Mathematical Optimization and Signal Processing Techniques for Cooperative Wireless Networks

by

Georgia Bournaka

A Doctoral Thesis submitted in partial fulfilment of the requirements for the award of the degree of Doctor of Philosophy (PhD), at Loughborough University under a Joint PhD Programme with National Center for Scientific Research "Demokritos", Greece.

November 2013

Advanced Signal Processing Group,
School of Electronic, Electrical and Systems Engineering,
Loughborough University, Loughborough
Leicestershire, UK, LE11 3TU.

CERTIFICATE OF ORIGINALITY

This is to certify that I am responsible for the work submitted in this thesis, that the original work is my own except as specified in acknowledgements or in footnotes, and that neither the thesis nor the original work contained therein has been submitted to this or any other institution for a degree.

.... (Signed)

Georgia Bournaka   (Candidate)

I dedicate this thesis to my dear parents Dimitra and Vassilio Bournaka.

Abstract

The rapid growth of mobile users and emergence of high data rate multimedia and interactive services have resulted in a shortage of the radio spectrum. Novel solutions are therefore required for future generations of wireless networks to enhance capacity and coverage. This thesis aims at addressing this issue through the design and analysis of signal processing algorithms. In particular various resource allocation and spatial diversity techniques have been proposed within the context of wireless peer-to-peer relays and coordinated base station (BS) processing.

In order to enhance coverage while providing improvement in capacity, peer-to-peer relays that share the same frequency band have been considered and various techniques for designing relay coefficients and allocating powers optimally are proposed. Both one-way and two-way amplify and forward (AF) relays have been investigated. In order to maintain fairness, a signal-to-interference plus noise ratio (SINR) balancing criterion has been adopted. In order to improve the spectrum utilization further, the relays within the context of cognitive radio network are also considered. In this case, a cognitive peer-to-peer relay network is required to achieve SINR balancing while maintaining the interference leakage to primary receiver below a certain threshold.

As the spatial diversity techniques in the form of multiple-input-multiple-output (MIMO) systems have the potential to enhance capacity significantly, the above work has been extended to peer-to-peer MIMO relay networks. Transceiver and relay beamforming design based on minimum mean-square error (MSE) criterion has been proposed. Establishing uplink-downlink MSE duality, an alternating algorithm has been developed.

A scenario where multiple users are served by both the BS and a MIMO relay is considered and a joint beamforming technique for the BS and the MIMO relay is proposed. With the motivation of optimising the transmission power at both the BS and the relay, an interference precoding design is presented that takes into account the knowledge of the interference caused by the relay to the users served by the BS.

Recognizing joint beamformer design for multiple BSs has the ability to reduce interference in the network significantly, cooperative multi-cell beamforming design is proposed. The aim is to design multi-cell beamformers to maximize the minimum SINR of users subject to individual BS power constraints. In contrast to all works available in the literature that aimed at balancing SINR of all users in all cells to the same level, the SINRs of users in each cell is balanced and maximized at different values. This new technique takes advantage of the fact that BSs may have different available transmission powers and/or channel conditions for their users.

Contents

Statement of Originality

The contributions of this thesis are mainly on the development and analysis of novel coordinated resource allocation and spatial multiplexing techniques for wireless networks. The following aspects of this thesis are believed to be originals:

- Spatial diversity and power allocation techniques for peer-to-peer relay networks have been proposed in Chapter 3. The novelty of the work in this chapter is the consideration of user fairness in the form of signal-to-interference and noise ratio (SINR) balancing for the design of peer-to-peer relays as supported by [4]-[7]. The algorithms were first proposed for multiple one-way amplify and forward (AF) relays [4] and then it has been extended for multiple two-way AF relays [5]. An SINR balancing technique for underlay cognitive peer-to-peer relay network, where the ratio between the achieved SINR and the target SINR is balanced for all secondary users while satisfying the primary users' interference and total transmit power constrains was also proposed [6]. The SINR balancing algorithm has been extended to a two-way relay based cognitive radio (CR) relay network [7].

