{i=1}^{N} w_i R_i$ is available. The contribution of this paper is to use this weighted sum rate maximization algorithm as a building block and devise another algorithm that maximize any utility function that satisfies the three properties above.

The following examples that are special cases in our formulation.

Example B.1. *(Proportional Fairness in Parallel Gaussian Broadcast Channel) Consider an N-user Gaussian broadcast channel (BC) described in Section 2.1. For proportional fair rate allocation [52], the joint utility function is*

$$U_{pf}(\mathbf{R}) \triangleq \sum_{n=1}^{N} \log(R_n). \tag{B.1}$$

Our objective is to maximize $U(\mathbf{R})$ over the capacity region. The capacity region, by definition is closed. It is clearly bounded, and hence compact. It is convex, since a convex combination of any two achievable points can be achieved by time sharing between the two original points. Also, as mentioned in Section 2.1, it is shown that every Pareto-optimal rate vectors in the capacity region is also an extreme point. Using the algorithm for maximizing the weighted sum rate presented in [101], we will apply the method proposed in this paper and extend the weighted sum rate maximization algorithm to compute the proportional fair rate allocation.

Example B.2. *(QoS Feasible Region for Multiuser System) The last example is the class of power control problems studied in [87]. There are N users sharing a single channel. Let p_n be the transmission power of user n and $\mathbf{p} = (p_1, p_2, \dots, p_N)$. Also, let $\mathcal{I}_n(\mathbf{p})$ be the interference experienced by user i for a given transmission power vector $\mathbf{p}$. In [87], the* QoS feasible region *is defined as*

$$\mathcal{C} = \{\mathbf{R} : f(\mathbf{R}) \leq 1\}, \tag{B.2}$$

where

$$f(\mathbf{R}) = \inf_{\mathbf{p}>0} \left(\max_{1 \leq n \leq N} \frac{(e^{R_n} - 1)\mathcal{I}_n(\mathbf{p})}{p_n} \right). \tag{B.3}$$

It can be regarded as the rate region in our system model. Furthermore, in [87], the interference function is assumed to satisfy the following axioms.

A1 $\mathcal{I}_n(\mathbf{p}) \geq 0$

A2 $\mathcal{I}_n(\alpha\mathbf{p}) = \alpha\mathcal{I}_n(\mathbf{p})$

A3 $\mathcal{I}_n(\mathbf{p}^{(1)}) \geq \mathcal{I}_n(\mathbf{p}^{(2)})$ *if* $\mathbf{p}^{(1)} \geq \mathbf{p}^{(2)}$

A4 $\mathcal{I}_n(e^{\mathbf{s}})$ *is log-convex on* $\mathbb{R}^N$*, where* $e^{\mathbf{s}} = \mathbf{p}$.

The rate region is a compact set. Theorem 2 in [87] shows that with the above axioms, $\mathcal{C}$ is a strictly convex set. Hence, the Pareto-optimal rate vectors are extreme points. Therefore, it satisfies the assumption of the rate region in our system model.

We again decompose the dual problem into two subproblems by introducing of an auxiliary rate vector

$$\tilde{\mathbf{R}} \triangleq (\tilde{R}_1, \tilde{R}_2, \dots, \tilde{R}_N), \tag{B.4}$$

and reformulate the problem into:

$$\max_{\tilde{\mathbf{R}} \in \mathcal{B}} U(\tilde{\mathbf{R}}) \tag{B.5}$$

subject to

$$\mathbf{R} \succeq \tilde{\mathbf{R}} \tag{B.6}$$

$$\mathbf{R} \in \mathcal{C} \tag{B.7}$$

where $\mathcal{B}$ is a closed rectangular box in $\mathbb{R}_+^N$ of the form

$$\mathcal{B} \triangleq \{\tilde{\mathbf{R}} : \tilde{R}_n \in [0, b_n], n = 1, 2, \dots, N\} \tag{B.8}$$

that contains $\mathcal{C}$. Such a bounding box $\mathcal{B}$ exists because $\mathcal{C}$ is assumed to be compact, and hence bounded. Using the property that $U(\tilde{\mathbf{R}})$ is monotonically increasing for every component of $\tilde{\mathbf{R}}$, we can see that the reformulated version is in fact equivalent to the original version.

We relax (B.6) to form the partial Lagrangian:

$$L(\mathbf{R}, \tilde{\mathbf{R}}, \boldsymbol{\mu}) \triangleq U(\tilde{\mathbf{R}}) + \sum_{n=1}^{N} \mu_n (R_n - \tilde{R}_n), \tag{B.9}$$

where μ_n's are non-negative Lagrange multipliers. Denote the vector $(\mu_1, \mu_2, \ldots, \mu_N)$ by $\boldsymbol{\mu}$. The partial Lagrangian can be rearranged as

$$L(\mathbf{R}, \tilde{\mathbf{R}}, \boldsymbol{\mu}) = \left[U(\tilde{\mathbf{R}}) - \boldsymbol{\mu} \cdot \tilde{\mathbf{R}}\right] + \boldsymbol{\mu} \cdot \mathbf{R}. \tag{B.10}$$

Define the partial dual function $q(\boldsymbol{\mu})$ by

$$q(\boldsymbol{\mu}) \triangleq \max_{\mathbf{R} \in \mathcal{C}, \tilde{\mathbf{R}} \in \mathcal{B}} L(\mathbf{R}, \tilde{\mathbf{R}}, \boldsymbol{\mu}). \tag{B.11}$$

The dual problem is

$$\min_{\boldsymbol{\mu} \succeq \mathbf{0}} q(\boldsymbol{\mu}) \tag{B.12}$$

with the minimum taken over all nonnegative μ_n's. Computation of the partial dual function amounts to solving two independent optimization subproblems,

$$q(\boldsymbol{\mu}) = \max_{\tilde{\mathbf{R}} \in \mathcal{B}} \left\{U(\tilde{\mathbf{R}}) - \boldsymbol{\mu} \cdot \tilde{\mathbf{R}}\right\} + \max_{\mathbf{R} \in \mathcal{C}} \boldsymbol{\mu} \cdot \mathbf{R}. \tag{B.13}$$

Again, we decompose the problem into the following two subproblems.

Problem B.1.

$$\max_{\tilde{\mathbf{R}} \in \mathcal{B}} \left\{U(\tilde{\mathbf{R}}) - \boldsymbol{\mu} \cdot \tilde{\mathbf{R}}\right\}. \tag{B.14}$$

Problem B.2.

$$\max_{\mathbf{R} \in \mathcal{C}} \boldsymbol{\mu} \cdot \mathbf{R}. \tag{B.15}$$

Let $\tilde{\mathbf{R}}^*(\boldsymbol{\mu})$ and $\mathbf{R}^*(\boldsymbol{\mu})$ be the optimal $\tilde{R}$ and $\mathbf{R}$ for a fixed $\boldsymbol{\mu}$ respectively. Following a similar proof in Section 4.7.2, we have the algorithmic framework to solve the joint utility maximization problem in Algorithm 1.

In Step 6 of Algorithm 1, μ_n is increased if $\tilde{R}_n^*(\boldsymbol{\mu}) > R_n^*(\boldsymbol{\mu})$. Otherwise, μ_n is decreased.

Algorithm 1 Framework for Joint Utility Maximization

Initialize $\boldsymbol{\mu}$

$n \leftarrow 1$

while $\boldsymbol{\mu}$ has not converged **do**

  Initialization for bisection search of mu_n

  while $\left\|\tilde{\mathbf{R}}^*(\boldsymbol{\mu}) - \mathbf{R}^*(\boldsymbol{\mu})\right\|^2 < \epsilon$ **do**

    Bisection update of μ_n and its upper and lower bounds

    $\tilde{\mathbf{R}}^*(\boldsymbol{\mu}) \leftarrow \arg\max_{\tilde{\mathbf{R}} \in \mathcal{B}} \left\{U(\tilde{\mathbf{R}}) - \boldsymbol{\mu} \cdot \tilde{\mathbf{R}}\right\}$.

