Distributed Beamforming in Wireless Relay Networks

by

Siavash Fazeli-Dehkordy

A thesis submitted to the

Department of Electrical and Computer Engineering

in conformity with the requirements for

the degree of Master of Science (Engineering)

Queen's University

Kingston, Ontario, Canada

September 2008

Abstract

In this thesis, we consider a wireless network consisting of d source-destination pairs and R relaying nodes. Each source wishes to communicate to its corresponding destination. By exploiting the spatial multiplexing capability of the wireless medium, we develop two cooperative beamforming schemes in order to establish wireless connections between multiple source-destination pairs through a collaborative relay network.

Our first communication scheme consists of two steps. In the first step, all sources transmit their signals simultaneously to the relay network. As a result, each relay receives a noisy faded mixture of all source signals. In the second step, each relay transmits an amplitude- and phase-adjusted version of its received signal, i.e., the relay received signals are multiplied by a set of complex coefficients and are retransmitted. Our goal is to obtain these complex coefficients (beamforming weights) through minimization of the total relay transmit power while the signal-to-interference-plus-noise ratio at the destinations are guaranteed to be above certain pre-defined thresholds.

Our second scheme is a distributed downlink beamforming technique which is performed in $d+1$ successive time slots. In the first d time slots, the d sources transmit their data to the relay network successively. The relay nodes receive and store the noisy faded versions of the source signals. In the $(d+1)$th time slot, the relays aim to collectively provide downlink connections to all d destinations. To do

so, each relay transmits a linear combination of the stored signals received during the first d time slots. Again, our goal is to determine the complex weights (used at the relaying nodes to linearly combine the source signals) by minimizing the total relay transmit power while satisfying certain quality of services at the destinations.

We use semi-definite relaxation to turn both problems into semi-definite programming (SDP) problems. Therefore, they can be efficiently solved using interior point methods. We showed that our proposed schemes significantly outperform orthogonal multiplexing schemes, such as time-division multiple access schemes, in a large range of network data rates.

Acknowledgments

It has been an exceptional experience to benefit from the supervision of Professor Saeed Gazor, and Professor Shahram Shahbazpanahi. I am very grateful to Professor Gazor for his insight, encouragement, and dedication to the personal and professional growth of his graduate students. My deepest gratitude goes to Professor Shahbazpanahi for his endless help, guidance, and inspiration during the course of my work. Despite the distance between Queen's University and UOIT, he has been a patient teacher and an excellent adviser. It is hard to imagine having done my M.Sc. without his unconditional support. I am also indebted to the other members of my thesis committee, Professor Il-Min Kim, Professor Glen Takahara, Professor Tom Dean, and Professor Richard Henriksen, for their valuable time and interest in my work.

Special thanks to my colleagues in the Advanced Multidimensional Signal Processing Laboratory for our inspiring discussions, friendship, and enjoyable moments over the past two years.

Finally, I would like to extend my deepest gratitude to my parents, my sisters, and my grandparents for their never-ending love, support, and encouragement.

Table of Contents

List of Figures

Chapter 1

Introduction

Wireless networking has been one of the fastest growing technologies in the last two decades which has triggered exponential research and investment. The growing demands of wireless networking, as well as new emerging applications, introduce many technical challenges in developing the next-generation wireless communication systems. Among different techniques proposed to address these challenges, the user cooperative schemes have been widely studied in recent years. Technological advances in communication hardware allow more complicated signal processing and coding techniques at mobile users and make the vision of cooperative communications a reality.

1.1 Cooperative Communications

Signal attenuation and interference are two major impairments of wireless networks. As the inherent properties of the wireless medium, they can lower the data rate and degrade the reliability of communication.

However, interference, caused by the broadcast nature of the wireless channel, can be beneficial by allowing mobile terminals to receive messages intended for other users in the network. By relaying messages for each other, mobile terminals can provide the final receiver with multiple replicas of the message signal arrived via different paths. These techniques, known as cooperative diversity [20, 21], are shown to significantly improve the network performance through mitigating the detrimental effects of signal fading.

Several schemes have been proposed to achieve spatial diversity through user cooperation [20], [26]. The most popular schemes are amplify-and forward (AF), decode-and-forward (DF), and coded cooperation [14].

1.2 Thesis Contributions

In this thesis, we consider a wireless network consisting of d source-destination pairs and R relaying nodes. Each pair wishes to communicate through the collaborative relay network.

A straightforward approach to establish such connections is to have different sources transmit their data over orthogonal or semi-orthogonal channels. Examples of such (semi-) orthogonal channels are orthogonal frequency division multiple access (OFDMA), time division multiple access (TDMA), or code division multiple access (CDMA) schemes. The relays are then required to receive the signal transmitted over each of these channels and amplify-and-forward that on the same channel. Each destination tunes to the corresponding channel to retrieve its data.

However, similar to traditional multiple antenna systems, the other potential benefit of cooperative communication is to increase the spectral efficiency of the wireless channel through providing additional degrees of freedom for communication (spatial multiplexing). We develop two cooperative beamforming schemes to exploit the spatial multiplexing capability of the wireless medium to establish wireless connections between multiple source-destination pairs through a relay network. The role of relays is to mitigate the cross-link interference and establish wireless connections between sources and their respective destinations. If the sources, destinations, and relays are distributed in the space, our proposed multiplexing schemes allow multiple source-destination pairs to share the communication resources through the spatial multiplexing [5, 6].

In Chapter 3, we propose a cooperative communication scheme consisting of two steps. In the first step, all sources transmit their data to the relays at the same time. In the second step, each relay transmits an amplified and phase-adjusted version of its received signal. Assuming that the second order statistics (i.e., correlation matrices) of all communication channels are available, we calculate the complex gains of the relays such that the total power dissipated by the relays is minimized, and at the same time, the signal-to-interference-plus-noise ratios (SINRs) at all destinations are above predefined thresholds.

Although this approach may seem similar to those used in the transmit (downlink) beamforming literature [1, 2, 4, 9, 28–32, 35], there are major differences between the former schemes and our communication scheme. One important difference is that in transmit beamforming, the signal intended for each user (destination) is exactly known at the transmitting antennas, while in our communication scheme, only noisy faded versions of source signals are available at the relays.

Another major difference is that in downlink beamforming schemes, the signal intended for each user is available separately to the transmitting antennas. This allows downlink beamforming schemes to use different weight vectors for different users, thereby forming one beam per user. However, in our approach, all the sources transmit their signals at the same time. Therefore, each relay receives a noisy mixture of the source signals. As a result, we can only adjust one weight vector (one beam) to cover all the destinations.

The cooperative scheme proposed in Chapter 4 is composed of $(d+1)$ time slots. In the first d time slot, all sources transmit their data to the relay network in successive time slots. In the next step, each relay multiplies each received signal by a complex weight, adds them all together and forwards the result toward the destinations. Our goal is to find the optimal complex weights at the relays to minimize the total power dissipated by the relay network. Again, the signal-to-interference- plus-noise ratios (SINRs) at all destinations are guaranteed to be above certain thresholds.

We show that using a semidefinite relaxation approach, the power minimization problems in Chapters 3 and 4 can be turned into semi-definite programming (SDP) problems, and therefore, they can be solved efficiently using interior point methods. The numerical examples verify that our proposed scheme significantly outperforms time-division multiplexing schemes in a large range of network data rates.

1.3 Thesis Outline

The remainder of this thesis is organized as follows. Chapter 2 provides some necessary background and a brief review of previous work in cooperative communications. Our proposed cooperative beamforming schemes are presented in Chapters 3 and 4. In Chapter 5, we conclude this thesis and discuss some limitations and practical issues to be considered in future research.

1.4 Notations

We use uppercase boldface letters to represent matrices and lowercase bold letters to denote vectors. We denote complex conjugate, transpose, and Hermitian (conjugate) transpose by $(\cdot)^*$, $(\cdot)^T$, and $(\cdot)^H$, respectively. We use $E\{\cdot\}$ to denote statistical expectation, $\delta_{rr'}$ to represent Kronecker's delta function, $\text{tr}\{\cdot\}$ and $\text{rank}(\cdot)$ represent the trace and the rank of a matrix, respectively, and $[\cdot]_{r,r'}$ denotes the (r, r') entry of a matrix. $\text{diag}(\mathbf{A})$ is a vector which contains the diagonal entries of the square matrix $\mathbf{A}$, $\text{diag}(\mathbf{a})$ denotes a diagonal matrix with the elements of the vector $\mathbf{a}$ as its diagonal entries, and $\text{vec}(\mathbf{A})$ stacks all columns of matrix $\mathbf{A}$ on top of each other. $\mathbf{A} \succeq 0$ means that $\mathbf{A}$ is a positive semi-definite matrix, $\odot$ stands for Schur-Hadamard (element-wise) multiplication of two matrices, and $\boldsymbol{\lambda} \succcurlyeq 0$ indicates that all entries of $\boldsymbol{\lambda}$ are non-negative.

Chapter 2

Background

2.1 Multi-Antenna Communication Systems

In addition to noise, interference, and other impairments which are inherent to all communication channels, wireless terminals suffer from random variation of the received signal power when they move from one place to another. In fact, during each transmission, there is significant probability that the wireless channel will go into a deep fade which likely results in the erroneous detection of the transmitted information. Increasing the transmitted power is a natural way to compensate for the fading and to satisfy reliability requirements suited for different applications (usually measured by bit error rate or frame error rate). However the error probabilities in different communication schemes are almost proportional to the inverse of the transmitted power. Thus, achieving tolerable BER constraints would result in huge power penalties on the wireless communication systems. A powerful technique to tackle the fading problem is diversity. In fact, if the information symbols pass through independently faded channels, we can retrieve the transmitted symbol if at least one of the paths is strong [37]. Since the probability that all paths experience a deep fade at the same time is much lower than a single path, diversity techniques drastically lower the error probability specially at the high SNR regime.

There are different ways to achieve diversity. Time diversity is achieved via coding and interleaving. If the separation between two successive code symbols, the size of the interleaver, is larger than the coherence time of the channel, different symbols experience almost independent fades. Interleaving spreads the error burst caused by a deep fade over different codewords which results in few symbol errors in each codeword. Then, these few errors can be corrected by the decoding step [8].

Frequency diversity is possible when the transmission bandwidth is larger than the coherence bandwidth of the wireless channel. In these channels, signals from different paths arrive at different symbol times. These independently faded versions of the transmitted symbol arrived from different paths can be resolved at the receiver with proper equalizers to achieve frequency diversity [8].

Another way of achieving independently faded symbols is to use multiple receive or transmit antennas at different locations in the space. To achieve independent fading paths, and hence to achieve the so called spatial diversity, in the wireless system, the receive or transmit antennas should have enough separation. Implementing multiple transmit antennas results in a similar system model and diversity gain when the channel gains are known. However, when the channel gains are unknown, space-time coding should be used to achieve spatial diversity [8].

In addition to the diversity gain, multiple antennas at the transmitter and receiver can dramatically increase the data rate by creating virtual decoupled channels over which independent data can be transmitted.

In fact, it is shown that the MIMO channel capacity grows linearly with the minimum number of transmit and receive antennas [25]. This result holds even when the channel matrix is unknown at the transmitter. This capacity gain over regular SISO channels is called the multiplexing gain [41].

Multiple antennas can also be used at the base station to significantly increase the spectral efficiency in the context of multiuser communications. In fact, exploiting additional degrees of freedom from having multiple antennas enables different users to transmit or receive at the same time in the same frequency band. This topic will be more explored in the next two sections.

2.2 Multiuser Systems

Multiple access and interference management are two major considerations in multiuser systems [37]. Multiple access techniques deal with how the communication resource is shared among different users; while interference management techniques try to mitigate the interference among the users transmitting at the same time.

The available communication resource can be divided up among multiple users in several ways: time-division multiple access (TDMA), frequency-division multiple access (FDMA), and code-division multiple access (CDMA). These schemes divide the signaling space along different dimensions (time, frequency, and code, respectively) [8].

In TDMA systems, the time domain is divided into consecutive TDMA frames. One TDMA frame consists of several time slots of the same length. Each time slot is then assigned to a different user. In FDMA systems, non-overlapping frequency bands are assigned to different users. In fact, TDMA and FDMA are orthogonal multiplexing methods in the time and frequency domains, respectively. In CDMA systems, the data signal is modulated by a pseudo noise sequence and transmitted over the whole system bandwidth. CDMA can be either orthogonal or non-orthogonal. Space is another communication resource that can be used to separate different users and channelize the wireless medium. Space division multiple access techniques will be considered in the next section.

2.3 Receive and Transmit Beamforming

As previously stated, multiple antennas provide additional degrees of freedom for wireless channels and hence, independent data streams can be sent over the same frequency band without significant interference among different streams. In several transmitter-receiver architectures proposed in the MIMO communication literature, there is no cooperation across transmit antennas and an independent data stream is transmitted at each transmit antenna. Therefore, a similar transceiver architecture can be used in a multiuser system with multiple mobile terminals and a base station with multiple receive antennas. In the uplink, each user can be viewed as a transmit antenna in a point-to-point MIMO system and the same receiver architecture can be used at the same base station to separate each user's data (receive beamforming). A similar strategy can be used in the downlink when the base station broadcasts independent data streams to different mobile users. In this case, different receive antennas are at different users and hence MIMO receiver structure can not be implemented. However, by exploiting an interesting duality between the uplink and the downlink, each downlink beamforming problem is turned into a virtual uplink problem and, as a result, each receive beamforming strategy in the uplink has a corresponding receive beamforming strategy in the downlink [1, 37].

The baseband model for a time-invariant uplink channel is:

$$\mathbf{y} = \sum_{k=1}^{d} \mathbf{h}_k s_k + \mathbf{w} \tag{2.1}$$

where the vector $\mathbf{y}$ represents the received signals at different receive antennas at the base station, s_k is the transmitted symbol by user k, d is the number of users in the system and w denotes additive independent and identically distributed (i.i.d.) noise. The vector $\mathbf{h}_k$ denotes the channel gain vector from user k to the receive antennas at the base station. In the beamforming literature, $\mathbf{h}_k$ is also called the spatial signature of user k.

The receiver at the base station can separate different users' transmitted signals because of their different spatial signatures on the receive antenna array. Furthermore, downlink or transmit beamforming can be employed at a base station with multiple transmit antennas to simultaneously transmit data to multiple users in the network (Figure 2.1). Each mobile user is assumed to have a single receive antenna. The vector of transmitted signals at the base station antenna array is given by:

$$\mathbf{x} = \sum_{k=1}^{d} \mathbf{w}_k s_k \tag{2.2}$$

where s_k is the signal intended for user k, and $\mathbf{w}_k$ is the kth beamforming vector. For simplicity, we assume that there is only one base station in the system. However, the extension of this problem to a system with multiple base stations is straightforward [1]. The received signal at user k is given by:

$$y_k = {\mathbf{h}_k}^H \mathbf{x} + v_k \tag{2.3}$$

where ${\mathbf{h}_k}^H$ represents the channel gain vector from the base station to user k, and v_k is additive noise that is i.i.d. in time with power ${\sigma_k}^2$. In order to design beamforming

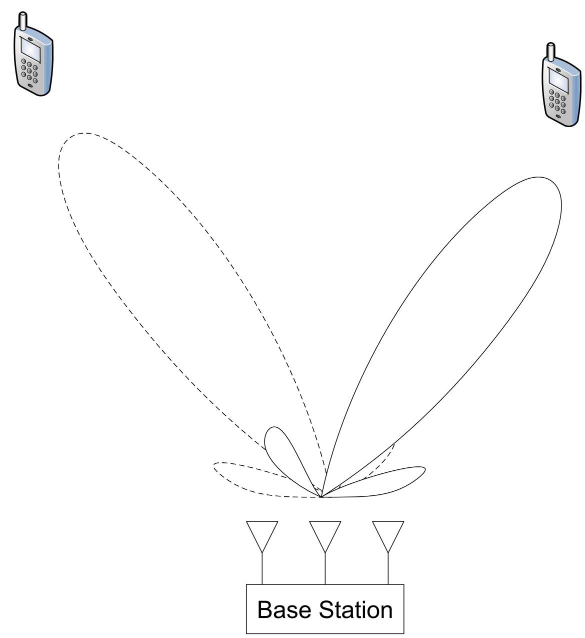

Figure 2.1: Downlink Beamforming

vectors, an estimate of the downlink channel is needed. In time division duplex (TDD) systems, the uplink channel can be measured at the base station and due to channel reciprocity, it provides an estimate of the downlink channel. However, this estimate can be erroneous due to the time distance between the uplink and downlink transmissions. In frequency division duplex (FDD) systems, the uplink and downlink channels are generally different and feedback from the mobile users is required. This would be impractical when the number of users or the number of transmit antennas is large.

Because of these limitations, it is rarely possible to obtain accurate estimates of the downlink channel vector. By averaging the uplink data over a long period of time (with respect to fast fading fluctuations), we can obtain a good estimate of the correlation matrix of the uplink channel, which is also a good estimate of the downlink channel in TDD systems. The averaging still preserves the information about the shadow fading and the direction of the incoming wave. However, some transformations may be required in order to obtain an estimate of the downlink channel from the uplink data [1].

The beamforming weight vectors in (2.2) should be found such that an acceptable quality of Service (QoS) is guaranteed at the receivers. The optimal beamforming vectors are those which minimize a cost function while satisfying QoS constraints. In the beamforming literature, Signal to Interference plus Noise Ratio (SINR) is commonly used as a measure of QoS and the cost function is the total transmit power at the base station. The main thrust of optimal beamforming is to maximize the transmitted energy toward the intended user as well as mitigating the interference to other users. The received signal at user i can be written as:

$$y_i = \underbrace{\mathbf{h}_i{}^H \mathbf{w}_i s_i}_{\text{desired signal}} + \underbrace{\mathbf{h}_i{}^H \sum_{k \neq i} \mathbf{w}_k s_k}_{\text{interference}} + \underbrace{v_i}_{\text{noise}}. \quad (2.4)$$

Assuming that only the correlation matrices of the downlink channel vectors are known at the base station, the average SINR at user i is given by:

$$\text{SINR}_i = \frac{\mathbf{w}_i{}^H \mathbf{R}_i \mathbf{w}_i}{\sum_{k \neq i} \mathbf{w}_k{}^H \mathbf{R}_i \mathbf{w}_k + \sigma_i{}^2}. \quad (2.5)$$

where $\mathbf{R}_i$ denotes the covariance matrix of the channel vector from the base station to user i. The data signals are assumed to be uncorrelated with unit power, i.e., $E\{s_i s_i{}^*\} = 1$. The total transmit power at the base station is given by:

$$P_T = \sum_{k=1}^{d} E\{\|\mathbf{w}_k s_k\|^2\} = \sum_{k=1}^{d} \mathbf{w}_k{}^H \mathbf{w}_k \quad (2.6)$$

and the optimal downlink beamforming problem becomes:

$$\begin{aligned} \min \quad & \sum_{k=1}^{d} \mathbf{w}_k{}^H \mathbf{w}_k \\ \text{subject to} \quad & \frac{\mathbf{w}_i{}^H \mathbf{R}_i \mathbf{w}_i}{\sum_{k \neq i} \mathbf{w}_k{}^H \mathbf{R}_i \mathbf{w}_k + \sigma_i{}^2} \geq \gamma_i, \text{ for } 1 \leq i \leq d, \end{aligned} \quad (2.7)$$

or,

$$\begin{aligned} \min \quad & \sum_{k=1}^{d} \mathbf{w}_k{}^H \mathbf{w}_k \\ \text{s.t.} \quad & \mathbf{w}_i{}^H \mathbf{R}_i \mathbf{w}_i - \gamma_i \sum_{k \neq i} \mathbf{w}_k{}^H \mathbf{R}_i \mathbf{w}_k \geq \gamma_i \sigma_i{}^2, \text{ for } 1 \leq i \leq d, \end{aligned} \quad (2.8)$$

where γ_i is the lower threshold on the received SINR at user i. Due to non-convex quadratic constraints, the optimization problem in (2.8) is not convex. Problems of this form are generally NP (nondeterministic polynomial-time)hard which means that they cannot be solved in a reasonable time [7]. However, by exploiting the specific structure of (2.8), two different algorithms have been proposed to solve (2.8) efficiently [1].