- A spectrum sharing peer-to-peer relay network has been proposed in Chapter 4, where multiple source nodes with multiple antennas communicate with their desired destination nodes with multiple antennas through a multiple-input and multiple-output (MIMO) relay. Initially the duality between uplink and downlink peer-to-peer MIMO channels with any number of antennas at each node has been established to demonstrate that the mean square error (MSE) of a downlink peer-to-peer network can be achieved in a virtual uplink network with the same total network transmission power. Using this MMSE duality result the weighted MSE minimization problem was solved using an

iterative algorithm that optimized the source, relay and receiver processing matrices. The novelty of this work is supported by [1].

- A new precoder design technique for a BS serving a number of users directly and another set of users through a MIMO wireless relay has been proposed in Chapter 5. The novelty of the work is the exploitation of the knowledge of the interference caused by the relay to the users served by the BS to perform more efficient interference mitigation which results into a lower overall network power consumption. The novelty is supported by [2].

- A coordinated multi-cell beamforming technique has been proposed for SINR balancing under multiple BS power constraints in Chapter 6. Instead of balancing SINR of all users in all cells to the same level as considered in the literature, a novel approach to balance SINR of users in various cells to different maximum possible values has been proposed. This has the ability to allow users in cells with relatively more transmit power or better channel condition to achieve a higher balanced SINR than that achieved by users in the worst case cells. This multi-level SINR balancing problem was solved using SINR constraints based SINR balancing criterion and subgradient method. The novelty of this work is supported by [3].

Journal Papers

1. G. Bournaka, K. Cumanan, S. Lambotharan and F. Lazarakis, "Transmitter receiver and relay optimisation for spectrum sharing multiple-input and multiple-output peer-to-peer users", IET Signal Processing, vol. 7, no. 5, pp. 411-419, July 2013

2. G. Bournaka, J. Tang, S. Lambotharan and F. Lazarakis, "A Coordinated Multiuser Relaying Technique through Interference Precoding

at the Base Station", IEEE Communication Letters, vol. 17, no. 6, pp. 1176-1179, June 2013.

3. G. Bournaka, Y. Rahulamathavan, K. Cumanan, S. Lambotharan and F. Lazarakis, "Cooperative Multicell Beamforming with Per Basestation SINR Balancing", submitted to IEEE Transactions on Wireless Communications.

Conference Papers

4. G. Bournaka, K. Cumanan, S. Lambotharan, J. A. Chambers and F. Lazarakis "An SINR balancing based multiuser relaying scheme", in Proc. IEEE International Conference on Communication Technology (ICCS), Singapore, pp. 274-278, 17-19 November 2010.

5. G. Bournaka, K. Cumanan, S. Lambotharan and F. Lazarakis, "An Iterative Semidefinite and Geometric Programming Technique for the SINR Balancing in Two-Way Relay Network", IEEE Global Communications Conference (Globecomm 2011), pp. 1-5, Houston, TX, 5-9 December 2011.

6. G. Bournaka, K. Cumanan, S. Lambotharan and F. Lazarakis, "SINR balancing techniques for a cognitive radio relay network with multiple peer-to-peer users", IEEE International Conference on Communication, Networks and Satellite (ComNetSat), pp. 108-112, Bali, Indonesia, July 2012.

7. G. Bournaka, K. Cumanan, S. Lambotharan and F. Lazarakis, "An SINR Balancing Technique for a Cognitive Two-Way Relay Network", IEEE Vehicular Technology Conference (VTC Fall), pp. 1-5, Quebec City, Canada, 3-6 September 2012.

Acknowledgements

Abbreviations

AF	Amplify and Forward
BER	Bit Error Rate
BS	Base Station
CDMA	Code Division Multiple Access
CR	Cognitive Radio
CSI	Channel State Information
dB	Decibel
DF	Decode and Forward
GP	Geometric Programming
ICI	Inter-cell Interference
ISI	Inter Symbol Interference
KKT	KarushKuhnTucker
MIMO	Multiple Input Multiple Output
ML	Maximum Likelihood
MMSE	Minimum Mean Square Error
MSE	Mean Square Error