    $\mathbf{R}^*(\boldsymbol{\mu}) \leftarrow \arg\max_{\mathbf{R} \in \mathcal{C}} \boldsymbol{\mu} \cdot \mathbf{R}$.

  end while

  $n \leftarrow n + 1$

  if $n > N$ **then**

    $n \leftarrow 1$

  end if

end while

Appendix C

Proofs of Theorems in Chapter 5

C.1 Proof of Lemma 5.1

It can be directly implied by the results below. Consider a degraded Gaussian broadcast channel with N users. Let g_i be the power gain of user i and we assume that $g_1 > g_2 > \ldots > g_N$. Let r_i and p_i be the rate and transmission power to user i respectively. The one-sided power spectral density is assumed to be N_0 and the bandwidth is assumed to be W. The rates r_i should satisfy the following constraints [18]

$$r_i \leq \log\left(1 + \frac{g_i p_i}{\sum_{j=1}^{i-1} g_j p_j + N_0 W}\right), \quad i = 1, 2, \ldots, N. \tag{C.1}$$

Since

$$r_1 \leq \log\left(1 + \frac{g_1 p_1}{N_0 W}\right), \tag{C.2}$$

we have

$$\frac{p_1}{N_0 W} \geq \frac{1}{g_1}\left(e^{r_1} - 1\right). \tag{C.3}$$

Then, consider $i = 2$.

$$r_2 \le \log\left(1 + \frac{g_2 p_2}{g_1 p_1 + N_0 W}\right) \tag{C.4}$$

$$= \log\left(1 + \frac{\frac{g_2 p_2}{N_0 W}}{1 + \frac{g_1 p_1}{N_0 W}}\right) \tag{C.5}$$

$$\left(\frac{1}{g_2} + \frac{p_1}{N_0 W}\right)\left(e^{r_2} - 1\right) \le \frac{p_2}{N_0 W} \tag{C.6}$$

$$\frac{1}{g_2}\left(e^{r_2} - 1\right) + \frac{p_1}{N_0 W} e^{r_2} \le \frac{p_1 + p_2}{N_0 W} \tag{C.7}$$

$$\frac{1}{g_2}\left(e^{r_2} - 1\right) + \frac{1}{g_1} e^{r_2}\left(e^{r_1} - 1\right) \le \frac{p_1 + p_2}{N_0 W}. \tag{C.8}$$

By repeating this computation for $i = 3, 4, \ldots, N$, we can deduce that

$$\frac{1}{N_0 W} \sum_{i=1}^{N} p_i \ge \sum_{i=1}^{N} \frac{1}{g_i} \exp\left(\sum_{j=i+1}^{N} r_j\right)\left(e^{r_i} - 1\right). \tag{C.9}$$

The lower bound can be achieved by superposition coding because equality holds for the above inequalities if superposition coding is used. Therefore, the lower bound is achieved.

C.2 Proof of Proposition 5.4

Since $q_i(\lambda)$ is a concave function, $\frac{dq_i}{d\lambda}$ is a monotonically decreasing function.

Now, we prove the second part of the proposition. If $\lambda = \lambda_{\min}$,

$$\lambda \le \frac{1}{Z_{i,i}^{(D)}} = \Upsilon_{i,i}^{(D)}(\mathbf{0}) \tag{C.10}$$

$$\lambda \le \frac{1}{Z_{i,j}^{(S)}} + K_{j,i} = \Upsilon_{i,j}^{(S)}(0, \mathbf{0}), \quad j \ne i. \tag{C.11}$$

Hence,

$$\left[\ln\left(\frac{\lambda}{\Upsilon_{i,i}^{(D)}(\mathbf{0})}\right)\right]^+ = 0 \tag{C.12}$$

$$\left[\ln\left(\frac{\lambda}{\Upsilon_{i,j}^{(S)}(0, \mathbf{0})}\right)\right]^+ = 0, \quad j \ne i. \tag{C.13}$$

That means, $R^*_{i,D_i}(\lambda) = R^*_{i,S_j}(\lambda) = 0$ and

$$\left.\frac{dq_i}{d\lambda}\right|_{\lambda=\lambda_{\min}} = R_i > 0. \tag{C.14}$$

If $\lambda = \lambda_{\max}$,

$$\lambda \geq \Upsilon^{(D)}_{i,i}\left(\tilde{\mathbf{R}}^{(S)}_i\right)\exp\left(R_i\right). \tag{C.15}$$

C.3 Proof of Proposition 5.5

By Lemma 5.1, we have

$$\Gamma_i = \sum_{j=1}^{2M-1} \frac{1}{Z_{i,\pi_i(j)}} \exp\left(\sum_{k=j+1}^{2M-1} R_{i,\pi_i(k)}\right)\left[\exp\left(R_{i,\pi_i(j)}\right) - 1\right]. \tag{C.16}$$

For sufficiently small x, according to the Taylor's expansion, $e^x \approx 1 + x$. Hence, for sufficiently small $\triangle x$ and x, $e^{x+\triangle x} - e^x \approx x$. Therefore,

$$\Gamma_i \approx \sum_{j=1,j\neq i}^{M} \frac{1}{Z_{i,S_j}} R_{i,S_j} + \sum_{j=1}^{M} \frac{1}{Z_{i,D_j}} R_{i,D_j} \tag{C.17}$$

$$= \sum_{j=1,j\neq i}^{M} \frac{1}{Z_{i,S_j}} R_{i,S_j} + \sum_{j=1,j\neq i}^{M} \frac{1}{Z_{i,D_j}} R_{j,S_i} + \frac{1}{Z_{i,D_i}} R_{i,D_i} \tag{C.18}$$

$$= \sum_{j=1,j\neq i}^{M} \frac{1}{Z_{i,S_j}} R_{i,S_j} + \sum_{j=1,j\neq i}^{M} \frac{1}{Z_{i,D_j}} R_{j,S_i} + \frac{1}{Z_{i,D_i}} R_i - \sum_{j=1,j\neq i}^{M} \frac{1}{Z_{i,D_i}} R_{i,S_j} \tag{C.19}$$

$$= \sum_{j=1,j\neq i}^{M} \left(\frac{1}{Z_{i,S_j}} - \frac{1}{Z_{i,D_i}}\right) R_{i,S_j} + \sum_{j=1,j\neq i}^{M} \frac{1}{Z_{i,D_j}} R_{j,S_i} + \frac{1}{Z_{i,D_i}} R_i. \tag{C.20}$$

This implies that for sufficiently small R_i's,

$$\sum_{i=1}^{M} \Gamma_i = \sum_{i=1}^{M} \sum_{j=1,j\neq i}^{M} \left(\frac{1}{Z_{i,S_j}} - \frac{1}{Z_{i,D_i}}\right) R_{i,S_j} + \sum_{i=1}^{M} \sum_{j=1,j\neq i}^{M} \frac{1}{Z_{i,D_j}} R_{j,S_i} + \sum_{i=1}^{M} \frac{1}{Z_{i,D_i}} R_i \tag{C.21}$$

$$= \sum_{i=1}^{M} \sum_{j=1,j\neq i}^{M} \left(\frac{1}{Z_{i,S_j}} - \frac{1}{Z_{i,D_i}}\right) R_{i,S_j} + \sum_{i=1}^{M} \sum_{j=1,j\neq i}^{M} \frac{1}{Z_{j,D_i}} R_{i,S_j} + \sum_{i=1}^{M} \frac{1}{Z_{i,D_i}} R_i \tag{C.22}$$

$$= \sum_{i=1}^{M} \sum_{j=1,j\neq i}^{M} \left(\frac{1}{Z_{i,S_j}} - \frac{1}{Z_{i,D_i}} + \frac{1}{Z_{j,D_i}}\right) R_{i,S_j} + \sum_{i=1}^{M} \frac{1}{Z_{i,D_i}} R_i. \tag{C.23}$$

Appendix D

Proofs of Close Form Solutions in Section 6.2.1

Firstly, we prove the following results.