The first algorithm is based on the duality between the uplink and the downlink. Problem (2.8) can be rewritten as a virtual downlink problem in which the receive beamformers at the base station and the transmit powers at the users are to be found. Therefore, a power control loop can be used to find the optimal downlink beamformers [30], [1]. However, distributed implementation of this algorithm is not possible for the downlink problem as for the uplink problem.

The other algorithm is based on semidefinite relaxation. If we let $\mathbf{W}_i = \mathbf{w}_i{\mathbf{w}_i}^H$, the optimization problem (2.8) can be written as:

$$\begin{aligned} \min \quad & \sum_{k=1}^{d} \operatorname{tr}\left(\mathbf{W}_k\right) \\ \text{s.t.} \quad & \operatorname{tr}\left(\mathbf{R}_i\mathbf{W}_i\right) - \gamma_i \sum_{k\neq i} \operatorname{tr}\left(\mathbf{R}_i\mathbf{W}_k\right) \geq \gamma_i {\sigma_i}^2 \\ & \mathbf{W}_i = {\mathbf{W}_i}^H \\ & \mathbf{W}_i \succeq \mathbf{0}, \text{and } \operatorname{rank}\left(\mathbf{W}_i\right) = 1, \text{ for } 1 \leq i \leq d, \end{aligned} \tag{2.9}$$

where tr(.) denotes the trace of a matrix. The last two constraints in (2.9) guarantee that the matrices $\mathbf{W}_i$ should be Hermitian and positive semidefinite. If we relax the rank-one constraint, the problem (2.9) becomes a semi-definite programming (SDP) problem which is a convex problem and can be solved efficiently using interior point methods [3, 36]. The optimal solutions to the relaxed problem ${\mathbf{W}_i}^{opt}$ are not generally rank-one and the relaxed problem only provides a lower bound on the original problem. However, the relaxed form of the downlink beamforming problem in (2.9) always has a rank-one solution and each optimal beamformer $\mathbf{w}_i$ is the eigenvector of the optimal matrix $\mathbf{W}_i$ with nonzero eigenvalue [1]. We exploit the semi-definite relaxation in the next two chapters to solve distributed beamforming problems in the context of relay networks.

2.4 Cooperative Schemes in Wireless Networks

As mentioned earlier, multiple antennas can be used either to improve the link reliability or to provide a higher throughput, thereby resulting in a higher spectral efficiency. Moreover, in multiple-antenna communications, transmit (downlink) and receive (uplink) beamforming are used to increase the data rate of the wireless channel through increasing the signaling range and mitigating the inter-user interference [8]. However, implementing multiple transmit antennas in mobile terminals is not always feasible due to size and complexity limitations. Furthermore, collocated antennas can rarely provide independently faded copies of the transmitted signal, unless in highly scattering environments. One approach to tackle these practical restrictions is to exploit user cooperative diversity schemes [12, 33, 34].

In user cooperative schemes, users share their communication resources, such as bandwidth and transmit power, to assist each other in data transmission. In fact, each user acts as a relay for other users during those time slots when that particular user is not transmitting its own information. Several cooperative schemes have been proposed in the literature [19, 27].

In the amplify-and-forward method, each relay receives a noisy faded version of the signal transmitted by its partner in the network. The relay then amplifies and forwards its received signal toward the final destination [20]. However, in the decode-and-forward method, each relay tries to decode the message transmitted by its partner. This decoded message is then re-encoded and retransmitted in the next step [20]. In the coded cooperation, each relay provides the final destination with incremental redundancy by transmitting a portion of its partner's codeword via a different path [14].

The amplify-and-forward (AF) approach is of particular interest due to its simplicity. By exploiting the amplify-and-forward approach, distributed space time coding strategies have also been proposed in the literature to achieve full spatial diversity in the context of relay networks [15, 17].

A two-step amplify-and-forward scheme is developed in [16, 18] where the instantaneous channel state information is assumed to be known at the relays and at the receiver. Each relay node in the network is assumed to a have a predefined power constraint. In this technique, the relay nodes try to maximize the receiver signal-to-noise ratio (SNR), not only by adjusting the phase of the received signal, but also by adjusting their transmit powers according to the channels' strength. It is shown that in order to achieve the maximum SNR at the receiver, some of the relay nodes may not use their maximum allowable power. Surprisingly, the relay transmit powers depend only on their own channel strengths except for a scalar which is broadcasted by the receiver. Another interesting aspect of the beamforming algorithm proposed in [16, 18] is that it enjoys a computational complexity that is linear in terms of the number of relaying nodes.

Two different distributed beamforming algorithms based on the second order statistics of all channel coefficients are presented in [11]. In the first approach, the beamformer design is based on the minimization of the total transmit power subject to the quality of service constraint at the destination node. This problem is shown to have a closed form solution. The second approach in [11] aims to maximize the receiver SNR subject to total power or individual relay power constraints. While closed-form solutions exist only in the case of total power constraint, the maximization problem can be written as a semi-definite programming problem in the case of individual power constraints. This convex feasibility problem is then solved using interior point methods.

In several previously published results related to cooperative communications, the relay nodes cooperate to establish a connection between a single source and a single destination, i.e., one pair of source-destination exists in the network [10, 12, 16, 33, 34, 40]. In the next two chapters, we develop two distributed beamforming schemes to establish pair wise communication links among multiple source-destination pairs through a relay network.

Chapter 3

A Multiple Peer-to-Peer Communication Scheme in Wireless Relay Networks

3.1 Introduction

In this chapter, we consider a wireless network consisting of d source-destination pairs and R relaying nodes. Each pair wishes to communicate through the collaborative relay network. The role of relays is to mitigate the cross-link interference and establish wireless connections between sources and their respective destinations. If the sources, destinations, and relays are distributed in the space, this multiplexing scheme allows multiple source-destination pairs to efficiently share the communication resources [5], [6].

Our cooperative scheme consists of two steps. In the first step, all sources transmit their data to the relays at the same time. In the second step, each relay transmits an amplified and phase-adjusted version of its received signal. Assuming that the second order statistics (i.e., correlation matrices) of all communication channels are available, we calculate the complex gains of the relays such that the total power dissipated by the relays is minimized, and at the same time, the signal-to-interference-plus-noise ratios (SINRs) at all destinations are above pre-defined thresholds.

We use a semi-definite relaxation approach to turn our power minimization problem into a semi-definite programming (SDP) optimization problem. These problems can be solved efficiently using interior point methods.

We present our data model and the power minimization problem in Sections 3.2 and 3.3, respectively. In Section 3.4, we impose additional per relay constraints to restrict the amount of power consumed by each relay. Simulation results are presented in Section 3.5.

3.2 Data Model

We consider a network with d source-destination pairs and R relays, as shown in Fig. 3.1[1]. In order to communicate to its respective destination, each source needs to transmit its data to the relay network. This assumption corresponds to the poor channel quality between the source and destination in each pair. This relay network is then responsible for delivering the data to the respective destinations. Each relay transmits an amplified and phase-steered version of its received signal that is obtained through multiplication of the relay's received signal by a complex weight.

Let f_{rp} denote the channel coefficient from the pth source to the rth relay and g_{rp} denote the channel coefficient from the rth relay to the pth destination. The rth

[1] We can also consider a network with a single common source and multiple destinations.

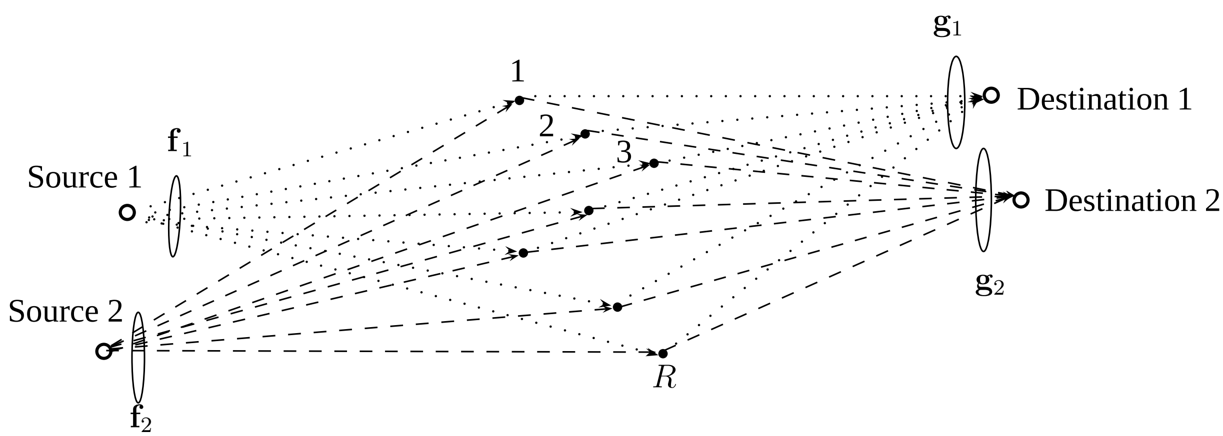

Figure 3.1: A network of R relays and 2 source-destination pairs.

relay received signal x_r is written as

$$x_r = \sum_{p=1}^{d} f_{rp} s_p + \nu_r \tag{3.1}$$

where s_p is the information symbol transmitted by the pth source and ν_r is the additive zero-mean noise at the rth relay node. We use the following assumptions throughout this chapter:

A1 The relay noise is spatially white, i.e., $E\{\nu_r \nu_{r'}^*\} = \sigma_\nu^2 \delta_{rr'}$, where σ_ν^2 is the relay noise power.

A2 The pth source uses its maximum power P_p, i.e., $E\{|s_p|^2\} = P_p$ for $p = 1, 2, \ldots, d$.

A3 The information symbols $\{s_p\}_{p=1}^d$ transmitted by different sources are uncorrelated, i.e., $E\{s_p s_q^*\} = P_p \delta_{pq}$.

A4 The information symbols $\{s_p\}_{p=1}^{d}$ and the rth relay noise ν_r are statistically independent.

Using vector notations, we can rewrite (3.1) as

$$\mathbf{x} = \sum_{p=1}^{d} \mathbf{f}_p s_p + \boldsymbol{\nu} \tag{3.2}$$

where the following definitions are used:

$$\mathbf{x} \triangleq [x_1 \;\; x_2 \;\; \ldots \;\; x_R]^T$$

$$\boldsymbol{\nu} \triangleq [\nu_1 \;\; \nu_2 \;\; \ldots \;\; \nu_R]^T$$

$$\mathbf{f}_p \triangleq [f_{1p} \;\; f_{2p} \;\; \ldots \;\; f_{Rp}]^T.$$

The rth relay multiplies its received signal by a complex weight coefficient w_r^*. As a result, the vector of the signals transmitted by all relays is given by

$$\mathbf{t} = \mathbf{W}^H \mathbf{x}, \tag{3.3}$$

where $\mathbf{W} \triangleq \text{diag}(w_1, w_2, \ldots, w_R)$ and $\mathbf{t}$ is an $R \times 1$ vector whose rth entry is the signal transmitted by the rth relay.

Let us denote the vector of the channel coefficients from the relays to the kth destination as $\mathbf{g}_k = [g_{1k} \;\; g_{2k} \;\; \ldots \;\; g_{Rk}]^T$. The kth destination received signal y_k is expressed as

$$\begin{aligned} y_k &= \mathbf{g}_k^T \mathbf{t} + n_k \\ &= \mathbf{g}_k^T \mathbf{W}^H \sum_{p=1}^{d} \mathbf{f}_p s_p + \mathbf{g}_k^T \mathbf{W}^H \boldsymbol{\nu} + n_k \\ &= \underbrace{\mathbf{g}_k^T \mathbf{W}^H \mathbf{f}_k s_k}_{\text{desired signal component}} + \underbrace{\mathbf{g}_k^T \mathbf{W}^H \sum_{p=1, p\neq k}^{d} \mathbf{f}_p s_p}_{\text{interference component}} + \underbrace{\mathbf{g}_k^T \mathbf{W}^H \boldsymbol{\nu} + n_k}_{\text{noise component}} \end{aligned} \tag{3.4}$$

where n_k is the zero-mean noise at the kth destination with a variance of σ_n^2. We further assume that:

A5 The channel coefficients $\{\mathbf{g}_k\}_{k=1}^d$, $\{\mathbf{f}_p\}_{p=1}^d$, the source signals $\{s_p\}_{p=1}^d$, the relay noise $\boldsymbol{\nu}$, and the destination noises $\{n_k\}_{k=1}^d$ are jointly independent.

3.3 Power Minimization

Our goal is to find the optimal beamforming weights $\{w_r{}^*\}_{r=1}^R$ such that the total transmit power dissipated by the relay network is minimized while the destinations' quality of services (QoSs) are kept above pre-defined thresholds. We use the SINR as a measure of QoS[2]. This measure is a standard criterion in the design of uplink and downlink beamformers. We aim to solve the following optimization problem[3]:

$$\begin{aligned} \min_{\mathbf{w}} \quad & P_T \\ \text{subject to} \quad & \text{SINR}_k \geq \gamma_k \ , \quad \text{for} \quad k = 1, 2, \ldots, d \end{aligned} \tag{3.5}$$

where P_T is the total relay transmit power and SINR_k is the SINR at the kth destination and is defined as

$$\text{SINR}_k = \frac{P_{\text{s}}^k}{P_{\text{i}}^k + P_{\text{n}}^k} \,. \tag{3.6}$$

Here, P_{s}^k, P_{i}^k, and P_{n}^k represent the desired signal component power, the interference power, and the noise power at the kth destination, respectively.

[2]In addition to the average SINR, the quality of each communication link in terms of bit error rate (BER) also depends on the statistical characteristics of the channel gains as well as the modulation type. However, these issues are too complicated to be considered in this thesis.

[3]Max min SINR is another commonly used formulation in the beamforming literature.

We now derive the expressions for the total transmit power P_T and SINR_k. Using (3.3), the total transmit power is given by

$$\begin{aligned} P_T &= E\left\{\mathbf{t}^H\mathbf{t}\right\} \\ &= E\left\{\mathbf{x}^H\mathbf{W}\mathbf{W}^H\mathbf{x}\right\} \\ &= \mathrm{tr}\left\{\mathbf{W}^H E\left\{\mathbf{x}\mathbf{x}^H\right\}\mathbf{W}\right\} \end{aligned} \tag{3.7}$$

Denoting the correlation matrix of the relay received signals by $\mathbf{R}_x \triangleq E\{\mathbf{x}\mathbf{x}^H\}$, we rewrite the total transmit power as

$$P_T = \mathrm{tr}\left\{\mathbf{W}^H\mathbf{R}_x\mathbf{W}\right\} = \sum_{r=1}^{R} |w_r|^2[\mathbf{R}_x]_{r,r} = \mathbf{w}^H\mathbf{D}\mathbf{w} \tag{3.8}$$

where $\mathbf{w} \triangleq \mathrm{diag}(\mathbf{W})$ and $\mathbf{D} \triangleq \mathrm{diag}([\mathbf{R}_x]_{1,1}, [\mathbf{R}_x]_{2,2}, \cdots, [\mathbf{R}_x]_{R,R})$. Note that using (3.2) as well as Assumptions A1-A4, the matrix $\mathbf{R}_x$ is expressed as

$$\begin{aligned} \mathbf{R}_x &= \sum_{p,q=1}^{d} E\left\{\mathbf{f}_p\mathbf{f}_q^H\right\} E\{s_p s_q^*\} + \sigma_\nu^2\mathbf{I} \\ &= \sum_{p=1}^{d} P_p E\left\{\mathbf{f}_p\mathbf{f}_p^H\right\} + \sigma_\nu^2\mathbf{I} \\ &= \sum_{p=1}^{d} P_p \mathbf{R}_f^p + \sigma_\nu^2\mathbf{I} \end{aligned} \tag{3.9}$$

where

$$\mathbf{R}_f^p \triangleq E\left\{\mathbf{f}_p\mathbf{f}_p^H\right\} \tag{3.10}$$

Note that the transmitted power P_T depends not only on the variances of the source-relay channel coefficients but also on the relay noise powers.

We now derive expressions for the desired signal component power P_{s}^k, the interference power P_{i}^k, and the noise power P_{n}^k in terms of $\{{w_r}^*\}_{r=1}^R$. Using (3.4) and

Assumption A5, we rewrite the noise power at the kth destination as

$$\begin{aligned} P_{\mathrm{n}}^k &= E\left\{\boldsymbol{\nu}^H \mathbf{W} \mathbf{g}_k^* \mathbf{g}_k^T \mathbf{W}^H \boldsymbol{\nu}\right\} + \sigma_n^2 \\ &= \mathrm{tr}\left\{\mathbf{W}^H E\left\{\boldsymbol{\nu}\boldsymbol{\nu}^H\right\} \mathbf{W} E\left\{\mathbf{g}_k^* \mathbf{g}_k^T\right\}\right\} + \sigma_n^2 \\ &= \sigma_\nu^2 \mathrm{tr}\left\{\mathbf{W}^H \mathbf{R}_g^k \mathbf{W}\right\} + \sigma_n^2 \end{aligned}$$

where

$$\mathbf{R}_g^k = E\left\{\mathbf{g}_k \mathbf{g}_k^H\right\} \tag{3.11}$$

As a result, the noise power P_{n}^k is given by

$$\begin{aligned} P_{\mathrm{n}}^k &= \sigma_\nu^2 \sum_{r=1}^{R} |w_r|^2 [\mathbf{R}_g^k]_{rr} + \sigma_n^2 \\ &= \mathbf{w}^H \mathbf{D}_k \mathbf{w} + \sigma_n^2 \end{aligned} \tag{3.12}$$

where $\mathbf{D}_k \triangleq \sigma_\nu^2 \ \mathrm{diag}([\mathbf{R}_g^k]_{11}, [\mathbf{R}_g^k]_{22}, \cdots, [\mathbf{R}_g^k]_{RR})$.