PU	Primary User
QoS	Quality of Service
SDP	Semi Definite Programming
SINR	Signal to Interference plus Noise Ratio
SNR	Signal to Noise Ratio
SOCP	Second Order Cone Programming
SU	Secondary User
TDD	Time Division Duplex
Tr	Trace
ZF	Zero Forcing

List of Symbols

Scalar variables are denoted by plain lower-case letters, (i.e., a), vectors by bold-face lower-case letters, (i.e., $\mathbf{a}$), and matrices by upper-case bold-face letters, (i.e., $\mathbf{A}$). Some frequently used notations are as follows:

$\mathrm{E}\{\cdot\}$	Statistical expectation
$(\cdot)^T$	Transpose
$(\cdot)^H$	Hermitian transpose
$\|.\|$	Euclidean norm
$\|\|.\|\|_{\mathcal{F}}$	Frobenius norm
$(.)^{-1}$	Matrix inverse
$\mathbf{1}_M$	$M \times 1$ vector of ones
$\mathbf{0}_M$	$M \times 1$ vector of zeros
$\mathrm{Diag}(\cdot)$	Diagonal matrix with the arguments as the diagonal elements
$\mathrm{Tr}\{\cdot\}$	Trace operator
$\odot$	Schur-Hadamard product
$\mathbf{I}$	Identity matrix
$\mathbf{A} \succeq 0$	Positive semidefinite matrix

$\otimes$ Kronecker product

Vec(.)   Forms a vector by stacking the columns of a matrix

$[\mathbf{A}]_{ij}$ (i, j)th element of $\mathbf{A}$

$[\mathbf{a}]_i$ ith element of $\mathbf{a}$

$x \sim N(\mu, \sigma^2)$ Random variable x is drawn from a Gaussian distribution with mean μ and variance σ^2

List of Figures

List of Tables

Chapter 1

INTRODUCTION

Traditionally, different parts of the radio spectrum are licensed exclusively for different applications and services. For example, ultra-low frequencies in the range of 300-3000 Hz are used for submarine communications and super high frequencies in the range of 3-30 GHz are used for satellite communications. The current radio frequency (RF) spectrum allocation in the United States of America is presented in Figure 1.1. As seen in Figure 1.1, the spectrum is very crowded and there is not much radio spectrum available for new applications and services. Generally, cellular networks use the ultra high frequencies in the range of 300 MHz to 3 GHz. This range of frequency spectrum is very valuable and called a "sweet spot" as it has many advantages in terms of good transmission capabilities compared to other parts of the radio spectrum. This is because at lower frequencies, the signals will experience very good propagation characteristics, however the available bandwidth is very limited, but at high frequencies, considerably high bandwidth is available, but the signals would experience very poor propagation characteristics. The frequencies in the range of 300 MHz to 3 GHz provide a good balance between propagation characteristics and bandwidth.

When multiple wireless devices operate in the same frequency band, they often interfere with each other. Spectrum management is essential for mitigating interference among various wireless devices. The spectrum management process is called frequency or spectrum allocation. The current

Figure 1.1. A visual representation of the congestion in radio spectrum allocation in the United States of America [3]

spectrum management regime is known as "command and control" [4]. In this approach, the spectrum is licensed by frequency regulators to various operators or service providers for predefined services for a longer period with strict rules on the physical layer technologies employed. However, this fixed spectrum assignment policy is a bottleneck for achieving efficient spectrum utilization. This has motivated the development of communication technologies that can exploit the wireless spectrum in a more intelligent and flexible

way.

The first breakthrough technology that enhanced significantly the utilization of the radio spectrum can be attributed to cellular architecture, as it offered high capacity without requiring any major technological changes [5]. This technique exploits the exponential propagation loss and reuse of the same frequency at spatially separated locations. More specifically, in this traditional cellular architecture, the coverage area is divided into smaller areas called cells and each cell typically consists of a base station (BS), often located at the center of the cell, providing coverage to mobile stations. This basic design of the cellular network has not changed considerably in the last 30 years. At the same time the growing adoption of rich multimedia services and smart phones has led to unprecedented demand for wireless data capacity as seen in Figure 1.2.