Lemma D.1. *If* $Z_{i,i}^{(D)} > Z_{i,3-i}^{(S)}$, $R_i^{(R)*} = 0$.

Proof. If $Z_{i,i}^{(D)} > Z_{i,3-i}^{(S)}$, $\delta_{i,i}^{(S)} = 0$. Then,

$$\begin{aligned}\frac{\partial \Gamma_i}{\partial R_i^{(R)}} &= -\frac{1}{Z_{i,i}^{(D)}} \exp\left(R_i^{(R)} + \delta_{i,3-i}^{(D)} R_{3-i}^{(R)}\right) + \frac{1}{Z_{i,3-i}^{(S)}} \exp\left(R_i^{(R)} + \delta_{i,3-i}^{(S)} R_{3-i}^{(R)}\right) \\ &+ \frac{1}{Z_{i,3-i}^{(D)}} \left(\delta_{i,3-i}^{(D)} - \delta_{i,3-i}^{(S)}\right) \exp\left[\left(1 - \delta_{i,3-i}^{(D)}\right) R + \left(\delta_{i,3-i}^{(D)} - \delta_{i,3-i}^{(S)}\right) R_i^{(R)}\right] \left(e^{R_{3-i}^{(R)}} - 1\right).\end{aligned} \tag{D.1}$$

For the cases $Z_{i,3-i}^{(D)} > Z_{i,i}^{(D)} > Z_{i,3-i}^{(S)}$, $Z_{i,i}^{(D)} > Z_{i,3-i}^{(D)} > Z_{i,3-i}^{(S)}$ and $Z_{i,i}^{(D)} > Z_{i,3-i}^{(S)} > Z_{i,3-i}^{(D)}$, by substituting the corresponding values of $\delta_{i,3-i}^{(D)}$ and $\delta_{i,3-i}^{(S)}$, we can easily show that $\frac{\partial \Gamma_i}{\partial R_i^{(R)}} > 0$ for all $0 \le R_i^{(R)} \le R$.

Moreover,

$$\Gamma_{3-i} = \frac{1}{Z^{(D)}_{3-i,3-i}} \exp\left[\left(1-\delta^{(S)}_{3-i,3-i}\right) R^{(R)}_{3-i} + \delta^{(D)}_{3-i} R^{(R)}_{i}\right]\left[\exp\left(R - R^{(R)}_{3-i}\right) - 1\right]$$
$$+ \frac{1}{Z^{(S)}_{3-i,i}} \exp\left[\delta^{(S)}_{3-i,3-i}\left(R - R^{(R)}_{3-i}\right) + \delta^{(S)}_{3-i} R^{(R)}_{i}\right]\left(e^{R^{(R)}_{3-i}} - 1\right)$$
$$+ \frac{1}{Z^{(D)}_{3-i,i}} \exp\left[\left(1-\delta^{(D)}_{3-i,i}\right) R + \left(\delta^{(D)}_{3-i,i} - \delta^{(S)}_{3-i,i}\right) R^{(R)}_{3-i}\right]\left(e^{R^{(R)}_{i}} - 1\right) \quad \text{(D.2)}$$

which is clearly a strictly increasing and convex function of $R^{(R)}_i$. Hence, $R^{(R)*}_i$ must be 0. □

Case I: $Z^{(D)}_{1,1} > Z^{(S)}_{1,2}$ and $Z^{(D)}_{2,2} > Z^{(S)}_{2,1}$

Direct consequence of Lemma D.1.

Case II: $Z^{(D)}_{i,3-i} > Z^{(D)}_{i,i} > Z^{(S)}_{i,3-i}$ and $Z^{(D)}_{3-i,3-i} < Z^{(S)}_{3-i,i}$

If $Z^{(D)}_{i,3-i} > Z^{(D)}_{i,i}$, $\delta^{(D)}_{i,3-i} = 0$. Also, by Lemma D.1, as $Z^{(D)}_{i,i} > Z^{(S)}_{i,3-i}$, $R^{(R)}_i = 0$. The optimization problem can thus be reduced to

$$\min_{R^{(R)}_2, R^{(D)}_2 \geq 0} \frac{1}{Z^{(D)}_{i,i}}\left(e^R - 1\right) + \frac{1}{Z^{(D)}_{i,3-i}} e^R \left[\exp\left(R^{(R)}_{3-i}\right) - 1\right]$$
$$+ \frac{1}{Z^{(D)}_{3-i,3-i}}\left[\exp\left(R^{(D)}_{3-i}\right) - 1\right] + \frac{1}{Z^{(S)}_{3-i,i}} \exp\left(R^{(D)}_{3-i}\right)\left[\exp\left(R^{(R)}_{3-i}\right) - 1\right] \quad \text{(D.3)}$$

subject to

$$R^{(R)}_{3-i} + R^{(D)}_{3-i} = R. \quad \text{(D.4)}$$

By eliminating the constant terms, (D.3) can be further simplified to

$$\min_{R^{(R)}_{3-i}, R^{(D)}_{3-i}} \left(\frac{1}{Z^{(D)}_{3-i,3-i}} - \frac{1}{Z^{(S)}_{3-i,i}}\right) \exp\left(R^{(D)}_{3-i}\right) + \frac{1}{Z^{(D)}_{i,3-i}} \exp\left(R + R^{(R)}_{3-i}\right). \quad \text{(D.5)}$$

Since the optimization problem is convex, the optimal solution, $R_{3-i}^{(R)*}$ and $R_{3-i}^{(D)*}$, satisfies the following Karush-Kuhn-Tucker (KKT) conditions:

$$\frac{1}{Z_{i,3-i}^{(D)}} \exp\left(R + R_{3-i}^{(R)}\right) + \mu - \lambda^{(R)} = 0 \tag{D.6}$$

$$\left(\frac{1}{Z_{3-i,3-i}^{(D)}} - \frac{1}{Z_{3-i,i}^{(S)}}\right) \exp\left(R_{3-i}^{(D)}\right) + \mu - \lambda^{(D)} = 0 \tag{D.7}$$

$$R_{3-i}^{(R)} + R_{3-i}^{(D)} = R \tag{D.8}$$

$$\lambda^{(R)} R_{3-i}^{(R)} = 0 \tag{D.9}$$

$$\lambda^{(D)} R_{3-i}^{(D)} = 0 \tag{D.10}$$

$$R_{3-i}^{(R)}, R_{3-i}^{(D)}, \lambda^{(R)}, \lambda^{(D)} \geq 0 \tag{D.11}$$

where μ is the Lagrange multiplier for (D.4), $\lambda^{(R)}$ and $\lambda^{(D)}$ are the Lagrange multipliers for $R_{3-i}^{(R)}$ and $R_{3-i}^{(D)}$ respectively.

It is obvious that $R_{3-i}^{(R)} \in [0, R]$. We consider three possible cases of $R_{3-i}^{(R)}$, namely, $R_{3-i}^{(R)} = 0$, $0 < R_{3-i}^{(R)} < R$ and $R_{3-i}^{(R)} = R$.