The kth desired signal power can be written as

$$\begin{aligned} P_{\mathrm{s}}^k &= E\left\{\mathbf{g}_k^T \mathbf{W}^H \mathbf{f}_k \mathbf{f}_k^H \mathbf{W} \mathbf{g}_k^*\right\} E\{|s_k|^2\} \\ &= P_k E\left\{\mathbf{w}^H \mathrm{diag}(\mathbf{g}_k) \mathbf{f}_k \mathbf{f}_k^H \mathrm{diag}(\mathbf{g}_k^*) \mathbf{w}\right\} \\ &= P_k E\left\{\mathbf{w}^H (\mathbf{g}_k \odot \mathbf{f}_k)(\mathbf{f}_k^H \odot \mathbf{g}_k^H) \mathbf{w}\right\} \\ &= P_k \mathbf{w}^H E\{\mathbf{h}_k \mathbf{h}_k^H\} \mathbf{w} \\ &= \mathbf{w}^H \mathbf{R}_h^k \mathbf{w} \end{aligned} \tag{3.13}$$

where

$$\begin{aligned} \mathbf{h}_k &\triangleq (\mathbf{g}_k \odot \mathbf{f}_k) = [f_{1k} g_{1k} \quad f_{2k}\, g_{2k} \quad \cdots \quad f_{Rk}\, g_{Rk}]^T \\ \mathbf{R}_h^k &\triangleq P_k E\{\mathbf{h}_k \mathbf{h}_k^H\}\,. \end{aligned} \tag{3.14}$$

It is worth mentioning that the vector $\mathbf{h}_k$ contains the total path gains from the kth source to its corresponding destination via different relays.

Denoting $\mathcal{D}_k = \{1, 2, \cdots, d\} - \{k\}$ and using (3.4), the interference power at the kth destination is given by

$$\begin{aligned}
P_{\mathrm{i}}^k &= E\left\{\mathbf{g}_k^T \mathbf{W}^H \left(\sum_{p,q\in\mathcal{D}_k} \mathbf{f}_p \mathbf{f}_q^H s_p s_q^*\right) \mathbf{W} \mathbf{g}_k^*\right\} \\
&= E\left\{\mathbf{w}^H \mathrm{diag}(\mathbf{g}_k) \left(\sum_{p\in\mathcal{D}_k} P_p \mathbf{f}_p \mathbf{f}_p^H\right) \mathrm{diag}(\mathbf{g}_k^*) \mathbf{w}\right\} \\
&= E\left\{\mathbf{w}^H \left(\sum_{p\in\mathcal{D}_k} P_p (\mathbf{g}_k \odot \mathbf{f}_p)(\mathbf{g}_k^H \odot \mathbf{f}_p^H)\right) \mathbf{w}\right\} \\
&= \mathbf{w}^H E\left\{\sum_{p\in\mathcal{D}_k} P_p \mathbf{h}_k^p (\mathbf{h}_k^p)^H\right\} \mathbf{w} \\
&= \mathbf{w}^H \mathbf{Q}_k \mathbf{w}
\end{aligned} \tag{3.15}$$

where $\mathbf{h}_k^p$ and $\mathbf{Q}_k$ are defined as

$$\begin{aligned}
\mathbf{h}_k^p &\triangleq \mathbf{g}_k \odot \mathbf{f}_p \\
\mathbf{Q}_k &\triangleq E\left\{\sum_{p\in\mathcal{D}_k} P_p \mathbf{h}_k^p (\mathbf{h}_k^p)^H\right\} .
\end{aligned} \tag{3.16}$$

The vector $\mathbf{h}_k^p$ contains the path coefficients from the pth source to the kth destination via R relays.

Uisng (3.8), (3.12) (3.13), and (3.15), we rewrite the optimization problem in (3.5) as

$$\begin{aligned}
\min_{\mathbf{w}} \quad & \mathbf{w}^H \mathbf{D} \mathbf{w} && (3.17) \\
\text{subject to} \quad & \frac{\mathbf{w}^H \mathbf{R}_h^k \mathbf{w}}{\mathbf{w}^H (\mathbf{Q}_k + \mathbf{D}_k) \mathbf{w} + \sigma_n^2} \geq \gamma_k, \ \text{for} \quad k = 1, 2, \ldots, d \,.
\end{aligned}$$

Since $\mathbf{w}^H(\mathbf{Q}_k + \mathbf{D}_k)\mathbf{w} + \sigma_n^2 \geq 0$, the above problem is equivalent to

$$\begin{aligned} \min_{\mathbf{w}} \quad & \mathbf{w}^H \mathbf{D} \mathbf{w} \\ \text{subject to} \quad & \mathbf{w}^H(\mathbf{R}_h^k - \gamma_k(\mathbf{Q}_k + \mathbf{D}_k))\mathbf{w} \geq \gamma_k \sigma_n^2, \ \text{for} \quad k = 1, 2, \ldots, d\,. \end{aligned} \tag{3.18}$$

The problem in (3.18) is not a convex optimization problem and may not have a solution with affordable computational complexity. We exploit a semi-definite relaxation approach to solve a relaxed version of (3.18). To do so, let us define $\mathbf{X} \triangleq \mathbf{w}\mathbf{w}^H$. Then, the optimization problem in (3.18) can be rewritten as

$$\begin{aligned} \min_{\mathbf{X}} \quad & \operatorname{tr}(\mathbf{D}\mathbf{X}) \\ \text{subject to} \quad & \operatorname{tr}(\mathbf{T}_k \mathbf{X}) \geq \gamma_k \sigma_n^2, \ \text{for} \quad k = 1, 2, \ldots, d \\ \text{and} \quad & \operatorname{rank}\,(\mathbf{X}) = 1, \quad \mathbf{X} \succeq 0 \end{aligned} \tag{3.19}$$

where $\mathbf{T}_k \triangleq \mathbf{R}_h^k - \gamma_k(\mathbf{Q}_k + \mathbf{D}_k)$. Only the rank constraint in (3.19) is not convex. Using semi-definite relaxation, we remove this non-convex constraint and aim to solve the following optimization problem

$$\begin{aligned} \min_{\mathbf{X}} \quad & \operatorname{tr}(\mathbf{D}\mathbf{X}) \\ \text{subject to} \quad & \operatorname{tr}(\mathbf{T}_k \mathbf{X}) \geq \gamma_k \sigma_n^2\,, \ \text{for} \quad k = 1, 2, \ldots, d \\ \text{and} \quad & \mathbf{X} \succeq 0\,. \end{aligned} \tag{3.20}$$

The optimization problem (3.20) is indeed convex and can be solved efficiently using interior point based software tools such as SeDuMi [36]. However, the matrix $\mathbf{X}_{\text{opt}}$, obtained by solving the optimization problem (3.20), is not necessarily of rank one, and the minimum value of the relaxed problem (3.20) only provides a lower bound on the minimum value of the original problem (3.18). Interestingly, it can be shown that the semidefinite relaxation provides the same lower bound to the original problem as the dual problem does (See Appendix A).

As it is shown in [13], we can always find a rank-one solution to the relaxed problem (3.20) as long as $d \leq 3$. Otherwise one might resort to randomization techniques to obtain a suboptimal rank-one solution. In these techniques, the optimal matrix $\mathbf{X}_{\text{opt}}$ is used to generate several suboptimal weight vectors, from which the best solution will be selected [23, 35, 38, 39]. Surprisingly, in all our numerical simulations (except for $d = 4$), the solution to the SDP problem turned out to be rank-one and hence its principal eigenvector is the optimal solution to the original problem.

Introducing slack variables $\alpha_k \geq 0$, for $k = 1, ..., d$, we can put the relaxed problem (3.20) into the standard SDP form [3]:

$$\begin{aligned}
\min_{\mathbf{X} \in \mathbb{C}^{R \times R}} \quad & \text{vec}(\mathbf{D})^T \text{vec}(\mathbf{X}) \\
\text{subject to} \quad & \text{vec}(\mathbf{T}_k)^T \text{vec}(\mathbf{X}) - \alpha_k = \gamma_k \sigma_n^2, \ \text{for} \ \ k = 1, 2, \ldots, d, \\
& \alpha_k \geq 0, \ \text{for} \ \ k = 1, 2, \ldots, d, \\
& \mathbf{X} \succeq \mathbf{0} \, .
\end{aligned} \tag{3.21}$$

The optimization problem (3.21) can now be efficiently solved using SeDuMi [36].

3.4 Individual Relay Power Constraints

In the previous section, we developed a computationally efficient technique for distributed multiplexing. This technique however does not guarantee that the relay powers are distributed fairly. As a result some of the relays may end up with significantly high transmit powers which is impractical due to power limitations of the transmit amplifiers. In this section, we impose individual constraints on relay powers to limit the maximum power consumed by each relay to a certain predefined value, say η_i. In fact, we aim to solve the following optimization problem[4]:

$$\begin{aligned} \min_{\mathbf{w}} \quad & \mathbf{w}^H \mathbf{D} \mathbf{w} \\ \text{subject to} \quad & \frac{\mathbf{w}^H \mathbf{R}_h^k \mathbf{w}}{\mathbf{w}^H (\mathbf{Q}_k + \mathbf{D}_k)\mathbf{w} + \sigma_n^2} \geq \gamma_k, \ \text{for} \quad k = 1, 2, \ldots, d \\ \text{and} \quad & [\mathbf{D}]_{rr} |w_r|^2 \leq \eta_r \ \text{for} \quad r = 1, 2, \ldots, R\,. \end{aligned} \tag{3.22}$$

Using the semi-definite relaxation approach, we turn (3.22) into the following SDP problem:

$$\begin{aligned} \min_{\mathbf{X}} \quad & \text{tr}(\mathbf{D}\mathbf{X}) \\ \text{subject to} \quad & \text{tr}(\mathbf{T}_k \mathbf{X}) \geq \gamma_k \sigma_n^2 \ , \text{for} \quad k = 1, 2, \ldots, d \\ \text{and} \quad & \mathbf{X} \succeq 0 \\ \text{and} \quad & \mathbf{X}_{rr} \leq \eta_r / [\mathbf{D}]_{rr} \ \text{for} \quad r = 1, 2, \ldots, R\,. \end{aligned} \tag{3.23}$$

In the next section, we numerically study the effects of the per-antenna power constraints on the power minimization performance.

3.5 Simulation Results

We consider two numerical examples. In both examples, the source transmit powers are the same. We require all destination SINRs to be above the same threshold value, i.e., $\gamma_k = \gamma$, for $k = 1, 2, \cdots, d$. The noise power at the relays and at the destinations are assumed to be equal to σ^2.

[4] In fact, in order to impose individual constraints on relay powers, the instantaneous values of the uplink channels must be known at each relay node.

In the first example, we assume that all channel coefficients are exactly known at a processing center where the beamforming weights for relays are to be determined. This center then broadcasts the beamforming weights to the relays. In each simulation run, the channel coefficients $\{f_{rp}\}$ and $\{g_{rp}\}$ are generated as i.i.d complex Gaussian random variables with variances σ_f^2 and σ_g^2, respectively. As the channel coefficients are known at the processing center, σ_f^2 controls the quality of uplink channels from the sources to the relays. Similarly, σ_g^2 controls the quality of the downlink channels from the relays to the destinations.

We first study the effect of the quality of the uplink and downlink channels on the performance of the proposed scheme. To do so, we consider a network consisting of 2 source-destination pairs and 20 relays. Fig. 3.2 shows the average minimum power (normalized by the noise power σ^2) consumed by all relaying nodes, versus γ, for $\sigma_g^2/\sigma^2 = 10$ dB and for different values of σ_f^2/σ^2. In this figure, the average minimum power is plotted only for those values of γ for which the beamforming problem is feasible. From this figure, we observe that an improvement in the quality of the uplink channels results in reduction of the average minimum power required to satisfy a certain QoS at the destinations.

In Fig. 3.3, we have plotted the probability of having a feasible solution for the distributed beamforming problem, for $\sigma_g^2/\sigma^2 = 10$ dB and for different values of σ_f^2/σ^2. This figure clearly shows that it becomes more likely for the beamforming problem to have a feasible solution as the quality of the uplink channels is improved. It is interesting to observe that for every 5 dB increase in σ_f^2/σ^2, the feasibility probability curves are shifted approximately by 5 dB to the right.

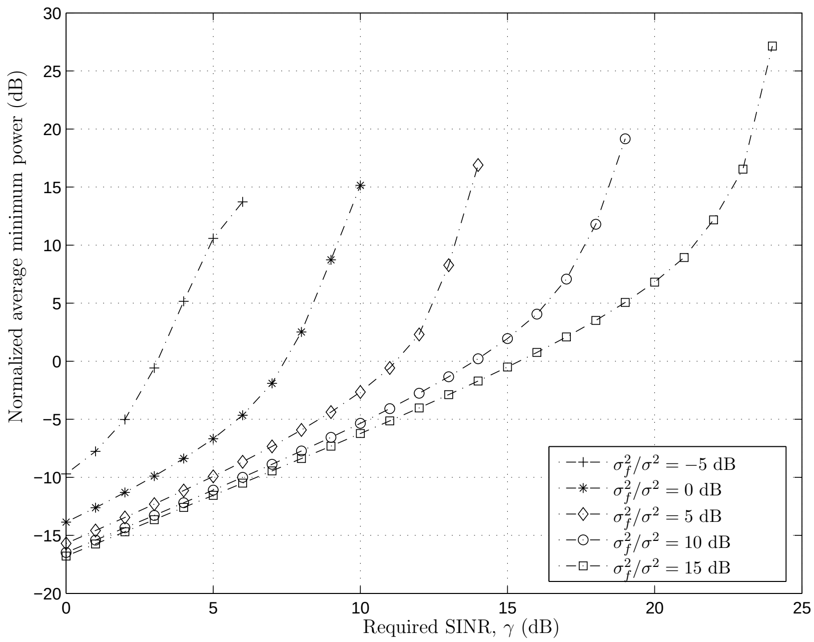

Figure 3.2: Normalized average minimum power versus SINR threshold γ, for $\sigma_g^2/\sigma^2 =$ 10 dB, and for different values of σ_f^2/σ^2, first example.

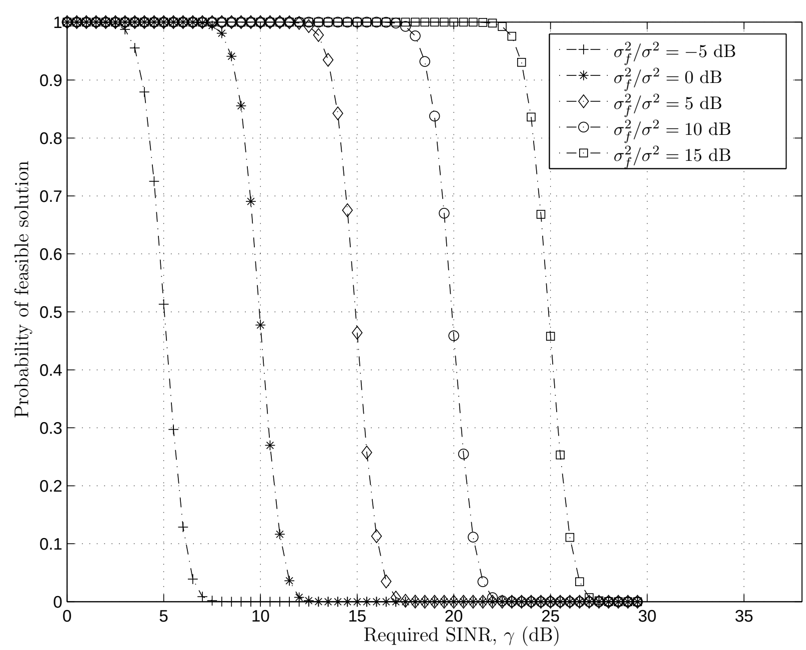

Figure 3.3: Probability of feasible solution versus SINR threshold γ, for $\sigma_g^2/\sigma^2 = 10$ dB, and for different values of σ_f^2/σ^2, first example.

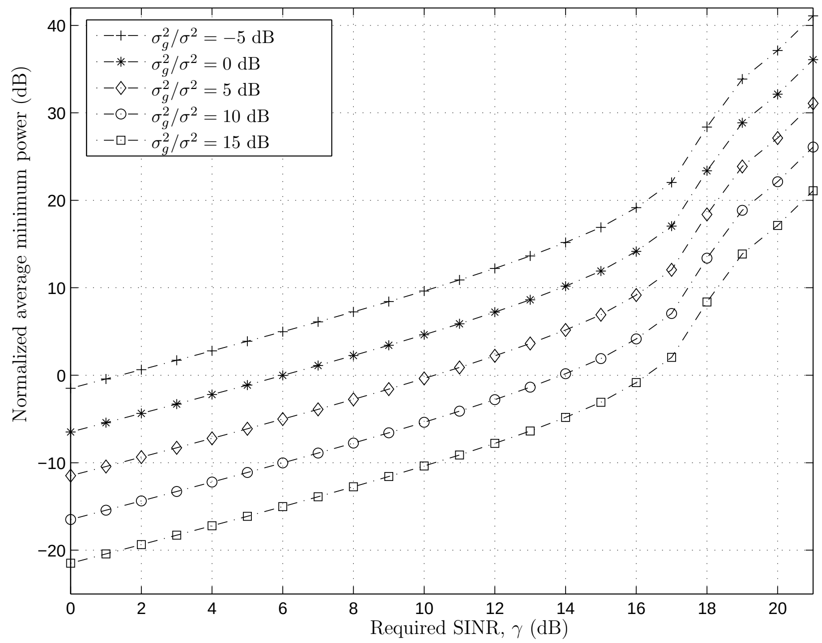

Figure 3.4: Normalized average minimum power versus SINR threshold γ, for $\sigma_f^2/\sigma^2 =$ 10 dB, and for different values of σ_g^2/σ^2, first example.

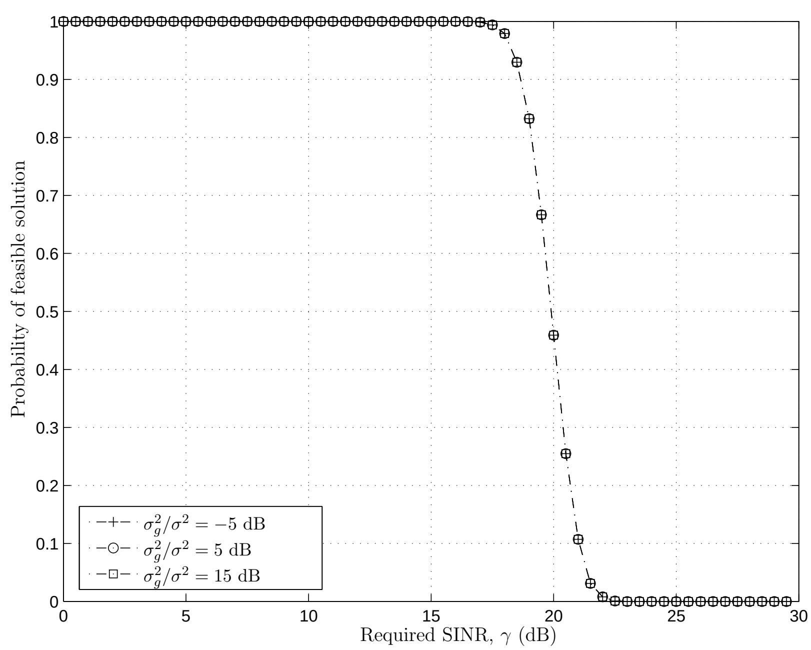

Figure 3.5: Probability of feasible solution versus SINR threshold γ, for $\sigma_f^2/\sigma^2 = 10$ dB, and for different values of σ_g^2/σ^2, first example.