The global mobile data traffic grew 70 percent in 2012 and it is ex-
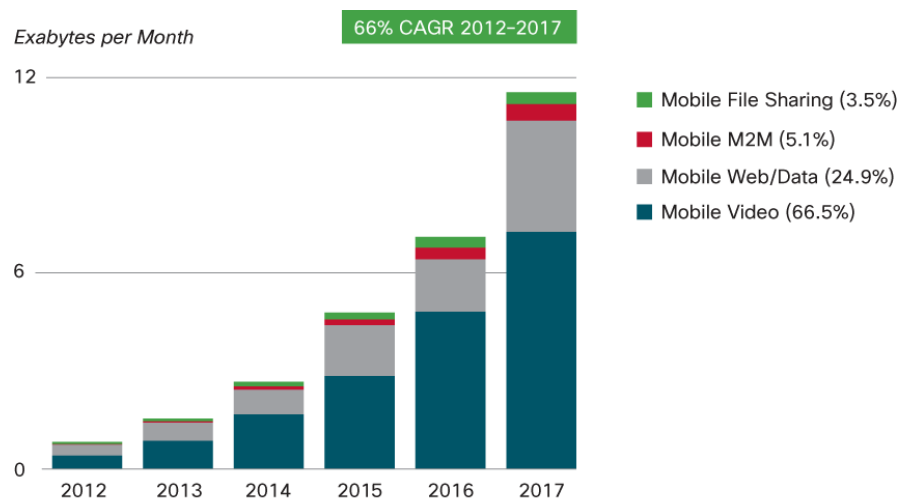
Figure 1.2. Mobile Video Will Generate Over 66 Percent of Mobile Data Traffic by 2017. Source: Cisco Mobile Forecast 2013 [1]

pected to grow to 11.2 exabytes per month by 2017, resulting in a 13-fold

Exabytes per Month

66% CAGR 2012–2017

12

6

0

0.9 EB — 2012
1.6 EB — 2013
2.8 EB — 2014
4.7 EB — 2015
7.4 EB — 2016
11.2 EB — 2017

Source: Cisco VNI Mobile Forecast, 2013
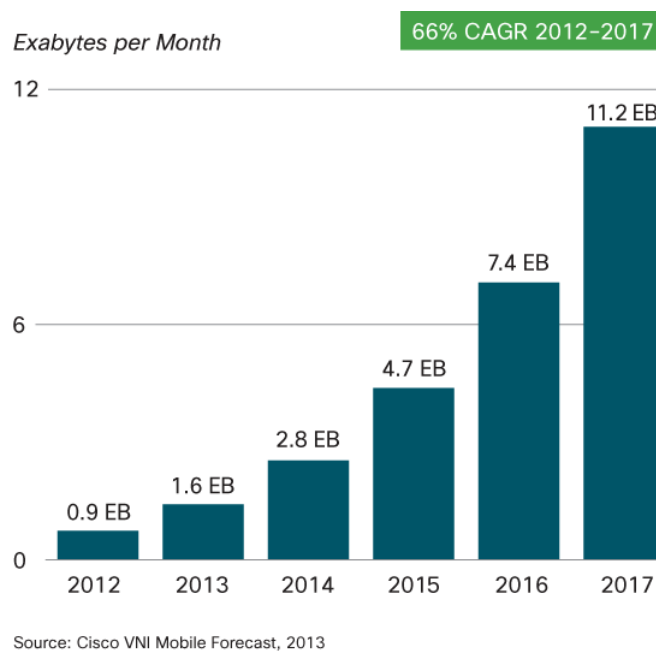
Figure 1.3. Cisco Forecasts 11.2 Exabytes per Month of Mobile Data Traffic by 2017. Source: Cisco Mobile Forecast 2013 [1]

increase (Figure 1.3 [1]). This rapid increase of the traffic volume and the subscriber number, has led to the need for more efficient way of utilizing the radio spectrum.