If $R_{3-i}^{(R)} = 0$, $R_{3-i}^{(D)} = R$ and thus $\lambda^{(D)} = 0$. Hence,

$$\mu = -\left(\frac{1}{Z_{3-i,i}^{(S)}} - \frac{1}{Z_{3-i,3-i}^{(D)}}\right) e^R \tag{D.12}$$

which implies that

$$\frac{1}{Z_{i,3-i}^{(D)}} e^R - \left(\frac{1}{Z_{3-i,i}^{(S)}} - \frac{1}{Z_{3-i,3-i}^{(D)}}\right) e^R = \lambda^{(R)} \geq 0 \tag{D.13}$$

$$\Rightarrow \log\left(\frac{1}{Z_{3-i,3-i}^{(D)}} - \frac{1}{Z_{3-i,i}^{(S)}}\right) - \log\left(\frac{1}{Z_{i,3-i}^{(D)}}\right) \leq 0. \tag{D.14}$$

Similarly, we can prove that if $0 < R_{3-i}^{(R)} < R$,

$$0 < R_{3-i}^{(R)} = \frac{1}{2}\left\{\log\left(\frac{1}{Z_{3-i,3-i}^{(D)}} - \frac{1}{Z_{3-i,i}^{(S)}}\right) - \log\left(\frac{1}{Z_{i,3-i}^{(D)}}\right)\right\} < R \tag{D.15}$$

and if $R_{3-i}^{(R)} = R$,

$$\frac{1}{2}\left\{\log\left(\frac{1}{Z_{3-i,3-i}^{(D)}} - \frac{1}{Z_{3-i,i}^{(S)}}\right) - \log\left(\frac{1}{Z_{i,3-i}^{(D)}}\right)\right\} \geq R. \tag{D.16}$$

This proves (6.12).

Case III: $Z_{i,i}^{(D)} > \max\left\{Z_{i,3-i}^{(S)}, Z_{i,3-i}^{(D)}\right\}$ **and** $Z_{3-i,3-i}^{(D)} < Z_{3-i,i}^{(S)}$

It can be proved in the same way as in the last case.

Case IV: $Z_{1,2}^{(S)} > Z_{1,1}^{(D)} > Z_{1,2}^{(D)}$ **and** $Z_{2,1}^{(S)} > Z_{2,2}^{(D)} > Z_{2,1}^{(D)}$

By substituting $\delta_{i,j}^{(S)} = \delta_{i,3-i}^{(D)} = 1$ for all i, j and some simplifications, the optimization problem becomes

$$\min_{R_i^{(R)}, R_i^{(D)}} \sum_{i=1}^{2} \left(\frac{1}{Z_{i,i}^{(D)}} - \frac{1}{Z_{i,3-i}^{(S)}} \right) \exp\left(R_i^{(D)} + R_{3-i}^{(R)} \right) + \sum_{i=1}^{2} \left(\frac{e^R}{Z_{i,3-i}^{(S)}} - \frac{1}{Z_{i,i}^{(D)}} + \frac{1}{Z_{i,3-i}^{(D)}} \right) \exp\left(R_{3-i}^{(R)} \right) \tag{D.17}$$

subject to

$$R_i^{(R)} + R_i^{(D)} = R, \quad i = 1, 2 \tag{D.18}$$

$$R_i^{(R)}, R_i^{(D)} \geq 0 \quad , i = 1, 2. \tag{D.19}$$

Since the optimization problem is convex, the optimal solution must satisfy the following

KKT conditions:

$$\left(\frac{1}{Z_{2,2}^{(D)}} - \frac{1}{Z_{2,1}^{(S)}}\right)\exp\left(R_2^{(D)} + R_1^{(R)}\right) + \left(\frac{e^R}{Z_{2,1}^{(S)}} - \frac{1}{Z_{2,2}^{(D)}} + \frac{1}{Z_{2,1}^{(D)}}\right)\exp\left(R_1^{(R)}\right) + \mu_1 - \lambda_1^{(R)} = 0 \tag{D.20}$$

$$\left(\frac{1}{Z_{1,1}^{(D)}} - \frac{1}{Z_{1,2}^{(S)}}\right)\exp\left(R_1^{(D)} + R_2^{(R)}\right) + \left(\frac{e^R}{Z_{1,2}^{(S)}} - \frac{1}{Z_{1,1}^{(D)}} + \frac{1}{Z_{1,2}^{(D)}}\right)\exp\left(R_2^{(R)}\right) + \mu_2 - \lambda_2^{(R)} = 0 \tag{D.21}$$

$$\left(\frac{1}{Z_{1,1}^{(D)}} - \frac{1}{Z_{1,2}^{(S)}}\right)\exp\left(R_1^{(D)} + R_2^{(R)}\right) + \mu_1 - \lambda_1^{(D)} = 0 \tag{D.22}$$

$$\left(\frac{1}{Z_{2,2}^{(D)}} - \frac{1}{Z_{2,1}^{(S)}}\right)\exp\left(R_2^{(D)} + R_1^{(R)}\right) + \mu_2 - \lambda_2^{(D)} = 0 \tag{D.23}$$

$$R_1^{(R)} + R_1^{(D)} = R \tag{D.24}$$

$$R_2^{(R)} + R_2^{(D)} = R \tag{D.25}$$

$$\lambda_1^{(R)} R_1^{(R)} = 0 \tag{D.26}$$

$$\lambda_2^{(R)} R_2^{(R)} = 0 \tag{D.27}$$

$$\lambda_1^{(D)} R_1^{(D)} = 0 \tag{D.28}$$

$$\lambda_2^{(D)} R_1^{(D)} = 0 \tag{D.29}$$

$$R_1^{(R)}, R_2^{(R)}, R_1^{(D)}, R_2^{(D)}, \lambda_1^{(R)}, \lambda_2^{(R)}, \lambda_1^{(D)}, \lambda_2^{(D)} \geq 0. \tag{D.30}$$

where μ_i is the Lagrange multiplier for the constraint $R_i^{(R)} + R_i^{(D)} = R$ and $\lambda_i^{(R)}$ and $\lambda_i^{(D)}$ are Lagrange multipliers for the nonnegative constraints of $R_i^{(R)}$ and $R_i^{(D)}$ respectively.

Firstly, we show that at least one of $R_1^{(R)}$ and $R_2^{(R)}$ is 0. Suppose $R_1^{(R)} \in (0, R]$, $\lambda_1^{(R)} = 0$. The first four KKT conditions become

$$\left(\frac{1}{Z_{2,2}^{(D)}} - \frac{1}{Z_{2,1}^{(S)}}\right) \exp\left(R_2^{(D)} + R_1^{(R)}\right) + \left(\frac{e^R}{Z_{2,1}^{(S)}} - \frac{1}{Z_{2,2}^{(D)}} + \frac{1}{Z_{2,1}^{(D)}}\right) \exp\left(R_1^{(R)}\right) + \mu_1 = 0 \tag{D.31}$$

$$\left(\frac{1}{Z_{1,1}^{(D)}} - \frac{1}{Z_{1,2}^{(S)}}\right) \exp\left(R_1^{(D)} + R_2^{(R)}\right) + \left(\frac{e^R}{Z_{1,2}^{(S)}} - \frac{1}{Z_{1,1}^{(D)}} + \frac{1}{Z_{1,2}^{(D)}}\right) \exp\left(R_2^{(R)}\right) + \mu_2 - \lambda_2^{(R)} = 0 \tag{D.32}$$

$$\left(\frac{1}{Z_{1,1}^{(D)}} - \frac{1}{Z_{1,2}^{(S)}}\right) \exp\left(R_1^{(D)} + R_2^{(R)}\right) + \mu_1 - \lambda_1^{(D)} = 0 \tag{D.33}$$

$$\left(\frac{1}{Z_{2,2}^{(D)}} - \frac{1}{Z_{2,1}^{(S)}}\right) \exp\left(R_2^{(D)} + R_1^{(R)}\right) + \mu_2 - \lambda_2^{(D)} = 0 \tag{D.34}$$