Fig. 3.4 illustrates the normalized average minimum power, consumed by all relaying nodes, versus γ, for $\sigma_f^2/\sigma^2 = 10$ dB, and for different values of σ_g^2/σ^2. As can be seen from Fig. 3.4, the higher the quality of the downlink channels, the lower the average minimum power required to meet a certain QoS.

Fig. 3.5 illustrates the probability of the distributed beamforming problem having a feasible solution versus γ for $\sigma_f^2/\sigma^2 = 10$ dB, and for different values of σ_g^2/σ^2. We can see from this figure that the probability of feasible solution is invariant with respect to the quality of the downlink channels quantified by σ_g^2/σ^2. This means that the existence of a solution to the optimization problem (3.18) does not depend on the quality of the downlink channels. This observation is justified by the fact that in optimization problem (3.18), matrices $\mathbf{R}_h^k$, $\mathbf{Q}_k$, and $\mathbf{D}_k$ are all linear in σ_g^2. Thus, as σ_g^2 decreases, the optimal solution $\mathbf{w}$ can be scaled up, thereby compensating for the loss of quality of the downlink channels. This, in turn, results in scaling up the objective function of (3.18), thereby increasing the total relay transmit power. This explains why the total transmit power curves in Fig. 3.4 differ from each other by exactly 5 dB steps as in these curves, σ_g^2/σ^2 changes in 5 dB steps.

We now study the effect of the number of relays. To do so, we consider different number of relays serving 2 source-destination pairs. In this experiment, we choose $\sigma_f^2/\sigma^2 = \sigma_g^2/\sigma^2 = 10$ dB. Figs. 3.6 and 3.7 show, respectively, the average minimum power and the probability of having feasible solution versus SINR threshold γ for different number of relays. These figures show that the higher the number of relays, the lower the minimum relay power and the higher the probability of feasible solution in order to meet a certain QoS measure.

To investigate the effect of the number of users on the performance of the network,

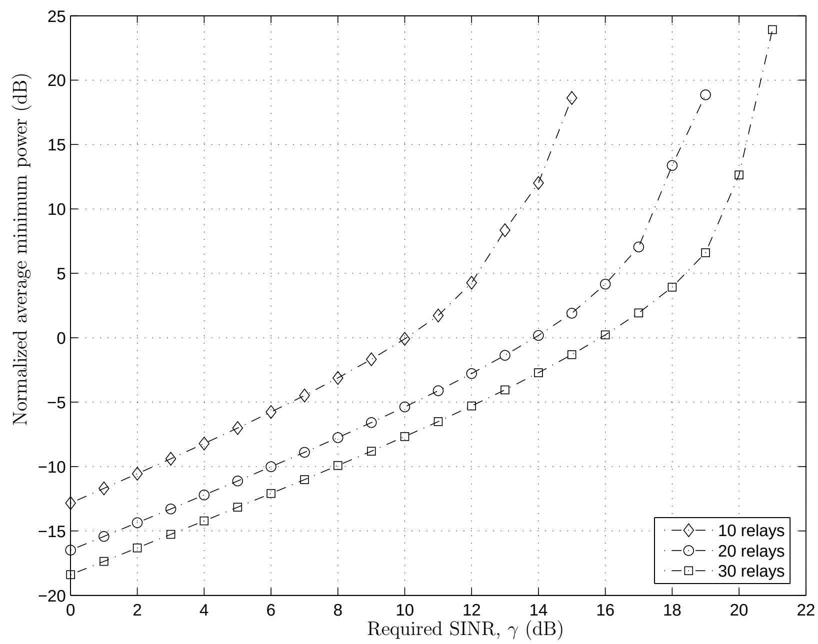

Figure 3.6: Normalized average minimum power versus SINR threshold γ for different number of relays, first example.

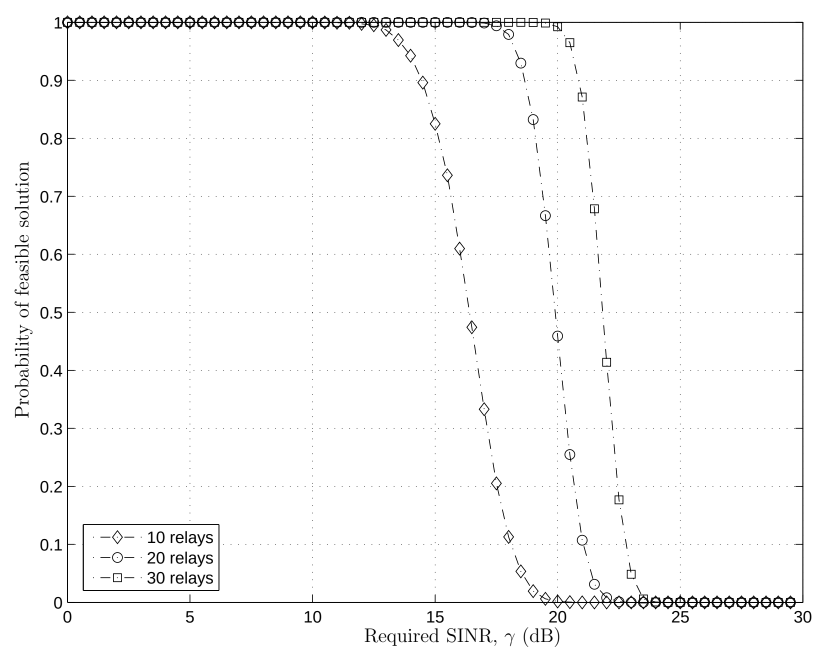

Figure 3.7: Probability of feasible solution versus SINR threshold γ for different number of relays, first example.

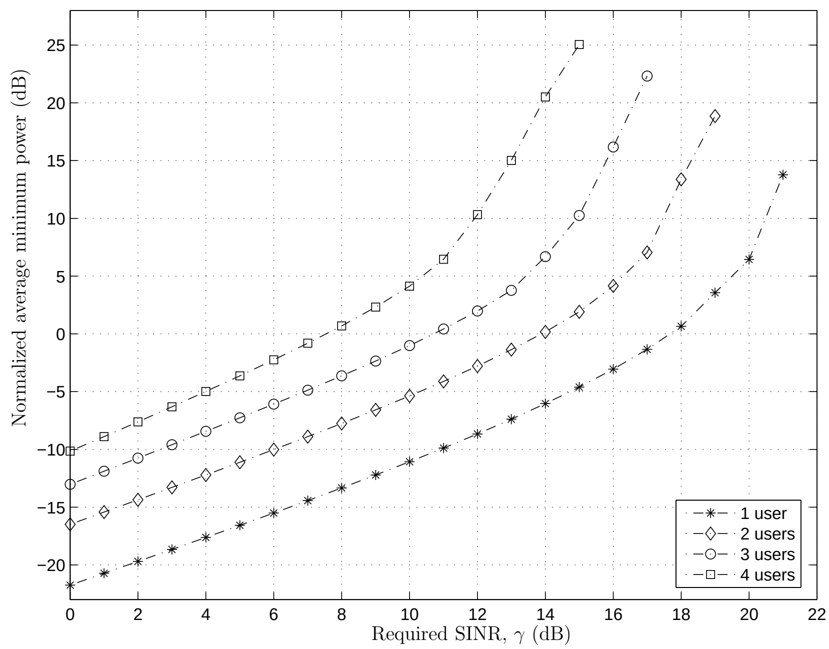

Figure 3.8: Normalized average minimum power versus SINR threshold γ for different number of source-destination pairs, first example.

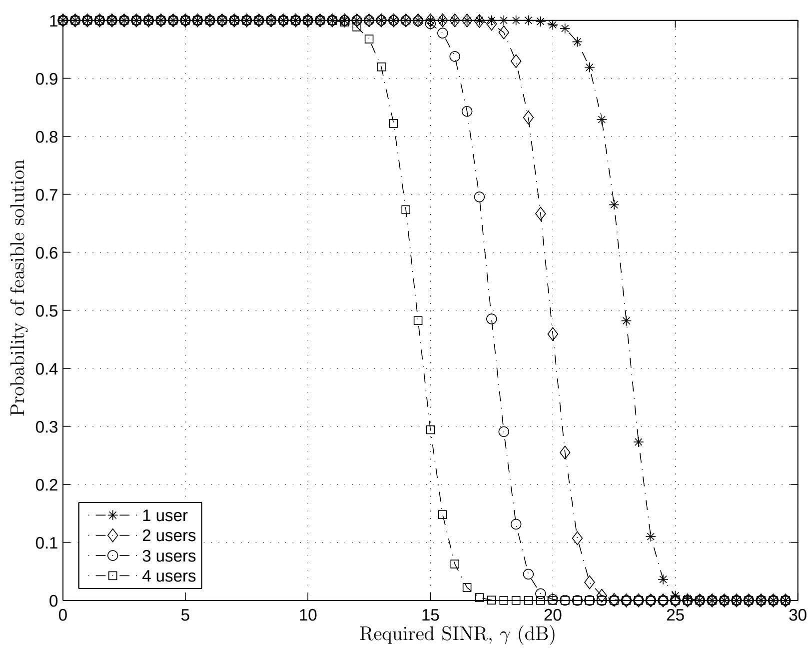

Figure 3.9: Probability of feasible solution versus SINR threshold γ, for different number of source-destination pairs, first example.

we consider four different networks which all consist of 20 relays but serve different numbers of source-destination pairs for $\sigma_f^2/\sigma^2 = \sigma_g^2/\sigma^2 = 10$ dB. Figs. 3.8 and 3.9 show, respectively, the average minimum power and the probability of feasible solution versus SINR threshold γ for different number of source-destination pairs. These figures show that for a certain QoS measure γ, as the number of source-destination pairs is increased, the minimum required relay power is increased, and the probability of feasible solution is decreased.

It is however worth mentioning that these networks are using the wireless medium for different network data rates. In other words, as the number of source-destination pairs is increased, the data rate, which the network is supposed to support, is increased. Therefore, it is a fair comparison to study the minimum relay power and the probability of feasible solution for fixed network data rates. This is exactly where our communication scheme can benefit from the spatial distribution of source-destination pairs, and therefore, can outperform (semi-)orthogonal communication schemes such as TDMA technique. To show this, we consider a network consisting of 4 source-destination pairs and 20 relays. We study three different schemes as they are described below.

- **Scheme 1:** We assume 8 different time slots, each with length T, are available for communication. In the first 4 time slots, the sources transmit their data to the relays, using a TDMA scheme. In each of the remaining 4 time slots, the relays transmit the beamformed version of the data they have received from each source to the corresponding destination. In this scheme different sources/destinations transmit/receive their data over temporally orthogonal channels, and therefore, no interference exists among different links.

- **Scheme 2:** We assume 4 different time slots, each with length $2T$, are available for communication. Each time slot is used to set up a connection between two source-destination pairs. In the first time slot, two sources transmit their data at the same time. In the second time slot, the remaining two sources transmit their data simultaneously. In the third time slot, the relays, transmit the amplitude and phase-adjusted version of the data they have received during the first time slot to the two destinations that correspond to the first two users. Finally, in the last time slot, the relays transmit the beamformed version of the data they have received in the second time slot toward the remaining two destinations.

- **Scheme 3:** We assume 2 time slots each with length $4T$ is available for communication. In the first time slot, all the four sources transmit their data toward the relays. In the second time slot, the relays transmit the beamformed version of the data they have received in the first time slot toward all 4 destinations. In this case the interference is maximal among different pairs.

We assume that in scheme 2, the source transmit powers are half of the source transmit powers in scheme 1. Taking into account that in scheme 2, the sources are transmitting for a duration which is twice the period of the source transmission in scheme 1, this will ensure that the average powers of scheme 1 and 2 are the same. Similarly, we assume that in scheme 3, the source powers are a quarter of the source powers in scheme 1. More specifically, the source transmit powers are 6 dB, 3 dB, and 0 dB above the noise level for schemes 1, 2, and 3, respectively.

In this example, we assume that source-destination pairs are required to have the same minimum data rate equal to D/d, where D is the network data rate. The SINR threshold γ is obtained through the following relationship between the rate and

SINR[5]:

$$\frac{D}{d} = \log(1+\gamma) \tag{3.24}$$

where d is the number of source-destination pairs which share the same time slot. That is d is 1, 2, and 4 for schemes 1, 2, and 3, respectively. The so-obtained γ is then used to obtain the beamforming weight vectors required in each scheme. Figs. 3.10 and 3.11 show, respectively, the normalized average minimum relay transmit power and the probability of feasible solution versus minimum network data rate for the three aforementioned schemes.

As can be seen from Fig. 3.10, as the minimum required network data rate is increased, the relay transmit power can be decreased by increasing the number of source-destination pairs which share the same time-slot[6].

It is noteworthy that in scheme 1, one beamformer is needed per source-destination pair. Therefore, the degrees of freedom (DoF) per destination is equal to the number of relays which is 20 in this example. As a result, there are 80 DoF in this TDMA network. Also, in scheme 1, only one QoS constraint is required per time slot. However, in scheme 3, only one beamformer is designed to cover all the four destinations while four QoS constraints are to be satisfied. Hence, the total DoF for scheme 3 is 20 which a quarter of that for scheme 1. As a result, one expects that scheme 1 outperforms scheme 3 over the whole range of required network data rate.

However, one can see that beyond a certain value of D, scheme 3 outperforms scheme 1. This phenomenon can be justified as it follows. For small values of data rate D, the additional degrees of freedom offered by scheme 1 results in a better

[5]Here, we have assumed that the interference is Gaussian.

[6]Only rank-one solutions have been considered in the case of $d = 4$. In almost 96 percent of our simulation runs, the solution to the SDP problem is rank-one

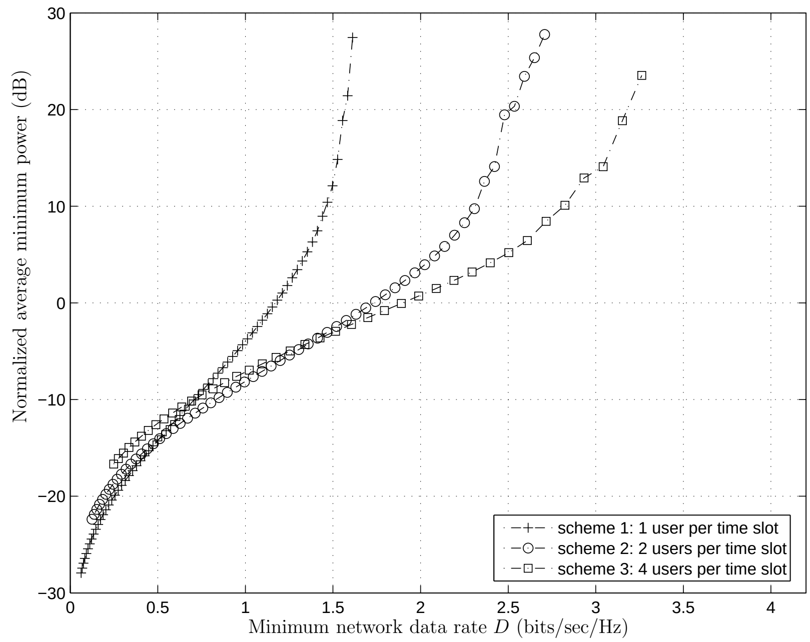

Figure 3.10: Normalized average minimum transmit power versus network data rate, for three different schemes, for $\sigma_f^2 = \sigma_g^2 = 10$ dB, first example.

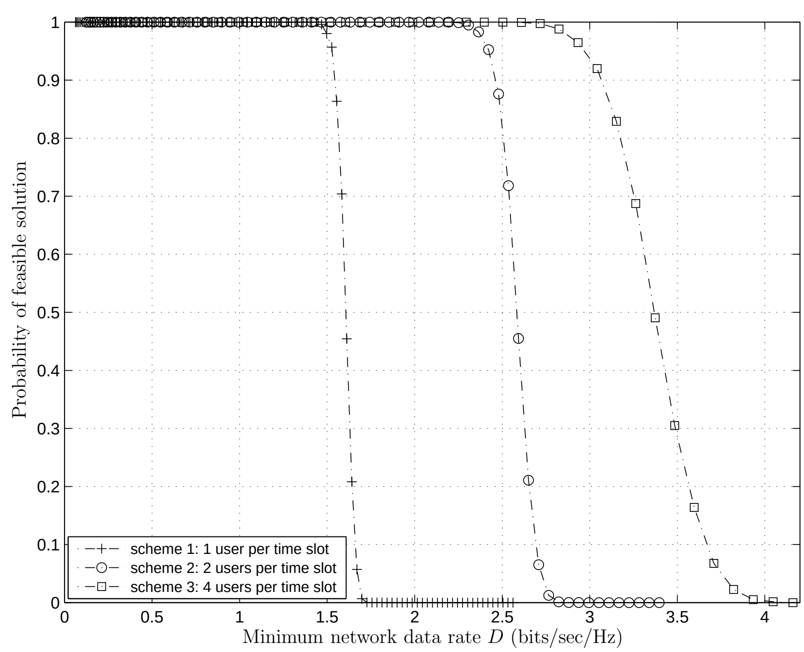

Figure 3.11: Probability of feasible solution versus network data rate, for three different schemes, for $\sigma_f^2 = \sigma_g^2 = 10$ dB, first example.

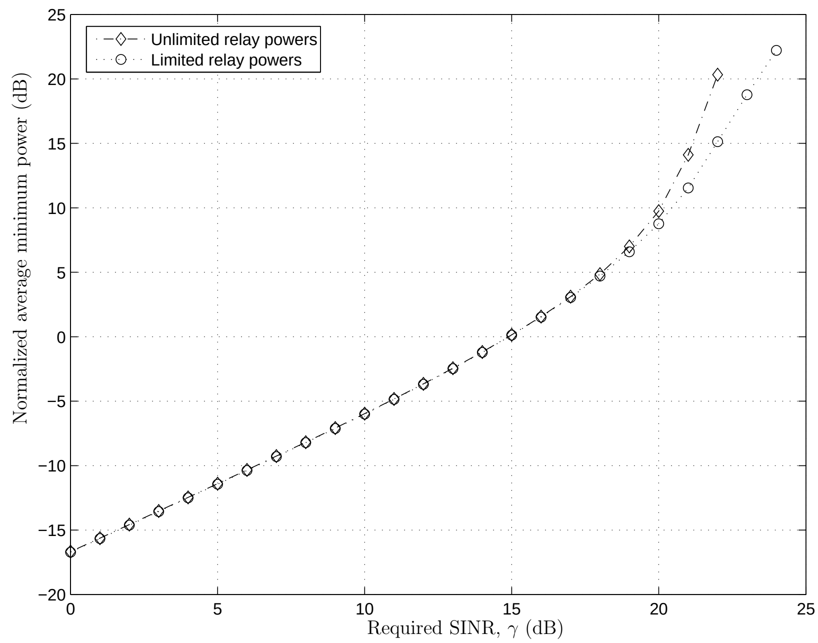

Figure 3.12: Normalized average minimum required transmit power versus SINR threshold γ, with and without per relay power constraints, second example.

performance as compared to scheme 3. As D is increased, the QoS constraints become increasingly difficult to satisfy. For scheme 1, we have that $\gamma = 2^D - 1$, while for scheme 3, $\gamma = 2^{D/4} - 1$ is chosen. Therefore, with the increase of D, the SINR threshold for scheme 1 is increased much faster than that for scheme 3. This effect overcomes the advantage of comparatively higher DoF offered by scheme 1. As a result, the average minimum power for scheme 1 increases rapidly with increasing D and the problem quickly becomes infeasible, while in scheme 3 the QoS constraints are less stringent and the network can serve the source-destination connections for a larger range of D. As can be seen from Fig. 3.11, scheme 3 remains feasible over a larger range of D as compared to scheme 1. The performance of scheme 2 is between those of schemes 1 and 3.