The possible ways of achieving this includes use of even smaller cells for delivering sufficient capacity. The gains from employing smaller cells come from higher area spectral efficiency [6]. Nonetheless, this approach leads to degradation of performance, in particular to the users at the cell edge who suffer from both high signal attenuation and severe inter-cell interference (ICI). One of the promising approaches to mitigate interference is array processing. The use of antenna arrays was an active area of research during World War II in radar systems [7]. Since 1970, more wider and sophisticated applications have emerged due to advances in digital signal processors.

One of the antenna array processing techniques that offers high diversity gains is beamforming [8]. In wireless communications, beamformers can be

employed at the transmitter as well as at the receiver. At the receiver side, the beamformer is designed in such a way to steer beam towards the direction of the interest and to attenuate signals arriving from other directions. Beamformer designs can be classified into different categories depending on various requirements and optimization criteria. The design based on the linearly constrained minimum variance beamformer (LCMV) [9], aims to put a constraint on the response of the beamformer so that signal from the direction of interest is passed with a specific gain and phase while interfering signals are attenuated. The weights of the beamformer are chosen to minimize output variance subject to the response constraint. This has the effect of preserving the desired signal, while minimizing the contributions to the output due to interfering signals arriving from directions other than the direction of interest. To determine this beamformer, the estimate of the angle of arrival is required. For uncorrelated sources, algorithms such as minimum variance distortionless response (MVDR) (also known as Capon) and MUSIC can be used. However, all of these methods i.e. LCMV, MVDR, MUSIC, require calibration of the array of antennas. In communication networks, generally pilot signals are available for the estimation of the channels or to design appropriate receivers directly. Hence, beamformers based on minimum-mean-square error (MMSE) can be designed with the aim of minimizing the difference between a desired response and the actual filter output. The optimum set of weights of the beamformer depends on the cross covariance between the beamformer input signal vector and the training signal [10]. This category of beamformer is distinguished from the LCMV from the fact that the direction of the desired signal is not required to be known. If the requirement is to maximize the signal to noise ratio (SNR), then for the design of the beamformer, the knowledge of both the desired and noise covariance matrices is required [11].

As mentioned already, the beamformers can be employed not only at the

receiver but also at the transmitter. The main motivation for the design of the beamformers at the transmitter is that the deployment of antenna arrays at the mobiles maybe impractical. In addition, BS's higher processing power capability enables easier deployment of the beamformers. The beam-pattern of each antenna array can be adjusted to maximize the signal in the direction of interest and minimize the induced interference to other receivers. For determining appropriate transmit beamformer, either a feedback channel is required to obtain the channel state information (CSI) from the receiver or the receiver needs to calculate the transmitter beamforming coefficients and inform to the BS [12]. In either way, the bandwidth required for transmission of information back to the BS is considerably high and for this reason, quantized information is required, for example Grassmannian Beamforming techniques can be used [13, 14]. The transmit and receive diversity have been widely adopted in third-generation (3G) code division multiple access (CDMA) cellular systems [15] as well as in 802.16 worldwide interpretability for microwave access (WiMaX).

In wireless networks, when a BS aims to transmit signals simultaneously to multiple users in the same frequency band, directional antennas has to be employed at the transmitter. This innovative means of using the available spectrum efficiently is called spatial multiplexing and it requires multiple antennas at the transmitter to steer multiple beams to multiple users. In this case, it is crucial that the BS allocates power optimally to each user. Joint power control and beamforming techniques have been proposed in the literature to satisfy different Quality of Services (QoS) criteria. There are numerous different optimizations that are of interest. Some examples include maximizing total throughput given a constraint on total transmission power, or minimizing transmission power subject to a set of QoS constraints. An overview of the problem of minimizing transmission power under a set of signal to interference plus noise ratio (SINR) constraints for each user can

be found in [16]. Another interesting problem that was studied in [17] is the user fairness based beamformer design. The aim is to maximize the SINR of the worst case receiver in a multiple user scenario subject to a total power constraint. This leads every user achieving the same SINR, namely SINR balancing.

Spectral efficiency, coverage and quality of wireless links can be improved further by employing multiple antennas at both the transmitter and the receiver [18]. This scheme is known as multiple-input-multiple-output (MIMO) system. The large spectral efficiencies associated with the MIMO systems is based on the assumption that a rich scattering environment provides independent paths from each transmit antenna to each receive antenna. The gain obtained is called spatial diversity gain. Space-time coding to exploit spatial diversity gain in a point-to-point MIMO channel has been studied in [19, 20].