From (D.31) and (D.34), we obtain

$$\lambda_2^{(D)} - \mu_2 + \left(\frac{e^R}{Z_{2,1}^{(S)}} - \frac{1}{Z_{2,2}^{(D)}} + \frac{1}{Z_{2,1}^{(D)}}\right) \exp\left(R_1^{(R)}\right) + \mu_1 = 0. \tag{D.35}$$

Then, we combine it with (D.32) to eliminate μ_2 and obtain

$$\left(\frac{1}{Z_{1,1}^{(D)}} - \frac{1}{Z_{1,2}^{(S)}}\right) \exp\left(R_1^{(D)} + R_2^{(R)}\right) + \left(\frac{e^R}{Z_{1,2}^{(S)}} - \frac{1}{Z_{1,1}^{(D)}} + \frac{1}{Z_{1,2}^{(D)}}\right) \exp\left(R_2^{(R)}\right) - \lambda_2^{(R)} + \lambda_2^{(D)} + \left(\frac{e^R}{Z_{2,1}^{(S)}} - \frac{1}{Z_{2,2}^{(D)}} + \frac{1}{Z_{2,1}^{(D)}}\right) \exp\left(R_1^{(R)}\right) + \mu_1 = 0. \tag{D.36}$$

Finally, we combine it with (D.33) to eliminate μ_1 and obtain

$$\left(\frac{e^R}{Z_{1,2}^{(S)}} - \frac{1}{Z_{1,1}^{(D)}} + \frac{1}{Z_{1,2}^{(D)}}\right) \exp\left(R_2^{(R)}\right) - \lambda_2^{(R)} + \lambda_2^{(D)} + \left(\frac{e^R}{Z_{2,1}^{(S)}} - \frac{1}{Z_{2,2}^{(D)}} + \frac{1}{Z_{2,1}^{(D)}}\right) \exp\left(R_1^{(R)}\right) + \lambda_1^{(D)} = 0 \tag{D.37}$$

which implies that

$$\lambda_2^{(R)} - \lambda_2^{(D)} = \left(\frac{e^R}{Z_{1,2}^{(S)}} - \frac{1}{Z_{1,1}^{(D)}} + \frac{1}{Z_{1,2}^{(D)}}\right)\exp\left(R_2^{(R)}\right) + \left(\frac{e^R}{Z_{2,1}^{(S)}} - \frac{1}{Z_{2,2}^{(D)}} + \frac{1}{Z_{2,1}^{(D)}}\right)\exp\left(R_1^{(R)}\right) + \lambda_1^{(D)} > 0. \tag{D.38}$$

That means, $\lambda_2^{(R)} > \lambda_2^{(D)} \geq 0$. This implies that $R_2^{(R)} = 0$. Similarly, we can also prove that if $R_2^{(R)} > 0$, $R_1^{(R)} = 0$. This implies that at least one of the $R_i^{(R)}$'s is 0. If $R_i^{(R)*} = 0$,

$$\Gamma_T = \left(\frac{e^R - 1}{Z_{i,i}^{(D)}} + \frac{1}{Z_{i,3-i}^{(D)}}\right)\exp\left(R_{3-i}^{(R)}\right) + \left(\frac{1}{Z_{3-i,3-i}^{(D)}} - \frac{1}{Z_{3-i,i}^{(S)}}\right)\exp\left(R_{3-i}^{(D)}\right) + \left(\frac{e^R}{Z_{3-i,i}^{(S)}} - \frac{1}{Z_{i,3-i}^{(D)}} - \frac{1}{Z_{3-i,3-i}^{(D)}}\right). \tag{D.39}$$

Now, Γ_T is a strictly convex function of $R_{3-i}^{(R)}$ and $R_{3-i}^{(D)}$ so the optimal $R_{3-i}^{(R)}$ and $R_{3-i}^{(D)}$ can be determined as follows. We substitute $R_{3-i}^{(D)} = R - R_{3-i}^{(R)}$ into the above equation. Then, we derive Γ_T with respect to $R_{3-i}^{(R)}$ and set the derivative to 0. Finally, we project the obtained $R_{3-i}^{(R)}$ into the interval $[0, R]$. Then, we can immediately draw the same conclusion as in (6.15).

Case V: $Z_{1,2}^{(D)} > Z_{1,2}^{(S)} > Z_{1,1}^{(D)}$ and $Z_{2,1}^{(D)} > Z_{2,1}^{(S)} > Z_{2,2}^{(D)}$

By substituting $\delta_{i,i}^{(S)} = 1, \delta_{i,3-i}^{(S)} = \delta_{i,3-i}^{(D)} = 0$ and some simplifications, the optimization problem becomes

$$\min_{R_i^{(R)}, R_i^{(D)}} \left(\frac{1}{Z_{1,1}^{(D)}} - \frac{1}{Z_{1,2}^{(S)}}\right)\exp\left(R_1^{(D)}\right) + \frac{e^R}{Z_{2,1}^{(D)}}\exp\left(R_1^{(R)}\right) + \left(\frac{1}{Z_{2,2}^{(D)}} - \frac{1}{Z_{2,1}^{(S)}}\right)\exp\left(R_2^{(D)}\right) + \frac{e^R}{Z_{1,2}^{(D)}}\exp\left(R_2^{(R)}\right) \tag{D.40}$$

subject to

$$R_i^{(R)} + R_i^{(D)} = R, \quad i = 1, 2 \tag{D.41}$$

$$R_i^{(R)}, R_i^{(D)} \geq 0, \quad i = 1, 2. \tag{D.42}$$

It is easy to see that $R_1^{(R)}$ and $R_1^{(D)}$ are independent to $R_2^{(R)}$ and $R_2^{(D)}$. Hence, we can independently optimize these two sets of variables.

To optimize $R_1^{(R)}$ and $R_1^{(D)}$, we solve the following optimization subproblem,

$$\min_{R_1^{(R)}, R_1^{(D)} \geq 0} \left(\frac{1}{Z_{1,1}^{(D)}} - \frac{1}{Z_{1,2}^{(S)}}\right)\exp\left(R_1^{(D)}\right) + \frac{e^R}{Z_{2,1}^{(D)}}\exp\left(R_1^{(R)}\right) \tag{D.43}$$

subject to

$$R_1^{(R)} + R_1^{(D)} = R, \tag{D.44}$$

which is a convex optimization problem. The KKT conditions are

$$\frac{e^R}{Z_{2,1}^{(D)}} \exp\left(R_1^{(R)}\right) + \mu - \lambda^{(R)} = 0 \tag{D.45}$$

$$\left(\frac{1}{Z_{1,1}^{(D)}} - \frac{1}{Z_{1,2}^{(S)}}\right) \exp\left(R_1^{(D)}\right) + \mu - \lambda^{(D)} = 0 \tag{D.46}$$

$$R_1^{(R)} + R_1^{(D)} = 0 \tag{D.47}$$

$$\lambda^{(R)} R_1^{(R)} = 0 \tag{D.48}$$

$$\lambda^{(D)} R_1^{(D)} = 0 \tag{D.49}$$

$$R_1^{(R)}, R_1^{(D)}, \lambda^{(R)}, \lambda^{(D)} \geq 0. \tag{D.50}$$

If $R_1^{(R)} = 0$, $R_1^{(D)} = R$. Hence, $\lambda^{(D)} = 0$ and

$$\mu = -\left(\frac{1}{Z_{1,1}^{(D)}} - \frac{1}{Z_{1,2}^{(S)}}\right) e^R. \tag{D.51}$$