To consider the effect of per relay power constraint, we simulated a network with

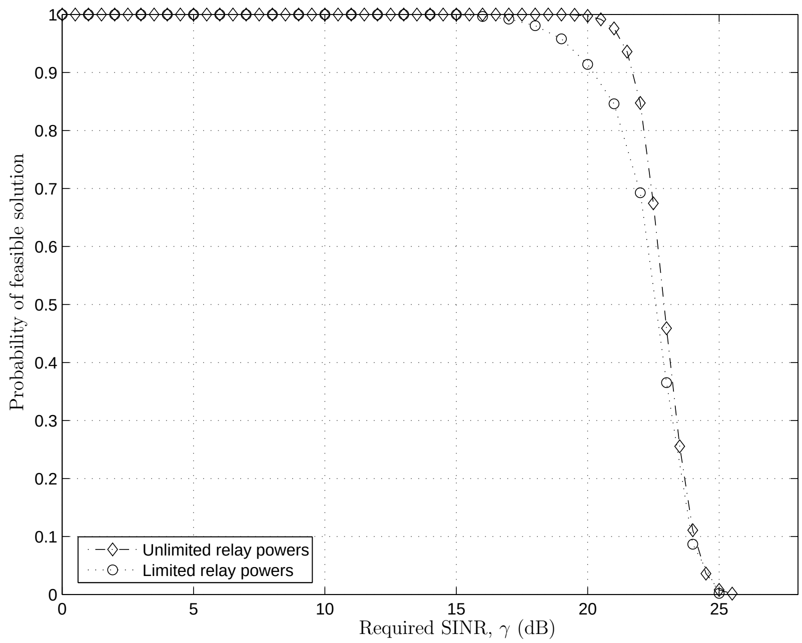

Figure 3.13: Probability of feasible solution versus SINR threshold γ, with and without per relay power constraints, second example.

2 source-destination pairs and 20 relays. The maximum allowable power for each relay is chosen to be 6 dB above the average power consumed by each relay in the unconstrained problem. Also, the source powers were assumed to be 3 dB above the noise power. We chose $\sigma_f^2 = \sigma_g^2 = 10$ dB. Figs.3.12 and 3.13 illustrate, respectively, the normalized average minimum power and the probability of feasible solution versus SNR threshold γ.

As can be seen from these figures, such per-relay power constraints do not affect the performance of our technique significantly for a wide range of γ up to 20 dB.

In our second numerical example, we assume that the second order statistics of the channel coefficients (rather than their instantaneous values) are available. We consider a network with $R = 20$ relay nodes. The channel coefficients f_{rp} and $g_{r'q}$ are assumed to be independent from each other for any p, q, r, and r'. We also

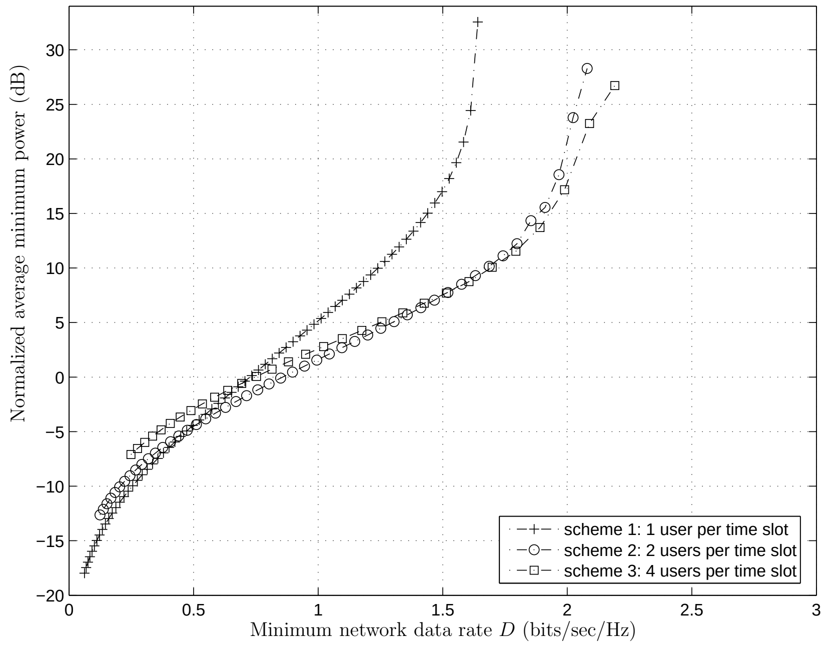

Figure 3.14: Normalized average minimum transmit power versus network data rate D, for three different schemes, for $\alpha_f = -20$ dB and $\alpha_g = -10$ dB, second example.

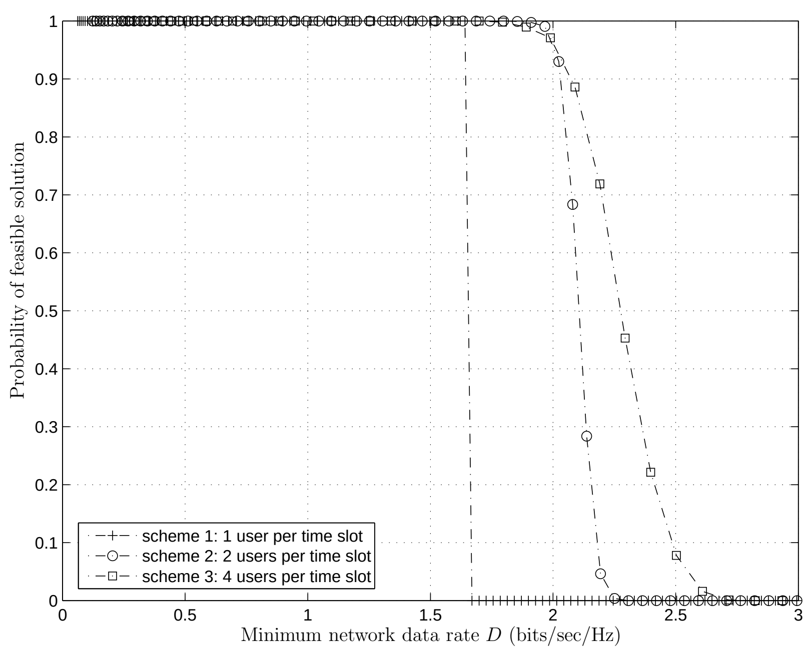

Figure 3.15: Probability of feasible solution versus network data rate, for three different schemes, $\alpha_f = -20$ dB and $\alpha_g = -10$ dB, second example.

assume that the channel coefficient f_{rp} can be written as $f_{rp} = \bar{f}_{rp} + \tilde{f}_{rp}$, where $\bar{f}_{rp}$ is a known estimate of the channel coefficient f_{rp} and $\tilde{f}_{rp}$ is a zero-mean random variable which represents the estimation error. We assume that $\tilde{f}_{rp}$ and $\tilde{f}_{r'p}$ are independent for $r \neq r'$. We choose $\bar{f}_{rp} = \frac{e^{j\theta_{rp}}}{\sqrt{1+\alpha_f}}$ and $\text{var}(\tilde{f}_{rp}) = \frac{\alpha_f}{1+\alpha_f}$, where θ_{rp} is a uniform random variable chosen from the interval $[0,\ 2\pi]$ and α_f is a parameter which determines the level of uncertainty in the channel coefficient f_{rp}. Note that as $E\{|f_{rp}|^2\} = 1$, if α_f is increased, the variance of the random component $\tilde{f}_{rp}$ is increased while the mean value, $\bar{f}_{rp}$ is decreased. This, in turn, means that the level of the uncertainty in the channel coefficient f_{rp} is increased. Similarly, we model the channel coefficient g_{rp} as $g_{rp} = \bar{g}_{rp} + \tilde{g}_{rp}$ where $\bar{g}_{rp}$ is the *known* mean of g_{rp} and $\tilde{g}_{rp}$ is a zero-mean random variable. We assume that $\tilde{g}_{rp}$ and $\tilde{g}_{r'p}$ are independent for $r \neq r'$. We choose $\bar{g}_{rp} = \frac{e^{j\phi_{rp}}}{\sqrt{1+\alpha_g}}$ and $\text{var}(\tilde{g}_{rp}) = \frac{\alpha_g}{1+\alpha_g}$, where ϕ_{rp} is a uniform random variable chosen from the interval $[0,\ 2\pi]$ and α_g is a parameter which determines the level of uncertainty in the channel coefficient g_{rp}. Based on this channel modeling, we can write the (r, r') entry of the matrices $\mathbf{R}_f^p$, $\mathbf{R}_g^k$, $\mathbf{R}_h^k$, and $\mathbf{Q}_k$, respectively, as

$$\begin{aligned}
\mathbf{R}_f^p(r,r') &= (\bar{f}_{rp}\bar{f}_{r'p}^* + \frac{\alpha_f}{1+\alpha_f}\delta_{rr'}) \\
\mathbf{R}_g^k(r,r') &= (\bar{g}_{rk}\bar{g}_{r'k}^* + \frac{\alpha_g}{1+\alpha_g}\delta_{rr'}) \\
\mathbf{R}_h^k(r,r') &= P_k\mathbf{R}_f^k(r,r')\mathbf{R}_g^k(r,r') \\
\mathbf{Q}_k(r,r') &= \sum_{p\in\mathcal{D}_k}^{d} P_p\mathbf{R}_f^p(r,r')\mathbf{R}_g^k(r,r')\,.
\end{aligned}$$

Also, throughout this numerical example, the source transmit powers are 16, 13, and 10 dB above the noise level for schemes 1, 2, and 3, respectively, and $\alpha_f = -20$ dB and $\alpha_g = -10$ dB. Figs. 3.14 and 3.15 illustrate the normalized average minimum power and the probability of feasible solution, respectively, for the three aforementioned schemes. As can be seen, in this example, scheme 3 outperforms the TDMA based scheme 1.

Chapter 4

A Distributed Downlink Beamforming Scheme in Wireless Relay Networks

4.1 Introduction

In this chapter, we consider a wireless network of d source-destination pairs and R relaying nodes. Each source communicates its data through the relay network. Our goal is to establish wireless connections between each source and its respective destination. Our cooperative scheme consists of $(d + 1)$ steps. In the first d steps, all sources transmit their data to the relay network in different time slots. So each relay obtains a noisy faded version of each source signal. In the next step, each relay multiplies each of its received signals by a complex weight, adds them all together and retransmits the result in the last $(d+1)$th time slot. We assume that the second order statistics of all communication channels are available. We aim to find the optimal

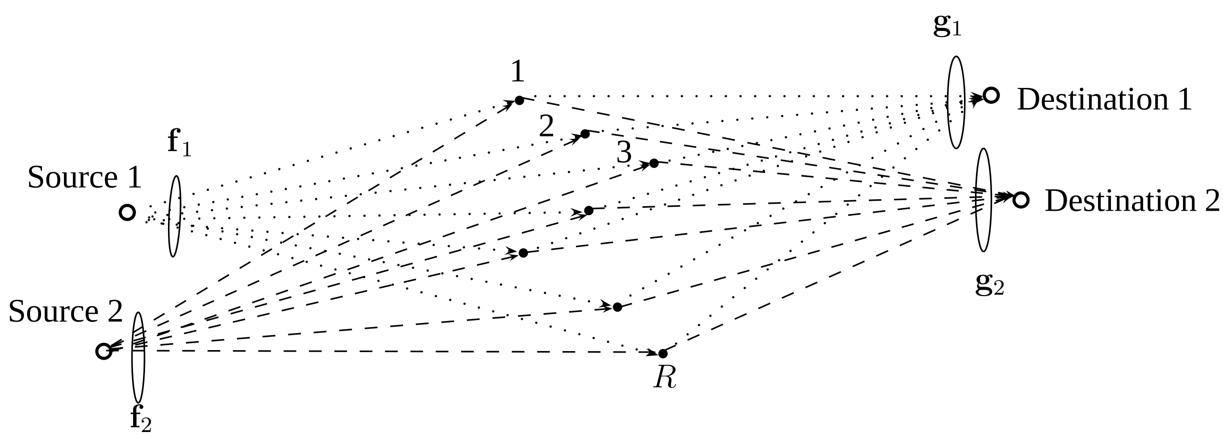

Figure 4.1: A network of R relays and 2 source-destination pairs.

complex weights at the relays to minimize the total power dissipated by the relay network, while meeting the QoS requirements at all destinations.

In contrast to those problems considered in the transmit (downlink) beamforming literature [1, 35], where the signals intended for different users are exactly known at the transmitting antennas, in our communication scheme, only noisy faded versions of source signals are available at the relays. We show that using a semi-definite relaxation approach [22], our power minimization problem can be turned into a semi-definite programming (SDP) problem, and therefore, it can be solved using interior point methods.

4.2 Problem Formulation

Consider a wireless network with d source-destination pairs and R relaying nodes, as shown in Fig. 4.1. Each source in the network wishes to communicate with its corresponding destination via the relay network. However, we assume that there is no direct link between any source and any destination. We consider a transmission scheme which consists of $d+1$ consecutive time slots. The d sources successively transmit their data during the first d time slots, so their signals do not interfere with each other at the relay nodes. The relay network is then responsible to deliver the data to the respective destinations in the $(d+1)$th slot by implementing a distributed downlink beamforming scheme.

Let f_{rp} denote the channel gain from the pth source to the rth relay and g_{rp} represent the channel gain from the rth relay to the pth destination. These channel gains are assumed to be fixed over the $d+1$ transmission time slots. Then, the rth relay received signal at time slot p denoted by x_{rp} is given by

$$x_{rp} = f_{rp}s_p + \nu_{rp}, \tag{4.1}$$

where ν_{rp} is the rth relay noise in the time slot p and s_p is the information symbol transmitted by the pth source. Again, we use the following assumptions:

A1. The relay noise is spatially and temporally white, i.e., $E\{\nu_{rp}\nu^*_{r'p'}\} = \sigma^2_\nu \delta_{rr'}\delta_{pp'}$, where σ^2_ν is the relay noise power.

A2. The power of the pth source is P_p, i.e., $E\{|s_p|^2\} = P_p$.

A3. The symbols transmitted by sources are uncorrelated, that is $E\{s_p s_q^*\} = P_p\delta_{pq}$.

The vector of the relay signals received at time slot p is then given by

$$\mathbf{x}_p = \mathbf{f}_p s_p + \boldsymbol{\nu}_p, \tag{4.2}$$

where $\mathbf{x}_p \triangleq [x_{1p}\ x_{2p}\ \ldots\ x_{Rp}]^T$, $\boldsymbol{\nu}_p \triangleq [\nu_{1p}\ \nu_{2p}\ \ldots\ \nu_{Rp}]^T$, and $\mathbf{f}_p \triangleq [f_{1p}\ f_{2p}\ \ldots\ f_{Rp}]^T$.

The rth relay multiplies the signal received from the pth source by the complex weight w_{rp}^*, adds them up, and transmits the result to the destinations at the $(d+1)$th time slot. Therefore, the $R \times 1$ vector $\mathbf{t}$ of the signals transmitted by all relays is expressed as

$$\mathbf{t} = \sum_{p=1}^{d} \mathbf{W}_p^H \left(\mathbf{f}_p s_p + \boldsymbol{\nu}_p \right), \tag{4.3}$$

where $\mathbf{W}_p \triangleq \text{diag}(\mathbf{w}_p)$, and $\mathbf{w}_p \triangleq [w_{1p}\ w_{2p}\ \ldots\ w_{Rp}]^T$ is the pth beamforming weight vector. Let us denote the vector of the channel gains from the relays to the kth destination as $\mathbf{g}_k \triangleq [g_{1k}\ g_{2k}\ \ldots\ g_{Rk}]^T$.

The signal received at the kth destination is given by

$$\begin{aligned} y_k &= {\mathbf{g}_k}^T \mathbf{t} + n_k \\ &= \underbrace{{\mathbf{g}_k}^T\, {\mathbf{W}_k}^H\, \mathbf{f}_k\, s_k}_{\text{desired signal}} + \underbrace{{\mathbf{g}_k}^T \sum_{p \in \mathcal{D}_K} {\mathbf{W}_p}^H \mathbf{f}_p s_p}_{\text{interference}} \\ &+ \underbrace{{\mathbf{g}_k}^T \sum_{p=1}^{d} {\mathbf{W}_p}^H \boldsymbol{\nu}_p + n_k}_{\text{noise}}, \end{aligned} \tag{4.4}$$

where n_k is the zero-mean additive noise at the kth destination with a variance of σ_n^2 and $\mathcal{D}_k \triangleq \{1, 2, \ldots, d\} - \{k\}$. We make another assumption:

A4 The source signals $\{s_p\}_{p=1}^d$, the relay noises $\{\boldsymbol{\nu}_p\}_{p=1}^d$, the destination noises $\{n_p\}_{p=1}^d$, and all channel gain vectors $\{\mathbf{g}_p\}_{p=1}^d$ and $\{\mathbf{f}_p\}_{p=1}^d$ are statistically independent.