Different assumptions can be made about the CSI at the transmitter and the receiver. Consider a MIMO channel with M transmit and N receiver antennas and assume that the channel side information is not known at the transmitter. This yields a system of N equations with M unknowns. The most simple way of estimating the transmitted signal at the receiver, is to process the received signal through the inverse of the MIMO channel matrix. Assuming the channel matrix is of full rank and has independent entries based on the assumption of a rich scattering environment, the capacity of this scheme scales linearly with $\min(M, N)$ compared to a system with one antenna at the transmitter and the receiver [21]. The capacity can be increased further if the channel is known at the transmitter. If the CSI is known to the transmitter via feedback from the receiver or due to the reciprocity of channels as in a time division duplex (TDD) system, the MIMO channel can be decomposed into parallel independent spatial sub-channels. The power can also be allocated optimally to each sub-channel based on wa-

ter filling techniques. MIMO is being widely considered for next-generation cellular systems such as IEEE 802.16/WiMAX and Third Generation Partnership Project (3GPP) Long Term Evolution (LTE).

For applications such as cellular telephony, MIMO systems can be deployed in an environment where a single BS must communicate with many users simultaneously. Multi-user systems with multiple antennas both at the BS and the users have the advantage to combine the high capacity achievable with MIMO processing, with the benefits of space-division-multiple access (SDMA). Adding multiple antennas at each user facilitates the transmission of multiple parallel data streams to multiple users . The CSI at the transmitter is very useful not only for achieving a high SNR at the receiver but also to reducing the interference introduced to various users.

For multiuser MIMO systems, two problems can be considered. The downlink, where the BS attempts to transmit signals to multiple users and the uplink where a group of users transmits data to the BS. In the uplink no coordination between users is assumed and for this reason the downlink problem differs from the uplink problem. In the uplink, the challenge is for the BS to separate the signals transmitted by different users using array processing, multi-user detection or other methods. In the downlink, the main challenge is the multiple access interference caused by the BS due to simultaneous transmissions of signals. The inter-user interference can be mitigated by intelligent beamforming techniques or by dirty paper coding techniques [22].

So far the uplink and downlink transmissions within the context of single cell based processing have been described. Due to characteristics of wireless propagation, signals transmitted from the BS do not confine within the corresponding cell only. Users will experience interference from other BSs as well. In the 3rd generation partnership project (3GPP) LTE systems, ICI coordination or avoidance has been proposed to deal with the ICI by

managing time or frequency or power resources available for each cell in a coordinated way [23]. Such approaches provide improvement in the SINR ratio.

Coordinated Multipoint Transmission (CoMP) is the first cooperative MIMO scheme proposed for cellular networks. The BSs have the ability to communicate to each other through high-speed reliable connection, possibly consisting of optical fiber links that is capable of having very high data rates. This enables them to cooperate and process different users' signals. These multi-BS MIMO cooperation is expected to play an important role in terms of interference mitigation. Coordinated multi-cell beamforming techniques can improve substantially the received signal quality and decrease the received spatial interference. Relays also form an important part of coordinated processing.

Instead of cooperating through backhaul links, a separate relay node that assists the direct communication within each cell has similar benefits to multi-cell MIMO network because it can strengthen the effective direct channel gain between the BS and the remote users and helps with ICI mitigation. Relays are also beneficial to reduce infrastructure deployment costs. This is essential under the demand for high data rates envisioned for fourth generation (4G) wireless systems.

Transmission power for relays are generally less than that is required by BSs, as relays are expected to cover a region smaller than that of a BS. The wireless relays receive the data from the BS and forward them to the users. The costs for provisioning the wired backhaul connection is not required for the wireless relays [24]. With the deployment of relays, the distance between the transmitter and receiver decreases, hence the propagation loss is smaller than if only the BS is employed. This becomes more important as future generation wireless networks are likely to use very high radio frequency resulting into higher channel attenuation, lower coverage while providing in-

creased bandwidth. Relays will form an important part of the network.