Then,

$$\frac{e^R}{Z_{2,1}^{(D)}} - \left(\frac{1}{Z_{1,1}^{(D)}} - \frac{1}{Z_{1,2}^{(S)}}\right) e^R = \lambda^{(R)} \geq 0 \tag{D.52}$$

which implies that

$$\log\left(\frac{1}{Z_{1,1}^{(D)}} - \frac{1}{Z_{1,2}^{(S)}}\right) - \log\left(\frac{1}{Z_{2,1}^{(D)}}\right) \leq 0. \tag{D.53}$$

Similarly, we can show that if $0 < R_1^{(R)} < R$,

$$0 < R_1^{(R)} = \frac{1}{2}\log\left(\frac{1}{Z_{1,1}^{(D)}} - \frac{1}{Z_{1,2}^{(S)}}\right) - \frac{1}{2}\log\left(\frac{1}{Z_{2,1}^{(D)}}\right) < R. \tag{D.54}$$

We can also show in the similar way that if $R_1^{(R)} = R$,

$$\frac{1}{2}\log\left(\frac{1}{Z_{1,1}^{(D)}} - \frac{1}{Z_{1,2}^{(S)}}\right) - \frac{1}{2}\log\left(\frac{1}{Z_{2,1}^{(D)}}\right) \geq R. \tag{D.55}$$

Hence, the optimal $R_1^{(R)}$ satisfies (6.16). By considering another optimization subproblem, we can also prove that the optimal $R_2^{(R)}$ satisfies (6.16).

Bibliography

[1] R. Ahlswede. Multi-way communication channels. In *Proc. IEEE ISIT 1971*, pages 103–135, 1971.

[2] R. Ahlswede, N. Cai, S.-Y. R. Li, and R. W. Yeung. Network information flow. *IEEE Trans. Inf. Theory*, 46(4):1204–1216, Ju 2000.

[3] E. Altman and Z. Altman. S-modular games and power control in wireless networks. *IEEE Trans. Autom. Control*, 50(2):291–303, May 2003.

[4] V. S. Annapureddy and V. V. Veeravalli. Gaussian interference networks: sum capacity in the low-interference regime and new outer bound on the capacity region. 55(7):3032–3050, July 2009.

[5] A. S. Avestimehr and D. N. C. Tse. Outage capacity of the fading relay in low snr regime. *IEEE Trans Inf. Theory*, 53(4):1401–1415, Apr 2007.

[6] K. Azarian, H. E. Gammal, and P. Schniter. On the achievable diversity-multiplexing tradeoff in half-duplex cooperative channels. *IEEE Trans. Inf. Theory*, 51(12):4152–4170, Dec 2005.

[7] C. Berge. *Topological Spaces.* The Macmillan Company, 1963.

[8] P. P. Bergmans. Random coding theorem for broadcast channels with degraded components. *IEEE Trans. Inf. Theory*, 19(2):197–207, 1973.

[9] P. P. Bergmans and T. M. Cover. Cooperative broadcasting. *IEEE Trans. Inf. Theory*, 20(3):317–324, May 1974.

[10] D. P. Bertsekas and J. N. Tsitsiklis. *Parallel and Distributed Computation Numerical Methods.* Prentice Hall, 1989.

[11] S. Boyd and L. Vandenberghe. *Convex Optimization.* Cambridge University Press, 2004.

[12] A. Brandenburger and B. Nalebuff. *Co-Opetition : A Revolution Mindset That Combines Competition and Cooperation.* Broadway Business, 1997.

[13] D. R. Brown III. Resource allocation for cooperative transmission in wireless networks with orthogonal users. In *Proc. Asilomar Conf. on Signals, Systems and Computers*, volume 2, pages 1473–1477, 2004.

[14] A. B. Carleial. A case where interference does not reduce capacity. *IEEE Trans. Inf. Theory*, IT-21:569–570, Sep 1975.

[15] Q. Chen and Z. Niu. A game-theoretical power and rate control for wireless ad hoc networks with step-up price. *IEICE Transactions on Communications*, E88-B(9):3515–3523, 2005.

[16] B. Cohen. Incentives build robustness in bittorrent. [Online] http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.14.1911, 2003.

[17] T. M. Cover and A. A. E. Gamal. Capacity theorems for the relay channel. *IEEE Trans Inf. Theory*, 25(9):572–584, Sep 1979.

[18] T. M. Cover and J. A. Thomas. *Elements of Information Theory.* Wiley-Interscience, 1991.

[19] T. Cui and T. Ho. Distributed space-time coding for two-way wireless relay networks. Technical report, California Institute of Technology, Sep 2007.

[20] T. Cui, T. Ho, and J. Kliewar. Some results on relay strategies for memoryless two-way relay channels. In *Proc. ITA 2008*, pages 158–164, Feb 2008.

[21] J. M. Danskin. *Theory of Max-Min and its Application in Weapon Allocation Problems.* Springer-Verlag, New York, 1967.

[22] G. B. Dantzig and M. N. Thapa. *Linear Programming.* Springer-Verlag, 1997.

[23] A. del Coso, U. Spagnolini, and C. Ibars. Cooperative distributed mimo channels in wireless sensor networks. *IEEE J. Sel. Areas in Commun.*, 25(2):402–414, Feb 2007.

[24] E. C. V. der Meulen. Three-terminal communication channels. *Adv. Appl. Prob.*, 3:120–154, 1971.

[25] K. B. Dey, S. Shatti, S. Jaggi, D. Katabi, and M. Médard. "Real" and "complex" network codes: Promises and challenges. In *Proc. IEEE NetCod 2008*, pages 1–6, Jan 2008.

[26] R. H. Etkin, D. N. C. Tse, and H. Wang. Gaussian interference channel capacity to within one bit. *IEEE Trans. Inf. Theory*, 54(12):5534–5562, Dec. 2008.

[27] N. Fawaz, D. Gesbert, and M. Debbah. When network coding and dirty paper coding meeet in a cooperative ad hoc network. *IEEE Trans. Wireless Commun.*, 7(5):1862–1867, May 2008.

[28] M. Florenzano and C. L. Van. *Finite Dimensional Convexity and Optimization.* Springer, 2001.

[29] G. Foschini. Layered space-time architecture for wireless communication in a fading environment when using multi-element antennas. *Bell Labs Tech. J.*, pages 41–59, 1996.

[30] H. N. Gabow. *Implementation of Algorithms for Maximum Matching on Nonbipartite Graphs.* PhD thesis, Standford University, 1973.

[31] R. G. Gallager. *Principles of Digital Communication.* Cambridge University Press, New York, 2008.

[32] A. E. Gamal and S. Zahedi. Capacity of a class of relay channels with orthogonal components. *IEEE Trans. Inf. Theory*, 51(5):1815–1817, May 2005.

[33] L. Grippo and M. Sciandrone. On the convergence of the block nonlinear gauss-seidel method under convex constraints. *Operations Research Letters*, 26:127–136, 2000.

[34] T. S. Han and K. Kobayashi. A new achievable rate region for the inteference channel. *IEEE Trans. Inf. Theory*, 27(1):49–60, Jan 1981.

[35] Z. Han, T. Himsoon, W. P. Siriwongpairat, and K. J. R. Liu. Energy-efficient co-operative transmission over multiuser ofdm networks: Who helps whom and how to cooperate. In *Proc. IEEE WCNC*, volume 2, pages 2131–2136, Mar 2005.