4.3 Power Minimization

In this section, we aim to find the relay weight vectors $\{\mathbf{w}_p\}_{p=1}^{d}$ such that the total relay transmit power is minimized while maintaining the destinations' quality of services (QoSs) above pre-defined thresholds. Using the SINR as the measure of QoS, our goal is to solve the following optimization problem:

$$\begin{aligned} \min_{\mathcal{W}} \quad & P_T \\ \text{subject to} \quad & \text{SINR}_k \geq \gamma_k \ , \quad \text{for} \quad 1 \leq k \leq d, \end{aligned} \tag{4.5}$$

where $\mathcal{W} \triangleq \{\mathbf{w}_p\}_{p=1}^{d}$, P_T is the total relay transmit power, and SINR_k is the SINR at the kth destination which is given by

$$\text{SINR}_k = \frac{P_{\text{s}}^k}{P_{\text{i}}^k + P_{\text{n}}^k}. \tag{4.6}$$

Here, P_{s}^k, P_{i}^k, and P_{n}^k are the desired signal power, the interference power, and the noise power at the kth destination, respectively. Using (4.3) and assumptions A1-A4, the total transmit power is written as

$$\begin{aligned} P_T &= E\left\{\mathbf{t}^H\mathbf{t}\right\} \\ &= E\left\{\sum_{p,q=1}^{d} \text{tr}\left\{\mathbf{W}_p{}^H\mathbf{f}_p\mathbf{f}_q{}^H\mathbf{W}_q\right\} s_p s_q^H\right\} \\ &+ E\left\{\sum_{p,q=1}^{d} \text{tr}\left\{\mathbf{W}_p{}^H\boldsymbol{\nu}_p\boldsymbol{\nu}_q{}^H\mathbf{W}_q\right\}\right\} \\ &= \sum_{p=1}^{d} \text{tr}\left\{\mathbf{W}_p{}^H\left(P_p\mathbf{R}_f^p + \sigma_\nu^2\mathbf{I}\right)\mathbf{W}_p\right\} \\ &= \sum_{p=1}^{d} \mathbf{w}_p{}^H\mathbf{D}_p\mathbf{w}_p, \end{aligned} \tag{4.7}$$

where $\mathbf{D}_p \triangleq P_p(\mathbf{I} \odot \mathbf{R}_f^p) + \sigma_\nu^2 \mathbf{I}$, $\mathbf{R}_f^p \triangleq E\left\{\mathbf{f}_p \mathbf{f}_p^H\right\}$. The desired signal power at the kth destination is given by

$$\begin{aligned}
P_{\mathrm{s}}^k &= E\left\{{\mathbf{g}_k}^T {\mathbf{W}_k}^H \mathbf{f}_k {\mathbf{f}_k}^H \mathbf{W}_k \mathbf{g}_k^*\right\} E\{|s_k|^2\} \\
&= P_k E\left\{{\mathbf{w}_k}^H \mathbf{G}_k \mathbf{f}_k {\mathbf{f}_k}^H {\mathbf{G}_k}^H \mathbf{w}_k\right\} \\
&= P_k E\left\{{\mathbf{w}_k}^H (\mathbf{f}_k \odot \mathbf{g}_k)(\mathbf{f}_k \odot \mathbf{g}_k)^H) \mathbf{w}_k\right\} \\
&= P_k {\mathbf{w}_k}^H \left(E\{\mathbf{f}_k {\mathbf{f}_k}^H\} \odot E\{\mathbf{g}_k {\mathbf{g}_k}^H\}\right) \mathbf{w}_k \\
&= P_k \mathbf{w}_k^H \left(\mathbf{R}_f^k \odot \mathbf{R}_g^k\right) \mathbf{w}_k, \qquad (4.8)
\end{aligned}$$

where $\mathbf{G}_k = \mathrm{diag}(\mathbf{g}_k)$. It is worth mentioning that the rth entry of the vector $\mathbf{f}_k \odot \mathbf{g}_k$ is the total path gain between the kth source and its corresponding destination via the rth relay (excluding the relay gains). Using (4.4) as well as assumptions A1-A4, the interference power at the kth destination is written as

$$\begin{aligned}
P_{\mathrm{i}}^k &= E\left\{{\mathbf{g}_k}^T \left(\sum_{p,q \in \mathcal{D}_k} {\mathbf{W}_p}^H \mathbf{f}_p {\mathbf{f}_q}^H \mathbf{W}_q s_p s_q^*\right) \mathbf{g}_k^*\right\} \\
&= E\left\{\sum_{p \in \mathcal{D}_k} P_p {\mathbf{w}_p}^H \mathbf{G}_k \mathbf{f}_p {\mathbf{f}_p}^H {\mathbf{G}_k}^H \mathbf{w}_p\right\} \\
&= E\left\{\sum_{p \in \mathcal{D}_k} P_p {\mathbf{w}_p}^H (\mathbf{f}_p \odot \mathbf{g}_k)(\mathbf{f}_p^H \odot \mathbf{g}_k^H) \mathbf{w}_p\right\} \\
&= \sum_{p \in \mathcal{D}_k} P_p {\mathbf{w}_p}^H \left(E\{\mathbf{f}_p {\mathbf{f}_p}^H\} \odot E\{\mathbf{g}_k {\mathbf{g}_k}^H\}\right) \mathbf{w}_p \\
&= \sum_{p \in \mathcal{D}_k} P_p {\mathbf{w}_p}^H \left(\mathbf{R}_f^p \odot \mathbf{R}_g^k\right) \mathbf{w}_p, \qquad (4.9)
\end{aligned}$$

where $\mathbf{R}_g^k \triangleq E\left\{\mathbf{g}_k {\mathbf{g}_k}^H\right\}$. The vector $\mathbf{f}_p \odot \mathbf{g}_k$ contains the total link gains from the pth source to the kth destination that are passing through the R relays.

Similarly the noise power at the kth destination is given by

$$\begin{aligned} P_{\mathrm{n}} &= E\left\{ \mathbf{g}_k{}^T \left(\sum_{p,q=1}^{d} \mathbf{W}_p{}^H \boldsymbol{\nu}_p \boldsymbol{\nu}_q{}^H \mathbf{W}_q \right) \mathbf{g}_k{}^* + |n_k|^2 \right\} \\ &= \sigma_\nu^2 E\left\{ \sum_{p=1}^{d} \left(\mathbf{g}_k{}^T \mathbf{W}_p{}^H \mathbf{W}_p \mathbf{g}_k{}^* \right) \right\} + \sigma_n^2 \\ &= \sigma_\nu^2 \sum_{p=1}^{d} \mathbf{w}_p{}^H \left(\mathbf{I} \odot E\left\{ \mathbf{g}_k \mathbf{g}_k{}^H \right\} \right) \mathbf{w}_p + \sigma_n^2 \\ &= \sigma_\nu^2 \sum_{p=1}^{d} \mathbf{w}_p{}^H \left(\mathbf{I} \odot \mathbf{R}_g^k \right) \mathbf{w}_p + \sigma_n^2. \end{aligned} \tag{4.10}$$

Using (4.7), (4.8), (4.9), and (4.10), the optimization problem in (3.5) can be rewritten as

$$\begin{aligned} \min_{\mathcal{W}} \quad & \sum_{p=1}^{d} \mathbf{w}_p{}^H \mathbf{D}_p \mathbf{w}_p \\ \text{subject to} \quad & \sum_{p=1}^{d} \mathbf{w}_p{}^H \mathbf{Q}_{kp} \mathbf{w}_p \geq \gamma_k \sigma_n^2, \ \text{for } 1 \leq k \leq d, \end{aligned} \tag{4.11}$$

where

$$\mathbf{Q}_{kp} \triangleq \begin{cases} (P_k \mathbf{R}_f^k - \gamma_k \sigma_\nu^2 \mathbf{I}) \odot \mathbf{R}_g^k\,, & \text{if } p = k \\ -\gamma_k (P_p \mathbf{R}_f^p + \sigma_\nu^2 \mathbf{I}) \odot \mathbf{R}_g^k)\,, & \text{otherwise.} \end{cases} \tag{4.12}$$

It is worth mentioning that at the optimum, all constraints hold with equality. Suppose that a link SINR is higher than the predefined threshold, then we can scale down the corresponding weight vector and hence decrease the total relay transmit power without violating any constraints, which is a contradiction[1]. In general, the optimization problem in (4.11) is not convex [3] and may not be amenable to a computationally affordable solution. We propose to use the semi-definite relaxation to

[1] In fact, scaling down the weight vector corresponding to each user is beneficial to other source-destination pairs in the network because of less interference from that user.

approximately solve (4.11). To do so, let us define $\mathbf{X}_p \triangleq \mathbf{w}_p {\mathbf{w}_p}^H$ for $1 \leq p \leq d$. In this case, the problem in (4.11) becomes

$$\begin{aligned} \min_{\mathcal{X}} \quad & \sum_{p=1}^{d} \operatorname{tr}\left(\mathbf{X}_p \mathbf{D}_p\right) \\ \text{subject to} \quad & \sum_{p=1}^{d} \operatorname{tr}\left(\mathbf{X}_k \mathbf{Q}_{kp}\right) \geq \gamma_k \sigma_n^2 \\ \text{and} \quad & \operatorname{rank}\left(\mathbf{X}_k\right) = 1, \quad \mathbf{X}_k \succeq \mathbf{0}, \ \text{for } 1 \leq k \leq d, \end{aligned} \tag{4.13}$$

where $\mathcal{X} \triangleq \{\mathbf{X}_p\}_{p=1}^{d}$. The rank constraints in (4.13) are not convex. We relax these constraints and solve the following convex problem

$$\begin{aligned} \min_{\mathcal{X}} \quad & \sum_{p=1}^{d} \operatorname{tr}\left(\mathbf{X}_p \mathbf{D}_p\right) \\ \text{subject to} \quad & \sum_{p=1}^{d} \operatorname{tr}\left(\mathbf{X}_k \mathbf{Q}_{kp}\right) \geq \gamma_k \sigma_n^2 \\ \text{and} \quad & \mathbf{X}_k \succeq \mathbf{0}\,, \ \text{for } 1 \leq k \leq d. \end{aligned} \tag{4.14}$$

The optimization problem in (4.14) is convex and can be solved using interior point based packages such as SeDuMi [36].

When using semi-definite relaxation, the solutions obtained by solving the relaxed problem may not necessarily be of rank-one. However, as the following lemma implies, we can always find a set of rank-one solutions $\{\mathbf{X}_p\}_{p=1}^{d}$ to the relaxed problem (4.14).

Lemma 1 *If the relaxed problem (4.14) is feasible, it always has at least one optimal solution where all* $\mathbf{X}_k$ *are rank-one.*

Lemma 1 can be proved following the guidelines of [1] when applied to our problem (see Appendix B). In practice, the solution to the problem (4.14) always turns out to be rank-one except for cases with exact symmetry. In this case, we can exploit Lemma 5 in [1] to find a rank-one solution.

4.4 Simulation Results

In our simulations we consider a network with 2 source-destination pairs and $R = 10$ relay nodes. The channel coefficient vectors $\mathbf{f}_p$ and $\mathbf{g}_q$ are assumed to be independent from each other for any p and q. We also assume that the channel coefficient vector $\mathbf{f}_p$ can be written as $\mathbf{f}_p = \bar{\mathbf{f}}_p + \tilde{\mathbf{f}}_p$, where $\bar{\mathbf{f}}_p$ is a known estimate of the channel vector $\mathbf{f}_p$ and $\tilde{\mathbf{f}}_k$ is a zero-mean random vector which represents the estimation error.

We choose $\bar{f}_{rp} = \dfrac{e^{j\theta_{rp}}}{\sqrt{1+\alpha_f}}$ and $\text{var}(\tilde{f}_{rp}) = \dfrac{\alpha_f}{1+\alpha_f}$, where $\bar{f}_{rp}$, and $\tilde{f}_{rp}$ are the rth entry of $\bar{\mathbf{f}}_p$ and $\tilde{\mathbf{f}}_p$, respectively, θ_{rp} is a uniform random variable chosen from the interval $[0 \;\; 2\pi]$, and α_f is a parameter which determines the level of uncertainty in the channel coefficient f_{rp}. Note that as $E\{|f_{rp}|^2\} = 1$, if α_f is increased, the variance of the random component of f_{rp} is increased while its mean is decreased. This, in turn, means that the level of the uncertainty in the channel coefficient f_{rp} is increased.

Similarly, we model the channel coefficient $\mathbf{g}_p$ as $\mathbf{g}_p = \bar{\mathbf{g}}_p + \tilde{\mathbf{g}}_p$, where $\bar{\mathbf{g}}_p$ is the (known) mean of $\mathbf{g}_p$, and $\tilde{\mathbf{g}}_p$ is a zero-mean random vector with independent entries. We choose $\bar{g}_{rp} = \dfrac{e^{j\phi_{rp}}}{\sqrt{1+\alpha_g}}$ and $\text{var}(\tilde{g}_{rp}) = \dfrac{\alpha_g}{1+\alpha_g}$, where $\bar{g}_{rp}$ and $\tilde{g}_{rp}$ are the rth entry of $\bar{\mathbf{g}}_p$ and $\tilde{\mathbf{g}}_p$, respectively, ϕ_{rp} is a uniform random variable chosen from the interval $[0 \;\; 2\pi]$, and α_g is a parameter which determines the level of uncertainty in the channel coefficient g_{rp}. Based on this modeling, the matrices $\mathbf{R}_f^p$ and $\mathbf{R}_g^p$ are respectively given by

$$\begin{aligned} \mathbf{R}_f^p &= \bar{\mathbf{f}}_p \bar{\mathbf{f}}_p^H + \frac{\alpha_f}{1+\alpha_f}\mathbf{I} \\ \mathbf{R}_g^p &= \bar{\mathbf{g}}_p \bar{\mathbf{g}}_p^H + \frac{\alpha_g}{1+\alpha_g}\mathbf{I} \end{aligned}$$

The relay and the destination noise powers are all assumed to be equal to 0 dBW, i.e., $\sigma_\nu^2 = \sigma_n^2 = 1$. We require all destinations' SINRs to be above the same threshold

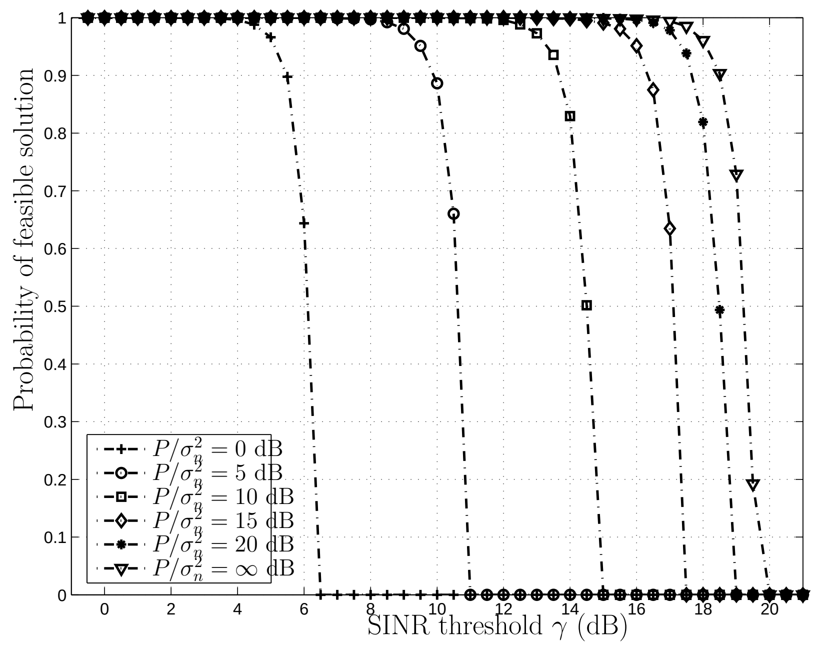

Figure 4.2: Probability of feasible solution versus SINR threshold, for different values of P/σ_n^2, for $\alpha_f = -20$ dB and $\alpha_g = -10$ dB.

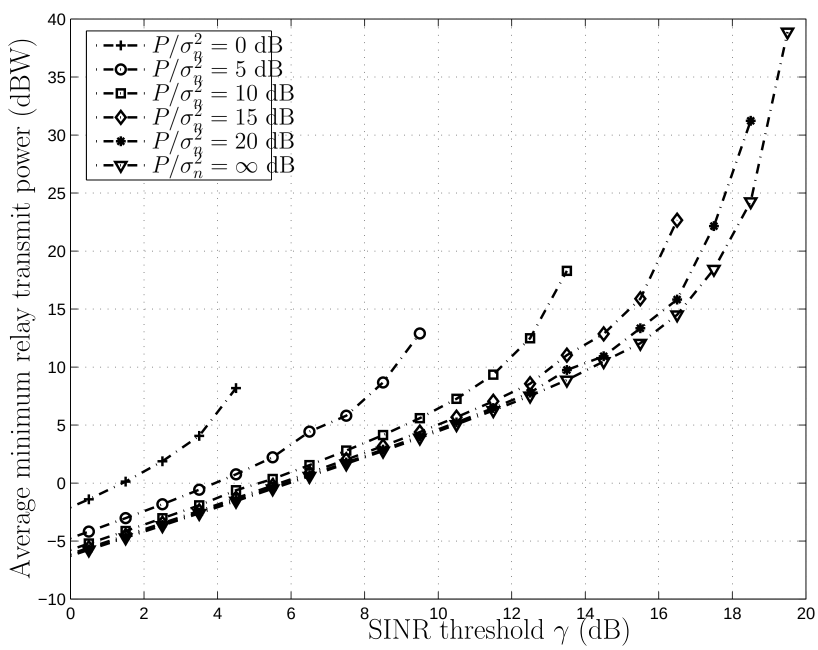

Figure 4.3: Average minimum relay transmit power versus SINR threshold, for different values of P/σ_n^2, for $\alpha_f = -20$ dB and $\alpha_g = -10$ dB.

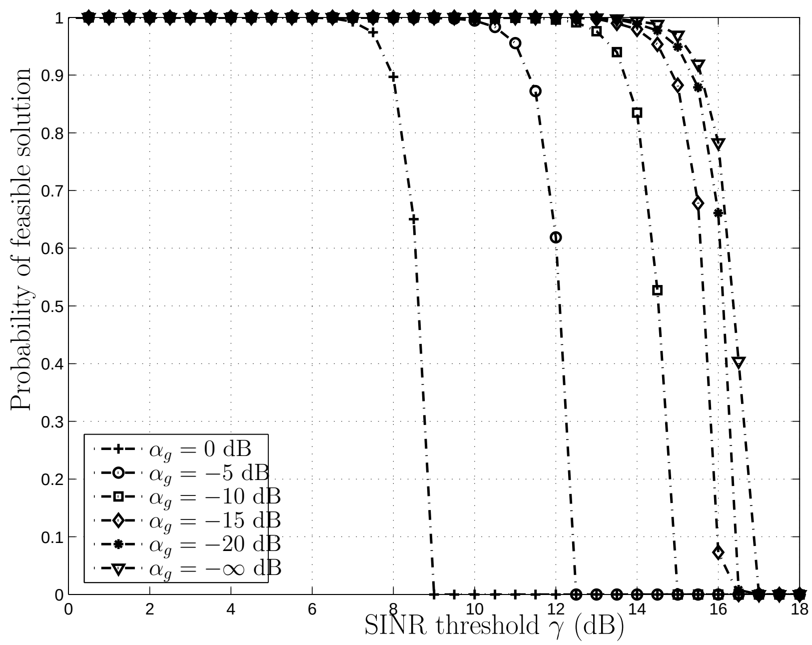

Figure 4.4: Probability of feasible solution versus SINR threshold, for different values of α_g and for $\alpha_f = -20$ dB.

value, that is, $\gamma_k = \gamma$, for $k = 1, 2, \cdots, d$. We also assume that the source transmit powers are all equal to P. Figs. 4.2 and 4.3 illustrate, respectively, the probability of feasible solution and the average minimum relay transmit power curves versus SINR threshold for different values of P/σ_n^2 and for $\alpha_f = -20$ dB and $\alpha_g = -10$ dB.

Figs. 4.4 and 4.5 illustrate the probability of feasible solution and the average minimum relay transmit power for different values of α_g, respectively. In these figures, we assume that $P/\sigma^2 = 10$ dB and $\alpha_f = -20$ dB. As one might expect, as the level of uncertainty in the channel coefficients is increased, the QoS constraints become exceedingly difficult to satisfy. The extreme case of $\alpha_g = -\infty$ dB corresponds to the case of perfectly known downlink channels.