Capacity can be increased further by exploiting spatial diversity and enabling simultaneous transmissions by both the BS and the relays. Regardless of the benefits, it should be understood that there are certain disadvantages as the relays may introduce additional delays and increase interference due to frequency reuse at the relays.

Even though relays and multi-cell processing have the ability to enhance capacity and coverage, the ever increasing demand for spectrum requires more advancement in terms of spectrum usage. As stated earlier, wireless technologies and users' demands are growing rapidly and the conventional fixed spectrum allocation approach is inefficient to meet this increasing demand. Cognitive radio (CR) is a promising technology that aims to mitigate the spectrum scarcity problem by allowing secondary (unlicensed) users (SUs) to access licensed frequency bands under the condition of protecting primary (licensed) users (PUs) from harmful interference. Various approaches have been developed regarding the way the SUs can access the licensed spectrum [25]. The first one is through opportunistic spectrum access, by using a frequency band only when it is detected to be idle. This is called an interweave approach, through which the utilization of the spectrum is improved by opportunistic reuse of temporary frequency voids, which are called spectrum holes. The second approach is through spectrum sharing, when the SUs coexist with the primary users by keeping the interference in the primary users below a certain acceptable threshold. This is also called underlay approach and allows only short-range communication for the SUs due to interference power constrains.

In summary, use of multiple antennas, exploitation of spatial diversity, intelligent resource allocation techniques, together with efficient interference management and coordinated multi-cell and relay processing techniques are the promising technical approaches for enhancing the spectral efficiency of

wireless networks. All these approaches have been exploited to propose various novel algorithms in this thesis.

Thesis Outline

Spectrum scarcity is one of the most important challenges that wireless communications need to address, especially due to increasing demand for wireless services and ever increasing number of users. The work in this thesis is focused on techniques that improve spectral efficiency, coverage and quality of wireless links. New resource allocation and spatial multiplexing techniques have been proposed for cooperative communication networks including peer-to-peer, relay and multi-cell networks.

In Chapter 2, a detailed literature survey is provided together with the necessary theoretical background. The chapter includes characterization of wireless channels followed by a detailed discussion on spatial diversity techniques, such as transmit and receive beamforming. Existing wireless relay techniques have also been summarized. Non-regenerative, regenerative and two-way relay techniques are introduced and the advantages and the disadvantages are discussed. Finally, a discussion on cooperative diversity schemes is provided.

In Chapter 3, spatial diversity and power allocation techniques for peer-to-peer relay networks, where the source and the destination nodes are equipped with single antenna have been investigated. An SINR balancing based relay signal forwarding scheme is proposed. The beamforming problem is formulated using a semidefinite programming (SDP) framework and the power allocation problem is solved using geometric programming (GP). Initially, this resource allocation problem is considered for multiple one-way amplify and forward relays and then it is extended for multiple two-way amplify and forward relays. SINR balancing technique for an underlay cognitive peer-to-peer relay network is also proposed, where the SINR is balanced for

all secondary users while satisfying the primary users' interference threshold and total transmit power constrains.

In Chapter 4, the work in Chapter 3 on peer-to-peer relay networks are extended for multiple antennas at the source and the destination nodes as well as the relay. The uplink-downlink duality for this problem is established. The MMSE of a downlink peer-to-peer network is shown to be achieved by a virtual uplink network with the same total network transmission power. Based on this result, an iterative algorithm is proposed to determine the source, relay and receiver processing matrices, such that the sum mean square error (MSE) of the retrieved signal at the receivers is minimized.

Chapter 5 investigates coordinated beamforming design in a network with one BS and a relay. Most of the works in the literature considered optimal beamformer design at the BS and the relays but they have not explicitly considered the backhaul overheads required to transmit the signal from the BS to the relay. The knowledge of the interference caused by the relay to the users served by the BS is exploited to enhance the overall network performance. The optimization problem is solved using second order cone programming (SOCP) and GP. The proposed method is shown to reduce significantly the transmission power by using interference precoding at the transmitter for various sets of data rates requirement.