[36] Z. Han and H. V. Poor. Coalition Games with Cooperative Transmission: A Cure for the Curse of Boundary Nodes in Selfish Packet-Forwarding Wireless Networks. *ArXiv e-prints*, 704, Apr. 2007.

[37] C. Hausl and J. Hagenauer. Iterative network and channel decoding for the two-way relay channel. In *Proc. IEEE ICC 2006*, page 1568—1573, 2006.

[38] T. Ho, R. Koetter, M. Médard, M. Effros, J. Shi, and D. Karger. A random linear network coding approach to multicast. *IEEE Trans. Inf. Theory*, 52(10):4413–4430, Oct 2006.

[39] T. Holliday, A. J. Goldsmith, and P. Glynn. Distributed power control for time varying wireless networks: Optimality and convergence. In *Proceedings: Allerton Conference on Communications, Control and Computing*, 2003.

[40] A. Høst-Madsen. On the capacity of wireless relaying. In *Proc. VTC 2002 Fall*, volume 3, pages 1333–1337, 2002.

[41] A. Høst-Madsen. Capacity bounds for cooperative diversity. *IEEE Trans. Inf. Theory*, 52(4):1522–1544, Apr 2006.

[42] A. Høst-Madsen and J. Zhang. Capacity bounds and power allocation in wireless relay channel. *IEEE Trans. Inf. Theory*, 51(6):2020–2040, Jun 2005.

[43] J. Huang, R. A. Berry, and M. L. Honig. Distributed interference compensation for wireless networks. *IEEE JSAC*, 24(5):1074–1084, May 2006.

[44] D. Hughes-Hartog. *The capacity of a degraded spectral Gaussian broadcast channel.* PhD thesis, Standford University, Jul 1995.

[45] T. E. Hunter, S. Sanayei, and A. Nosratinia. Outage analysis of coded cooperation. *IEEE Trans. Inf. Theory*, 52(2):375–391, Feb 2006.

[46] O. Ileri, S. C. Mau, and N. B. Mandayam. Pricing for enabling forwarding in self-configuring ad hoc networks. *IEEE JSAC*, 23:151–162, Jan 2005.

[47] M. K. Jensen. Stability of pure strategy nash equilibrium in best-reply potential games. Technical report, University of Birmingham, Sep 2007.

[48] N. Jindal, U. Mitra, and A. Goldsmith. Capacity of ad-hoc networks with node cooperations. In *Proc. IEEE Int. Symp. Information Theory*, page 271, Jun 2004. detailed in http://www.standford.edu/class/ee360/chris.pdf.

[49] Y. S. Jung and J. H. Lee. Partner assignment algorithm for cooperative diversity in mobile communication systems. In *Proc. IEEE VTC 2006 Spring*, volume 4, pages 1610–1614, 2006.

[50] S. Katti, I. Marić, A. Goldsmith, D. Katabi, and M. Médard. Joint relaying and network coding in wireless networks. In *Proc. IEEE ISIT 2007*, pages 1101–1105, 2007.

[51] O. Kaya and S. Ulukus. Power control for fading cooperative multiple access channels. *IEEE Trans. Wireless Commun.*, 6(8):2915–2923, Aug 2007.

[52] F. Kelly. Charging and rate control for elastic traffic. *European Transactions on Telecommunications*, 8:33–37, 1997.

[53] M. A. Khojastepour, A. Sabharwal, and B. Aazhang. On capacity of gaussian 'cheap' relay channel. In *IEEE GLOBECOM'03*, volume 3, pages 1776 – 1780, Dec 2003.

[54] M. A. Khojastepour, A. Sabharwal, and B. Aazhang. Bounds on achievable rates for general multi-terminal networks with practical constraints. In F. Zhao and L. Guibas, editors, *Information Processing in Sensor Networks (IPSN '03)*, volume 2634 of *Lecture*

Notes in Computer Sciences, pages 146–161, Berlin Heidelberg, Apr. 2004. Springer-Verlag.

[55] Y.-H. Kim. Capacity of a class of deterministic relay channels. *IEEE Trans Inf. Theory*, 54(3):1328–1329, Mar 2008.

[56] J. N. Laneman. *Cooperative diversity in wireless networks: algorithms and architectures.* PhD thesis, Massachusetts Institute of Technology, Cambridge MA, 2002.

[57] 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*, 50(12):3062–3080, Dec 2004.

[58] J. N. Laneman and G. W. Wornell. Distributed space-time-coded protocols for exploiting cooperative diversity in wireless networks. *IEEE Trans. Inf. Theory*, 49(10):2415–2425, Oct 2003.

[59] A. Lapidoth, İ. E. Telatar, and R. Urbanke. On wide-band broadcast channels. 49(2):3250–3258, Dec. 2003.

[60] E. G. Larsson and B. R. Vojcic. Cooperative transmit diversity based on superposition coding. *IEEE Commun. Lett.*, 9(9):778–780, Sep 2005.

[61] M. Levorato, S. Tomasin, and M. Zorzi. Cooperative spatial multiplexing for ad hoc networks with hybrid arq: System design and performance analysis. *IEEE Trans. Commun.*, 56(9):1545–1555, Sep 2008.

[62] X. Lin and N. B. Shroff. Utility maximization for communication networks with multipath routing. *IEEE Trans. Automatic Control*, 51(5):766–781, May 2006.

[63] C. Long, Q. Zhang, B. Li, H. Yang, and X. Guan. Non-cooperative power control for wireless ad hoc networks with repeated games. *IEEE JSAC*, 25(6):1101–1112, Aug 2007.

[64] J. Luo, R. S. Blum, L. J. Greenstein, L. J. Cimini, and A. M. Haimovich. New approaches for cooperative use of multiple antennas in ad hoc wireless networks. In *Proc. IEEE VTC 2004-Fall*, volume 4, pages 26–29, 2004.

[65] V. Mahinthan, L. Cai, J. W. Mark, and X. Shen. Partner selection based on optimal power allocation in cooperative-diversity systems. *IEEE Trans. Veh. Technol.*, 57(1):511–520, Jan 2008.

[66] I. Marić, R. D. Yates, and G. Kramer. Capacity of interference channels with partial transmitter cooperation. *IEEE Trans. Inf. Theory*, 53(10):3536–3548, Oct 2007.

[67] S. Marthur, L. Sankar, and N. B. Mandayam. Coalitions in cooperative wireless networks. *IEEE J. Sel. Areas in Commun.*, 26(7):1104–1115, Sep 2008.

[68] J. Mo and J. Walrand. Fair end-to-end window-based congestion control. *IEEE/ACM Trans. on Networking*, 8:556–567, Oct. 2000.

[69] A. S. Motahari and A. K. Khandani. Capacity bounds for the Gaussian interference channel. 55(2):620–643, Feb. 2009.

[70] R. U. Nabar, H. Bolcskei, and F. W. Kneubuhler. Fading relay channels: performance limits and space-time signal design. *IEEE J. Sel. Areas Commun.*, 22(6):1099– 1109, Aug 2004.

[71] C. T. K. Ng and A. J. Goldsmith. Transmitter cooperation in ad-hoc wireless networks: Does dirty-payer coding beat relaying? In *Proc. IEEE ITW 2004*, pages 277–282, 2004.

[72] C. Y. Ng and T. M. Lok. Grouping algorithm for partner selection in cooperative transmission. To appear in IEEE ICUFN 2010.

[73] C. Y. Ng, T. M. Lok, and T. F. Wong. Pricing games for distributed cooperative transmission. Accepted by IEEE Trans. Veh. Technol.