We now compare our proposed beamforming scheme with the TDMA scheme and

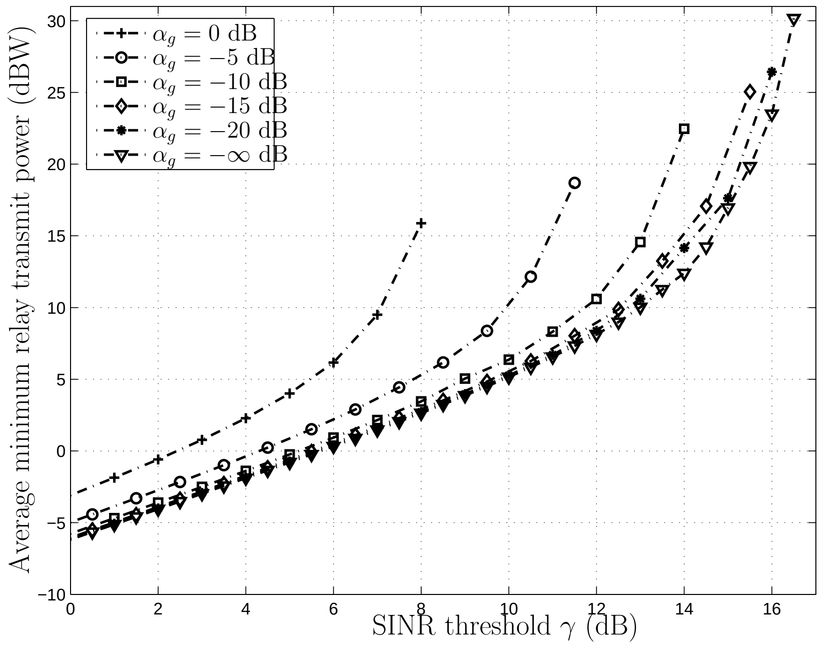

Figure 4.5: Average minimum relay transmit power versus SINR threshold, for different values of α_g and for $\alpha_f = -20$ dB.

the communication scheme proposed in Chapter 3. Note that due to the amplify-and-forward approach, different time slots have equal lengths and each source transmits in only one time slot. Again we assume that the average source transmit powers are the same. Again, if we assume that the interference is Gaussian, we can write

$$\begin{aligned} D &= \frac{d}{d+1}\log(1+\gamma) \quad \text{or} \\ \gamma &= 2^{\frac{(d+1)D}{d}} - 1. \end{aligned} \tag{4.15}$$

where D is the total network data rate, and d is the number of sources transmitting at the same time. In our simulations we consider a network with four source-destination pairs and twenty relays. We assume that the relay and the destination noises are Gaussian with the power of 0 dB and the average source powers are equal to 10 dB. We also assume that α_f and α_g are equal to −20 dB. Fig. 4.6 and Fig. 4.7 illustrate

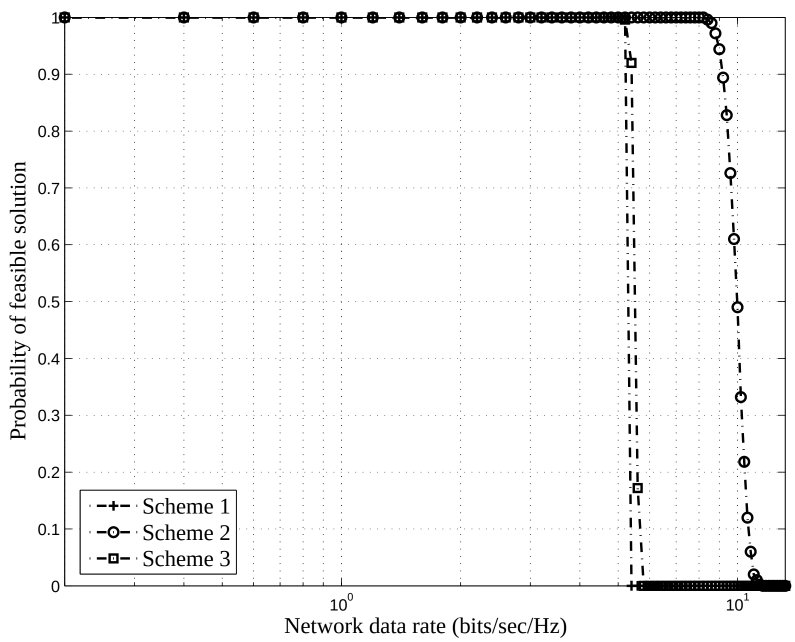

Figure 4.6: Probability of feasible solution versus network data rate.

the probability of feasible solution and the average minimum relay transmit-power versus network data rate for three aforementioned schemes. Scheme 1, Scheme 2, and Scheme 3 represent the orthogonal TDMA scheme and the proposed beamforming schemes in Chapter 3 and 4, respectively. As can be seen from these figures, the second scheme significantly outperforms the other two schemes when the required data rate is high. Although the SINR threshold increases exponentially with the data rate in all three schemes, the increase in γ is much slower for the second scheme. This results in the better performance of the second scheme at high data rates. In the third scheme, one beam is designed per destination, which results in higher degrees of freedom and better performance over the other two schemes at low to medium data rate. The first scheme is preferred when the complexity is the major concern.

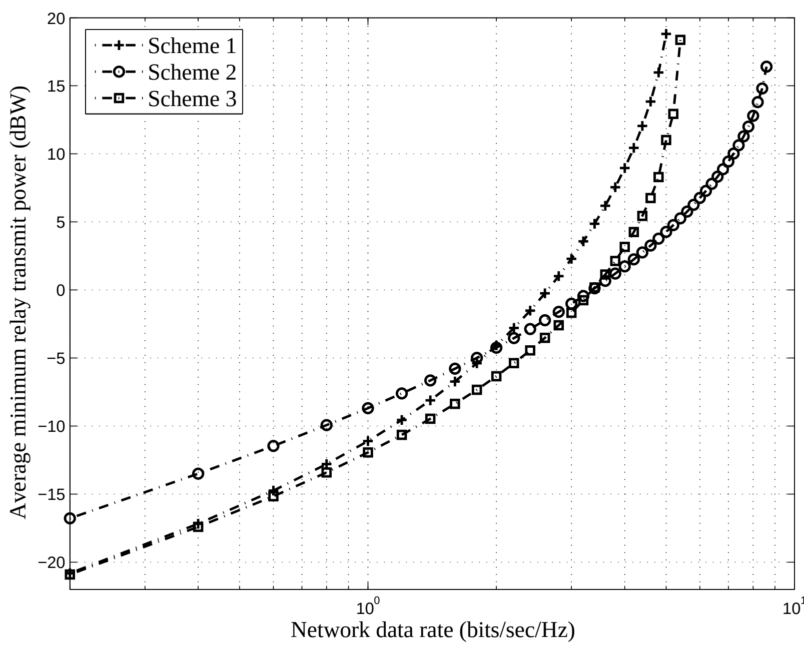

Figure 4.7: Average minimum relay transmit power versus network data rate.

Chapter 5

Conclusions and Future Work

Due to its simplicity, the amplify-and-forward approach has been widely studied in the literature to gain better insight into cooperative communications. We proposed two cooperative beamforming schemes to establish wireless connections between multiple source-destination pairs in a relay network. By exploiting the inherent spatial diversity of wireless users, our proposed schemes result in a significant improvement in the spectral efficiency of the wireless network.

In Chapter 3, we considered a network consisting of d source-destination pairs and R relaying nodes. We studied the problem of multiple peer-to-peer communication using an amplify-and-forward approach. In our scheme, first, the d sources transmit their information symbols simultaneously to the relay network. Then, each relay transmits an amplitude- and phase-adjusted version of its received signal. These adjustments are made by multiplication of the relays' received signal by a complex weight. In this scheme, one beam covers all the destinations and mitigates the cross-link interference. The optimal weight vector is in fact a receive and a transmit beamformer at the same time. This vector is obtained through the minimization of the total relay transmit power under quality of services constraints at all links. We used semi-definite relaxation to convert this optimization problem into a semi-definite programming problem and solved the resulting convex problem efficiently using interior point methods. By exploiting the spatial diversity of sources, relays, and destinations, our method allows multiple source-destination pairs to share the available resources.

In Chapter 4, we studied the problem of distributed downlink beamforming in a relay network where the relaying nodes cooperate to establish multiple peer-to-peer communication links. Having the idea of downlink beamforming in mind, our goal was to minimize the total power dissipated by the relay network subject to quality of service constraints, measured by SINR, at the destination nodes. Our communication scheme consists of two phases. In the first phase, the sources transmit their information using a TDMA scheme. In the second phase, the relay nodes multiply their received signals by complex weights, and thereby, build one beam per destination collectively. We formulated the problem of relay transmit power minimization and showed that it can be solved using a semi-definite relaxation approach.

We compared our proposed schemes with the conventional TDMA scheme in terms of probability of feasible solution and average minimum relay transmit power. We showed that the proposed schemes significantly outperform the TDMA scheme, specially, at high data rates.

However, optimal beamforming schemes result in more complexity as well as more sensitivity to channel uncertainties. Furthermore, the assumption that all channel statistics are available at a processing center may result in high system overhead. Therefore, developing suboptimal beamforming schemes that are more robust and require only local channel information is of particular interest for future research.

Another important question is how to best group the source-destination pairs into cooperating groups and how to choose the best (possibly small) set of relays to guarantee the QoS constraints, while keeping the complexity of the problem as low as possible.

Bibliography

[1] M. Bengtsson and B. Ottersten. Optimal and suboptimal transmit beamforming. In L. C. Godara, editor, *Handbook of Antennas in Wireless Communications*, chapter 18. CRC Press, Boca Raton, FL, August 2001.

[2] M. Biguesh, S. Shahbazpanahi, and A. B. Gershman. Robust downlink power control in wireless cellular systems. *EURASIP Journal on Wireless Communications and Networking*, 2004:261–272, December 2004.

[3] S. Boyd. *Convex Optimization.* Cambridge University Press, 2004.

[4] B. K. Chalise, S. Shahbazpanahi, A. Czylwik, and A. B. Gershman. Robust downlink beamforming based on outage probability specifications. *IEEE Transactions on Wireless Communications*, 6(10):3498–3503, October 2007.

[5] S. Fazeli-Dehkordy, S. Gazor, and S. Shahbazpanahi. Distributed peer-to-peer multiplexing using ad hoc relay networks. *Acoustics, Speech and Signal Processing, 2008. IEEE International Conference on*, pages 2373–2376, 2008.

[6] S. Fazeli-Dehkordy, S. Shahbazpanahi, and S. Gazor. Multiple peer-to-peer communications using a network of relays. *submitted to Signal Processing, IEEE Transactions on*, 2008.

[7] M.R. Garey and D.S. Johnson. *Computers and Intractability: A Guide to the Theory of NP-Completeness*. WH Freeman & Co. New York, NY, USA, 1979.

[8] A. Goldsmith. *Wireless Communications*. Cambridge University Press, 2005.

[9] D. Hammarwall, M. Bengtsson, and B. Ottersten. On downlink beamforming with indefinite shaping constraints. *Signal Processing, IEEE Transactions on*, 54(9):3566–3580, September 2006.

[10] I. Hammerstrom, M. Kuhn, and A. Wittneben. Impact of relay gain allocation on the performance of cooperative diversity networks. In *Proc. The IEEE 60th Vehicular Technology Conference, Fall 2004.*, volume 3, pages 1815–1819, Los Angeles, CA, September 2004.

[11] V. Havary-Nassab, S. Shahbazpanahi, A. Grami, and Z.-Q. Luo. Distributed beamforming for relay networks based on second order statistics of the channel state information. *Signal Processing, IEEE Transactions on*, 2008, to appear, see also Proc. *ICASSP* 2008.

[12] Y.-W. Hong, W.-J. Huang, F.-H. Chiu, and C.-C. J. Kuo. Cooperative communications in resource-constrained wireless networks. *IEEE Signal Processing Magazine*, 24:47–57, 2007.

[13] Y. Huang and S. Zhang. Complex Matrix Decomposition and Quadratic Programming. *Mathematics of Operations Research*, 32(3):758, 2007.

[14] M. Janani, A. Hedayat, T. E. Hunter, and A. Nosratinia. Coded cooperation in wireless communications: space-time transmission and iterative decoding. *Signal Processing, IEEE Transactions on*, 52:362–371, February 2004.

[15] Y. Jing and B. Hassibi. Distributed space-time coding in wireless relay networks. *IEEE Transactions on Wireless Communications*, 5(12):3524–3536, 2006.

[16] Y. Jing and H. Jafarkhani. Network beamforming using relays with perfect channel information. In *Proc. The IEEE International Conference on Acoustics, Speech and Signal Processing*, volume 3, Honolulu, HI, April 2007.

[17] Y. Jing and H. Jafarkhani. Distributed differential space-time coding for wireless relay networks. *Communications, IEEE Transactions on*, 56(7):1092–1100, 2008.

[18] Y. Jing and H. Jafarkhani. Network beamforming using relays with perfect channel information. Submitted, Dec. 2006.

[19] G. Kramer, M. Gastpar, and P. Gupta. Cooperative Strategies and Capacity Theorems for Relay Networks. *Information Theory, IEEE Transactions on*, 51(9):3037–3063, 2005.

[20] J. N. Laneman, D. N. C. Tse, and G. W. Wornell. Cooperative diversity in wireless networks: Efficient protocols and outage behavior. *Information Theory, IEEE Transactions on*, 50:3062–3080, December 2004.

[21] J. N. Laneman and G. W. Wornell. Distributed space-time-coded protocols for exploiting cooperative diversity in wireless networks. *Information Theory, IEEE Transactions on*, 49:2415–2425, October 2003.

[22] Z. Q. Luo and W. Yu. An introduction to convex optimization for communications and signal processing. *IEEE Journal on Selected Areas in Communications*, 24:1426–1438, August 2006.

[23] W.-K. Ma, T. N. Davidson, K. M. Wong, Z.-Q. Luo, and P.-C. Ching. Quasi-maximum-likelihood multiuser detection using semi-definite relaxation with application to synchronous CDMA. *Signal Processing, IEEE Transactions on*, 50:912–922, April 2002.

[24] C.D. Meyer. *Matrix Analysis and Applied Linear Algebra*. Society for Industrial Mathematics, 2000.

[25] A.L. Moustakas, S.H. Simon, and A.M. Sengupta. MIMO capacity through correlated channels in the presence of correlated interferers and noise: a (not so) large N analysis. *Information Theory, IEEE Transactions on*, 49(10):2545–2561, 2003.

[26] R. U. Nabar, H. Bolcskei, and F. W. Kneubuhler. Fading relay channels: performance limits and space-time signal design. *IEEE Journal on Selected Areas in Communications*, 22:1099–1109, August 2004.

[27] A. Nosratinia, TE Hunter, and A. Hedayat. Cooperative communication in wireless networks. *Communications Magazine, IEEE*, 42(10):74–80, 2004.

[28] F. Rashid-Farrokhi, K. J. R. Liu, and L. Tassiulas. Transmit beamforming and power control for cellular wireless systems. *IEEE Journal on Selected Areas in Communications*, 16(8):1437–1450, October 1998.

[29] F. Rashid-Farrokhi, KJR Liu, and L. Tassiulas. Downlink power control and base station assignment. *Communications Letters, IEEE*, 1(4):102–104, 1997.

[30] F. Rashid-Farrokhi, L. Tassiulas, and KJR Liu. Joint optimal power control and beamforming in wireless networksusing antenna arrays. *Communications, IEEE Transactions on*, 46(10):1313–1324, 1998.

[31] M. Schubert and H. Boche. Solution of the multiuser downlink beamforming problem with individual SINR constraints. *Vehicular Technology, IEEE Transactions on*, 53(1):18–28, January 2004.

[32] M. Schubert and H. Boche. Iterative multiuser uplink and downlink beamforming under SINR constraints. *Signal Processing, IEEE Transactions on*, 53(7):2324–2334, July 2005.

[33] A. Sendonaris, E. Erkip, and B. Aazhang. User cooperation diversity-part I: System description. *Communications, IEEE Transactions on*, 51:1927–1938, November 2003.

[34] A. Sendonaris, E. Erkip, and B. Aazhang. User cooperation diversity-part II: Implementation aspects and performance analysis. *Communications, IEEE Transactions on*, 51:1939–1948, November 2003.

[35] N. D. Sidiropoulos, T. N. Davidson, and Z.-Q. Luo. Transmit beamforming for physical-layer multicasting. *Signal Processing, IEEE Transactions on*, 54:2239–2251, June 2006.

[36] J. F. Sturm. Using SeDuMi 1.02, a Matlab toolbox for optimization over symmetric cones. *Optimization Methods and Software*, 11-12:625–653, 1999.

[37] D. Tse and P. Viswanath. *Fundamentals of Wireless Communication*. Cambridge University Press, 2005.

[38] P. Tseng. Further results on approximating nonconvex quadratic optimization by semidefinite programming relaxation. *SIAM J. Optim.*, 14(1):268–283, jul 2003.

[39] S. Zhang. Quadratic maximization and semidefinite relaxation. *Math. Program.*, 87:453–465, 2000.

[40] Y. Zhao and R. Adve. Improving amplify-and-forward relay networks: Optimal power allocation versus selection. In *Proc. IEEE International Symposium on Information Theory,*, pages 1234–1238, Seattle, WA, July 2006.

[41] L. Zheng and D. N. C. Tse. Diversity and multiplexing: a fundamental tradeoff in multiple-antenna channels. *Information Theory, IEEE Transactions on*, 49(5):1073–1096, May 2003.

Appendix A

Semidefinite Relaxation and The Duality Gap

We further explore the relationship between the original problem (3.18) and the relaxed problem (3.20). The Lagrangian of the optimization problem (3.18) can be written as

$$\begin{aligned} L(\mathbf{w}, \boldsymbol{\lambda}) &= \mathbf{w}^H \mathbf{D} \mathbf{w} + \sum_{k=1}^{d} \lambda_k (\gamma_k \sigma_n^2 - \mathbf{w}^H \mathbf{T}_k \mathbf{w}) \\ &= \mathbf{w}^H \left(\mathbf{D} - \sum_{k=1}^{d} \lambda_k \mathbf{T}_k \right) \mathbf{w} + \sum_{k=1}^{d} \lambda_k \gamma_k \sigma_n^2 \end{aligned} \tag{A.1}$$

where $\{\lambda_k\}_{k=1}^d$ are the non-negative dual variables and $\boldsymbol{\lambda} \triangleq [\lambda_1 \ \lambda_2 \ \ldots, \lambda_d]^T$. Therefore, the corresponding dual objective function is given by

$$g(\boldsymbol{\lambda}) = \min_{\mathbf{w} \in \mathbb{C}^R} L(\mathbf{w}, \boldsymbol{\lambda}) \,. \tag{A.2}$$

It is easy to show that the quadratic term in (A.1) is unbounded from below unless $\mathbf{D} - \sum_{k=1}^{d} \lambda_k \mathbf{T}_k \succeq \mathbf{0}$. Hence, the dual problem is

$$\begin{aligned} \max_{\boldsymbol{\lambda} \succcurlyeq \mathbf{0}} \quad & \sum_{k=1}^{d} \gamma_k \sigma_n^2 \lambda_k \\ \text{subject to} \quad & \mathbf{D} - \sum_{k=1}^{d} \lambda_k \mathbf{T}_k \succeq 0 \\ & \lambda_k \geq 0, \text{ for } \quad k = 1, 2, \ldots, d\,. \end{aligned} \tag{A.3}$$

It can be easily shown that the dual of the relaxed problem in (3.20) is exactly the same as (A.3). This means that the dual problems for both the original problem (3.18) and the relaxed problem (3.20) are the same. Since the maximum value of the dual problem provides a lower bound for the minimum value of the original problem, the maximum value of (A.3) is a lower bound for both optimal values of (3.18) and (3.20). On the other hand, because the strong duality holds for the relaxed problem [3], [22], the minimum value of (3.20) is the same as the maximum value of (A.3). Therefore, we conclude that the semidefinite relaxation provides the same lower bound to the original problem as the dual problem does. That is the gap between (3.18) and (3.20) is equal to the duality gap.