[74] C. Y. Ng, T. M. Lok, and T. F. Wong. Pricing game for selfish link cooperation. In *Proc. IEEE ISIT 2008*, pages 2533–2537, Jul 2008.

[75] C. Y. Ng, K. W. Shum, C. W. Sung, and T. M. Lok. Joint utility optimization in parallel gaussian broadcast channel. Accepted by IEEE Trans. Veh. Technol.

[76] C. Y. Ng, K. W. Shum, C. W. Sung, and T. M. Lok. Rate allocation for cooperative orthogonal-division channels. Accepted by IEEE Trans. Commun.

[77] C. Y. Ng, C. W. Sung, and K. W. Shum. Rate allocation for cooperative transmission in parallel channels. In *IEEE GLOBECOM '07*, pages 3921–3925, 2007.

[78] N. Nisan, T. Roughgarden, E. Tardos, and V. V. Vazirani. *Algorithmic Game Theory.* Cambridge University Press, 2007.

[79] J. Nocedal and S. J. Wright. *Numerical Optimization.* Springer-Verlag, 1999.

[80] A. Nostratinia and T. E. Hunter. Grouping and partner selection in cooperative wireless networks. *IEEE J. Sel. Area. in Comm.*, 25(2):369–378, Feb 2007.

[81] B. Rankov and A. Wittneben. Spectral efficient signaling for half duplex relay channels. In *Proc. of Asilomar Conference on Signals, Systems and Computers*, pages 1066–1071, Oct 2005.

[82] B. Rankov and A. Wittneben. Achievable rate regions for the two-way relay channel. In *Proc. IEEE ISIT 2006*, pages 1668–1672, Jul 2006.

[83] A. Ribeiro, N. D. Sidiropoulos, G. B. Giannakis, and Y. Yu. Achieving wireline random access throughput in wireless networking via user cooperation. *IEEE Trans. Inf. Theory*, 53(2):732–758, Feb 2007.

[84] R. T. Rockafellar. *Convex Analysis.* Princeton University Press, New Jersey, 1970.

[85] C. U. Saraydar, N. B. Mandayam, and D. J. Goodman. Efficient power control via pricing in wireless data networks. *IEEE Trans. Commun.*, 50(2):291–303, Feb 2002.

[86] H. Sato. The capacity of the Gaussian interference channel under strong interference. 6(27):786–688, Nov. 1981.

[87] M. Schubert, H. Boche, and S. Stanczak. Strict convexity of the qos feasible region for log-convex interference functions. In *Proc. IEEE ACSSC 2006*, pages 478–482, Oct 2006.

[88] G. Scutari, S. Barbarossa, and D. P. Palomar. Potential games: A framework for vector power control problems with coupled constraints. In *Proc. IEEE ICASSP 2006*, volume 4, pages 241–244, May 2006.

[89] A. Sendonaris, E. Erkip, and B. Aazhang. Increasing uplink capacity via user cooperation diversity. In *Proc. IEEE Int. Symp. Information Theory*, page 156, Cambridge, M.A., Aug 1998.

[90] X. Shang, G. Kramer, and B. Chen. A new outer bound and he noisy-interference sum-rate capacity for the Gaussian interference channel. 2(55):689–699, Feb. 2009.

[91] C. E. Shannon. Two-way communication channels. In *Proc. 4th Berkeley Symp. Math. Stat. and Prob.*, volume 1, pages 611–644, 1961.

[92] N. Shastry and R. S. Adve. Stimulating cooperative diversity in wireless ad hoc networks through pricing. In *IEEE ICC '06*, 2006.

[93] K. Shum and C. W. Sung. Transmitter cooperation by recycling dirty paper. In *Proc. IEEE ISIT 2008*, pages 777–781, 2008.

[94] K. W. Shum and C. W. Sung. The capacity region of two-way collision networks. In *Proc. NetCod 2008*, pages 1–6, Jan 2008.

[95] W. P. Siriwongpairat, A. K. Sadek, and K. J. R. Liu. Bandwidth-efficient ofdm cooperative protocol with applications to uwb communications. In *Proc. IEEE WCNC 2006*, pages 1729–1734, 2006.

[96] A. Stefanov and E. Erkip. On the performance analysis of cooperative space-time coded systems. In *Proc. IEEE WCNC 2003*, volume 2, pages 729–734, 2003.

[97] Y. L. Sun, W. Yu, Z. Han, and K. J. R. Liu. Information theoretic framework of trust modeling and evaluation for ad hoc networks. *IEEE J. Sel. Areas. of Commun.*, 24(2):305–317, Feb 2006.

[98] C. W. Sung and W. S. Wong. A noncooperative power control game for multirate cdma data network. *IEEE Trans. Wireless Commun.*, 2(1):186–195, Jan 2003.

[99] W. P. Tam, T. M. Lok, and T. F. Wong. Power-minimizing rate allocation in cooperative uplink systems. *IEEE Trans. Veh. Technol.*, 58(9):4919–4929, Nov 2009.

[100] D. Tse and P. Viswanath. *Fundamentals of Wireless Communication.* Cambridge University Press, 2005.

[101] D. N. C. Tse. Optimal power allocation over parallel gaussian broadcast channels. URL: http://www.eecs.berkeley.edu/~dtse/broadcast2.pdf.

[102] S. P. Uryasev and P. M. Pardalos. *Stochastic Optimization: Algorithms and Applications.* Kluwer Academic Publishers, 2001.

[103] M. Voorneveld. Best-reply potential games. *Economics Letters*, 66:289–295, 2000.

[104] B. Wang, J. Zhang, and L. Zheng. Achievable rates and scaling laws of power-constrained wireless sensory relay networks. *IEEE Trans Inf. Theory*, 52(9):4084–4104, Sep 2006.

[105] H. Weingarten, Y. Steinberg, and S. Shamai. The capacity region of the gaussian multiple-input multiple-output broadcast channel. *IEEE Trans Inf. Theory*, 52(9):3936–3964, Sept 2006.

[106] F. M. J. Willems. The discrete memoryless multiple access channel with partially cooperating encoders. *IEEE Trans. Inf. Theory*, 29(3):441–445, May 1983.

[107] T. F. Wong, T. M. Lok, and J. M. Shea. Flow-optimized cooperative transmission for the relay channel, 2007.

[108] Y. Wu, P. A. Chou, and S.-Y. Kung. Information exchange in wireless networks with network coding and physical-layer broadcast. Technical report, MSR-TR-2004-78, Aug 2004.

[109] Y. Wu and S. Y. Kung. Distributed utility maximization for network coding based multicasting: A shortest path approach. *IEEE J. Sel. Areas Commun.*, 24(8):1475–1488, 2006.

[110] M. Yu and J. Li. Is amplify-and-forward practically better than decode-and-forward or vice versa. In *Proc. IEEE ICASP*, pages 365–368, 2005.

[111] W. Yu and J. M. Cioffi. Trellis precoding for the broadcast channel. In *Proc. IEEE GLOBECOM*, volume 2, pages 1344–1348, Nov 2001.

[112] R. Zamir, S. Shamai (Shitz), and U. Erez. Nested linear/lattice codes for structured multiterminal binning. *IEEE Trans. Inf. Theory*, 48(6):1250–1276, Jun 2002.

[113] S. Zhang, S. C. Liew, and P. P. Lam. Hot topic: physical-layer network coding. In *Proc. ACM Mobicom 2006*, pages 358–365, 2006.

[114] L. Zheng and D. Tse. Diversity and multiplexing: A fundamental tradeoff in multiple antenna channels. *IEEE Trans Inf. Theory*, 49(5):1073–1096, May 2003.