Appendix B

Proof of Lemma 1

To prove Lemma 1, we need to introduce some new notations. We define

$$\bar{\mathbf{R}}_{k,p} = \mathbf{D}_p^{-H/2}\mathbf{Q}_{kp}\mathbf{D}_p^{-1/2}, \tag{B.1}$$

and

$$\mathbf{D}_p^{1/2}\mathbf{w}_p = \sqrt{z_p}\mathbf{u}_p, \quad z_p \geq 0 \quad \text{for} \quad 1 \leq p \leq d, \tag{B.2}$$

where z_p is a non negative scaler and $\mathbf{u}_p$ is a unit vector for $1 \leq p \leq d$. The optimization problem (4.11) can be rewritten in the form

$$\begin{aligned} &\min \sum_{p=1}^{d} z_p \mathbf{u}_p{}^H \mathbf{u}_p \\ &\text{subject to } \mathbf{u}_k{}^H \left(z_k \bar{\mathbf{R}}_{k,k}\right) \mathbf{u}_k + \sum_{p \neq k} \mathbf{u}_p{}^H \left(z_p \bar{\mathbf{R}}_{k,p}\right) \mathbf{u}_p \\ &-\gamma_k \sigma_n^2 \geq 0, \ \text{for} \quad 1 \leq k \leq d, \end{aligned} \tag{B.3}$$

and if we let $\mathbf{U}_p = z_p \mathbf{u}_p {\mathbf{u}_p}^H$, the relaxed problem (4.14) is written as

$$\begin{aligned} &\min \sum_{p=1}^{d} \operatorname{tr}(\mathbf{U}_p) \\ &\text{subject to } \operatorname{tr}\left(\bar{\mathbf{R}}_{k,k}\mathbf{U}_k\right) + \sum_{p \neq k} \operatorname{tr}\left(\bar{\mathbf{R}}_{k,p}\mathbf{U}_p\right) \\ &\qquad -\gamma_k \sigma_n^2 \geq 0, \text{ and } \mathbf{U}_k \succeq \mathbf{0}, \text{ for } \quad 1 \leq k \leq d. \end{aligned} \tag{B.4}$$

First, we provide some useful lemmas.

Lemma 2 *The Lagrangian dual problem for the relaxed problem (B.4) is given by*

$$\begin{aligned} &\max_{\lambda_p} \sum_{p=1}^{d} \lambda_p \gamma_p \sigma_n^2 \\ &\textit{subject to } \mathbf{I} - \lambda_k \bar{\mathbf{R}}_{k,k} - \sum_{p \neq k} \lambda_p \bar{\mathbf{R}}_{p,k} \succeq \mathbf{0} \\ &\qquad \textit{for} \quad 1 \leq k \leq d. \end{aligned} \tag{B.5}$$

Proof: The Lagrangian of the optimization problem (B.4) is given by

$$\begin{aligned} L(\mathcal{U}, \boldsymbol{\lambda}) = &\sum_{p=1}^{d} \operatorname{tr}(\mathbf{U}_p) + \sum_{k=1}^{d} \lambda_k \gamma_k \sigma_n^2 \\ &- \sum_{k=1}^{d} \lambda_k \operatorname{tr}\left(\bar{\mathbf{R}}_{k,k}\mathbf{U}_k\right) - \sum_{k=1}^{d} \sum_{p \neq k} \lambda_k \operatorname{tr}\left(\bar{\mathbf{R}}_{k,p}\mathbf{U}_p\right), \end{aligned} \tag{B.6}$$

where $\mathcal{U} \triangleq \{\mathbf{U}_p \succeq \mathbf{0}\}_{p=1}^{d}$ and $\boldsymbol{\lambda} = [\lambda_1, \ldots, \lambda_d]^T$. It can be written as

$$\begin{aligned} L(\mathcal{U}, \boldsymbol{\lambda}) = &\sum_{p=1}^{d} \lambda_p \gamma_p \sigma_n^2 + \sum_{k=1}^{d} \operatorname{tr}(\mathbf{U}_k) \\ &- \sum_{k=1}^{d} \lambda_k \operatorname{tr}\left(\bar{\mathbf{R}}_{k,k}\mathbf{U}_k\right) - \sum_{k=1}^{d} \sum_{p \neq k} \lambda_p \operatorname{tr}\left(\bar{\mathbf{R}}_{p,k}\mathbf{U}_k\right), \end{aligned} \tag{B.7}$$

or

$$L\left(\mathcal{U},\boldsymbol{\lambda}\right)=\sum_{p=1}^{d}\lambda_p\gamma_p\sigma_n^2 + \sum_{k=1}^{d}\operatorname{tr}\left((I-\lambda_k\bar{\mathbf{R}}_{k,k}-\sum_{p\neq k}\lambda_p\bar{\mathbf{R}}_{p,k})\mathbf{U}_k\right) \tag{B.8}$$

The Lagrange dual function is written as

$$g(\boldsymbol{\lambda})=\inf_{\mathcal{U}} L\left(\mathcal{U},\boldsymbol{\lambda}\right) \tag{B.9}$$

The term in (B.8) is unbounded from below unless we have

$$\mathbf{I}-\lambda_k\bar{\mathbf{R}}_{k,k}-\sum_{p\neq k}\lambda_p\bar{\mathbf{R}}_{p,k}\succeq\mathbf{0} \quad \text{for } 1\leq k\leq d. \tag{B.10}$$

Hence, the Lagrangian dual problem for the relaxed problem (B.4) is given by

$$\max_{\lambda_p}\ \sum_{p=1}^{d}\lambda_p\gamma_p\sigma_n^2 \quad \text{subject to } \mathbf{I}-\lambda_k\bar{\mathbf{R}}_{k,k}-\sum_{p\neq k}\lambda_p\bar{\mathbf{R}}_{p,k}\succeq\mathbf{0} \quad \text{for} \quad 1\leq k\leq d. \tag{B.11}$$

Lemma 3 *The following problems have the same optimal cost function and the same optimal vector* $\boldsymbol{\lambda}$*.*

$$\min\ \sum_{p=1}^{d}\lambda_p{\mathbf{u}_p}^H\mathbf{u}_p \quad \textit{subject to } {\mathbf{u}_k}^H\left(\lambda_k\bar{\mathbf{R}}_{k,k}\right)\mathbf{u}_k+\sum_{p\neq k}{\mathbf{u}_p}^H\left(\lambda_p\bar{\mathbf{R}}_{k,p}\right)\mathbf{u}_p -\gamma_k\sigma_n^2\geq 0,\ \textit{for} \quad 1\leq k\leq d. \tag{B.12}$$

$$
\begin{aligned}
&\min \sum_{k=1}^{d} \lambda_k \gamma_k \sigma_n^2 \\
&\textit{subject to } \mathbf{I} - \lambda_p \bar{\mathbf{R}}_{p,p} - \sum_{k \neq p} \lambda_k \bar{\mathbf{R}}_{k,p} \nsucc \mathbf{0} \\
&\textit{and } \lambda_p \geq 0, \quad \textit{for} \quad 1 \leq p \leq d.
\end{aligned}
\tag{B.13}
$$

Proof: If we define

$$
[\mathbf{A}]_{k,p} = \mathbf{u}_p{}^H \bar{\mathbf{R}}_{k,p} \mathbf{u}_p \tag{B.14}
$$

the optimization problem (B.12) becomes

$$
\begin{aligned}
&\min \boldsymbol{\lambda}^T \omega \\
&\text{subject to } \mathbf{A}\boldsymbol{\lambda} = \mathbf{b} \\
&\lambda_p \geq 0 \\
&\|\mathbf{u}_p\| = 1 \;\; \text{for} \quad 1 \leq p \leq d,
\end{aligned}
\tag{B.15}
$$

where

$$
\omega = \left[\|\mathbf{u}_1\|^2, \ldots, \|\mathbf{u}_d\|^2\right]^T, \tag{B.16}
$$

$$
\mathbf{b} = \left[\gamma_1 \sigma_n^2, \ldots, \gamma_d \sigma_n^2\right]^T. \tag{B.17}
$$

We can write $\mathbf{A} = \mathbf{A}_1 - \mathbf{A}_2$ where

$$
[\mathbf{A}_1]_{k,p} = \begin{cases} \mathbf{u}_k{}^H \bar{\mathbf{R}}_{k,k} \mathbf{u}_k & p = k \\ 0 & p \neq k \end{cases} \tag{B.18}
$$

$$
[\mathbf{A}_2]_{k,p} = -\mathbf{u}_p{}^H \bar{\mathbf{R}}_{k,p} \mathbf{u}_p \tag{B.19}
$$

For any matrix $\mathbf{A}$ satisfying the constraints in (B.15) we have

$$\mathbf{A}\boldsymbol{\lambda} = \mathbf{b}, \tag{B.20}$$

$$(\mathbf{A}_1 - \mathbf{A}_2)\,\boldsymbol{\lambda} = \mathbf{b}, \tag{B.21}$$

$$\left(\mathbf{I} - \mathbf{A}_1{}^{-1}\mathbf{A}_2\right)\boldsymbol{\lambda} = \mathbf{A}_1{}^{-1}\mathbf{b}, \tag{B.22}$$

$$\lambda_p = \left[\mathbf{A}_1{}^{-1}\mathbf{b}\right]_p + \left[\mathbf{A}_1{}^{-1}\mathbf{A}_2\boldsymbol{\lambda}\right]_p > 0, \tag{B.23}$$

$$\frac{\left[\mathbf{A}_1{}^{-1}\mathbf{A}_2\boldsymbol{\lambda}\right]_p}{\lambda_p} = 1 - \frac{\left[\mathbf{A}_1{}^{-1}\mathbf{b}\right]_p}{\lambda_p} < 1. \tag{B.24}$$

From Collatz-Wielandt formula (see [24]) we have

$$r\left(\mathbf{A}_1{}^{-1}\mathbf{A}_2\right) \leq \max_{1\leq p\leq d} \frac{\left[\mathbf{A}_1{}^{-1}\mathbf{A}_2\boldsymbol{\lambda}\right]_p}{\lambda_p} < 1. \tag{B.25}$$

where r denotes the spectral radius of a matrix. From Perron-Frobenius theorem for irreducible non-negative matrices and because the spectral radius of matrix $\mathbf{A}_1{}^{-1}\mathbf{A}_2$ is less than 1 (see [24]), we conclude that $\left(\mathbf{I} - \mathbf{A}_1{}^{-1}\mathbf{A}_2\right)^{-1}$ exists and is non-negative. So we can write

$$\mathbf{A}^{-1} = \left(\mathbf{A}_1\left(\mathbf{I} - \mathbf{A}_1^{-1}\mathbf{A}_2\right)\right)^{-1} = \left(\mathbf{I} - \mathbf{A}_1^{-1}\mathbf{A}_2\right)^{-1}\mathbf{A}_1^{-1}. \tag{B.26}$$

Hence, $\mathbf{A}^{-1}$ also exists and is non-negative. So we can solve for $\boldsymbol{\lambda}$:

$$\boldsymbol{\lambda} = \mathbf{A}^{-1}\mathbf{b} = \left(\mathbf{I} - \mathbf{A}_1^{-1}\mathbf{A}_2\right)^{-1}\mathbf{A}_1^{-1}\mathbf{b}. \tag{B.27}$$

The optimization problem (B.15) can be rewritten as

$$\begin{aligned} &\min\ \mathbf{b}^T\mathbf{A}^{-T}\mathbf{1} \\ &\text{subject to } \mathbf{A}\boldsymbol{\lambda} = \mathbf{b} \\ &\qquad \lambda_p \geq 0 \\ &\qquad \|\mathbf{u}_p\| = 1 \ \text{ for } \ 1 \leq p \leq d. \end{aligned} \tag{B.28}$$

On the other hand, for any invertible matrix $\mathbf{A}$ with non-negative inverse we can always find $\boldsymbol{\lambda} \geq \mathbf{0}$ such that $\mathbf{A}\boldsymbol{\lambda} = \mathbf{b}$. Thus we can replace the first and the second constraint in (B.28) to write it as

$$\begin{aligned} &\min \mathbf{b}^T \mathbf{A}^{-T} \mathbf{1} \\ &\text{subject to } \mathbf{A} \in \mathcal{A} \\ &\|\mathbf{u}_p\| = 1. \end{aligned} \tag{B.29}$$

where $\mathcal{A}$ denotes the set of invertible matrices with non-negative inverse. Following the above discussion, we can rewrite (B.29) as

$$\begin{aligned} &\min \mathbf{b}^T \mathbf{A}^{-T} \mathbf{1} \\ &\text{subject to } \mathbf{A}^T \boldsymbol{\lambda} = \mathbf{1} \\ &\lambda_p \geq 0 \\ &\|\mathbf{u}_p\| = 1 \text{ for } \quad 1 \leq p \leq d, \end{aligned} \tag{B.30}$$

or equivalently

$$\begin{aligned} &\min \sum_{k=1}^{d} \lambda_k \gamma_k \sigma_n^2 \\ &\text{subject to } \mathbf{u}_p{}^H \left(\mathbf{I} - \lambda_p \bar{\mathbf{R}}_{p,p} - \sum_{k \neq p} \lambda_k \bar{\mathbf{R}}_{k,p} \right) \mathbf{u}_p = 0 \\ &\lambda_p \geq 0 \\ &\|\mathbf{u}_p\| = 1 \text{ for } \quad 1 \leq p \leq d. \end{aligned} \tag{B.31}$$

We can rewrite problem (B.31) as

$$\begin{aligned}
\min \ & \sum_{k=1}^{d} \lambda_k \gamma_k \sigma_n^2 \\
\text{subject to} \ & \frac{\mathbf{u}_p{}^H \lambda_p \bar{\mathbf{R}}_{p,p} \mathbf{u}_p}{1 - \sum_{k \neq p} \lambda_k \mathbf{u}_p{}^H \bar{\mathbf{R}}_{k,p} \mathbf{u}_p} = 1 \\
& \lambda_p \geq 0 \\
& \|\mathbf{u}_p\| = 1 \ \text{ for } \ \ 1 \leq p \leq d,
\end{aligned} \tag{B.32}$$

which is a virtual uplink problem except for the weight factors $\gamma_k \sigma_n^2$. The first d constraints are QoS requirements at the base station, one for each user. Note that we can relax the QoS constraints in (B.32) to obtain

$$\begin{aligned}
\min \ & \sum_{k=1}^{d} \lambda_k \gamma_k \sigma_n^2 \\
\text{subject to} \ & \frac{\mathbf{u}_p{}^H \lambda_p \bar{\mathbf{R}}_{p,p} \mathbf{u}_p}{1 - \sum_{k \neq p} \lambda_k \mathbf{u}_p{}^H \bar{\mathbf{R}}_{k,p} \mathbf{u}_p} \geq 1 \\
& \lambda_p \geq 0 \\
& \|\mathbf{u}_p\| = 1 \ \text{ for } \ \ 1 \leq p \leq d.
\end{aligned} \tag{B.33}$$

This is due to the fact that the SINR constraint for each user is always met with equality at the optimum. Otherwise, the transmit power at the corresponding user can be reduced without violating any constraint. This results in a lower cost function, which is a contradiction. According to this argument, we can relax the problem (B.31) to obtain

$$\begin{aligned} \min \; & \sum_{k=1}^{d} \lambda_k \gamma_k \sigma_n^2 \\ \text{subject to } & \mathbf{u}_p{}^H \left(\mathbf{I} - \lambda_p \bar{\mathbf{R}}_{p,p} - \sum_{k \neq p} \lambda_k \bar{\mathbf{R}}_{k,p} \right) \mathbf{u}_p \leq 0 \\ & \lambda_p \geq 0 \\ & \|\mathbf{u}_p\| = 1 \;\; \text{for} \quad 1 \leq p \leq d, \end{aligned} \tag{B.34}$$

which can be written as

$$\begin{aligned} \min \; & \sum_{k=1}^{d} \lambda_k \gamma_k \sigma_n^2 \\ \text{subject to } & \mathbf{I} - \lambda_p \bar{\mathbf{R}}_{p,p} - \sum_{k \neq p} \lambda_k \bar{\mathbf{R}}_{k,p} \not\succ \mathbf{0} \\ & \lambda_p \geq 0 \\ & \|\mathbf{u}_p\| = 1 \;\; \text{for} \quad 1 \leq p \leq d. \end{aligned} \tag{B.35}$$

Lemma 4 *Problems (B.11) and (B.35) have the same optimal solution* $\boldsymbol{\lambda}^*$.

Proof: First we show that the optimal solution to problem (B.11) is a feasible solution to problem (B.35) and vice versa. If we assume that the matrix in the kth constraint in Problem (B.35) has some negative eigenvalues at the optimum, then we can reduce the cost function by decreasing λ_k without violating any of the constraints, which is a contradiction. Similarly, if we assume that the matrix in the kth constraint in Problem (B.11) has only positive eigenvalues at the optimum, then we can increase the objective function by increasing λ_k while satisfying all the constraints, which is again a contradiction.

Now we assume that both $\boldsymbol{\lambda}^1$ and $\boldsymbol{\lambda}^2$ satisfy the constraints in both (B.11) and (B.35). Without loss of generality, we assume that there exists an index k and a constant $\alpha \geq 1$ such that $\lambda_k^1 = \alpha\lambda_k^2$ and $\boldsymbol{\lambda}^1 \leq \alpha\boldsymbol{\lambda}^2$. From the kth constraint in (B.11) we have:

$$\begin{aligned} \alpha\lambda_k^2\bar{\mathbf{R}}_{k,k} = \lambda_k^1\bar{\mathbf{R}}_{k,k} &\preceq \mathbf{I} - \sum_{p\neq k}\lambda_p^1\bar{\mathbf{R}}_{p,k} \\ &\preceq \mathbf{I} - \sum_{p\neq k}\alpha\lambda_p^2\bar{\mathbf{R}}_{p,k} \prec \alpha(\mathbf{I} - \sum_{p\neq k}\lambda_p^2\bar{\mathbf{R}}_{p,k}) \end{aligned} \tag{B.36}$$

which contradicts that $\boldsymbol{\lambda}^2$ satisfies the kth constraint in (B.35). Therefore, only a single point satisfies all the constraints in (B.11) and (B.35) which is the optimal solution to both problems.

Proof of Lemma 1: As it is shown in Lemma 4, the optimal solution of the Lagrangian dual of the relaxed problem (B.11) is the same as the optimal solution of (B.35), which is algebraically equivalent to the original problem (B.12), as a result of Lemma 3. The dual problem always provides a lower bound for the relaxed problem which in turn provides a lower bound for the original problem. Consequently all three problems have the same optimal solution that results in the following fact: whenever the problem (4.14) is feasible, there is at least one rank-one solution.