In Chapter 6, a coordinated multi-cell beamforming technique is proposed for SINR balancing under multiple BS power constraints. Instead of balancing SINR of all users in all cells to the same level, a new approach is proposed to balance SINR of users in various cells to different maximum possible values. This has the ability to allow users in cells with relatively more transmit power or better channel condition to achieve a higher balanced SINR than that achieved by users in the worst case cells. This multi-level SINR balancing problem is solved using SINR constraints based SINR bal-

ancing criterion and subgradient method.

Conclusions are drawn in Chapter 7. Possible future research directions are also discussed.

Chapter 2

SPATIAL DIVERSITY AND WIRELESS RELAY NETWORKS

2.1 Introduction

In this chapter a literature review and the underlying technical background required for the contributing chapters are provided. Starting with the description of various communication channel models, various techniques that are important to maximize spectrum efficiency are provided. In particular, the focus is on spatial diversity techniques, such as beamforming, multiuser spatial diversity, MIMO technologies, wireless relay technologies including one-way and two-way relays and cooperative communication networks. Particular emphasis is placed on cooperative relay networks and coordinated multi-cell processing.

2.2 Characterization of the channel

Due to non-ideal transmission media that could also vary over time and frequency, the transmitted signal gets distorted. For example, when a pulse is transmitted, the received signal may appear as a pulse train due to multi-paths. For this case, a typical impulse response may appear as depicted in
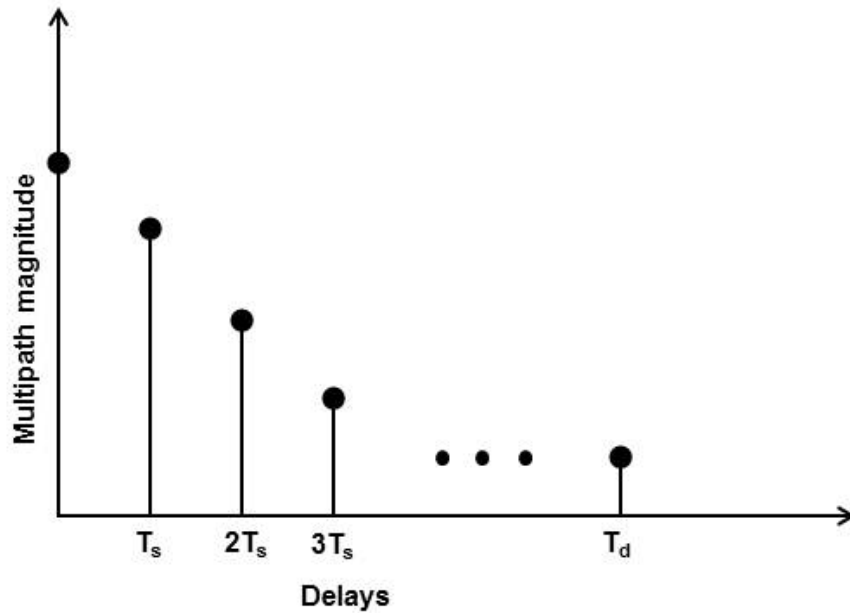
Figure 2.1. An example of delay profile

Figure 2.1. In addition, the nature of multipath could vary over time. This variation happens because either the transmitter or the receiver moves and also possibly as a result of variations of the location of the reflectors in the transmission path. In this case, the received signal strength could change as depicted in Figure 2.2. Generally, the time-variant channel response $h(\tau, t)$ can be modeled as a complex-valued Gaussian random process in time variable t and delay τ. The envelope $|h(\tau, t)|$ at any instant t is generally characterized by Rayleigh or Ricean distribution. If the real and complex parts of the channel impulse response are modeled by a zero-mean Gaussian random process, then this channel is said to be Rayleigh fading. In addition if there is a strong line of sight (LOS) path, then the channel will obey Ricean fading [26]. However, throughout this thesis, the channel is assumed to be Rayleigh fading.

The channel impulse response is generally assumed to be wide-sense stationary with uncorrelated scattering $h(\tau, t)$, where τ represents the impulse
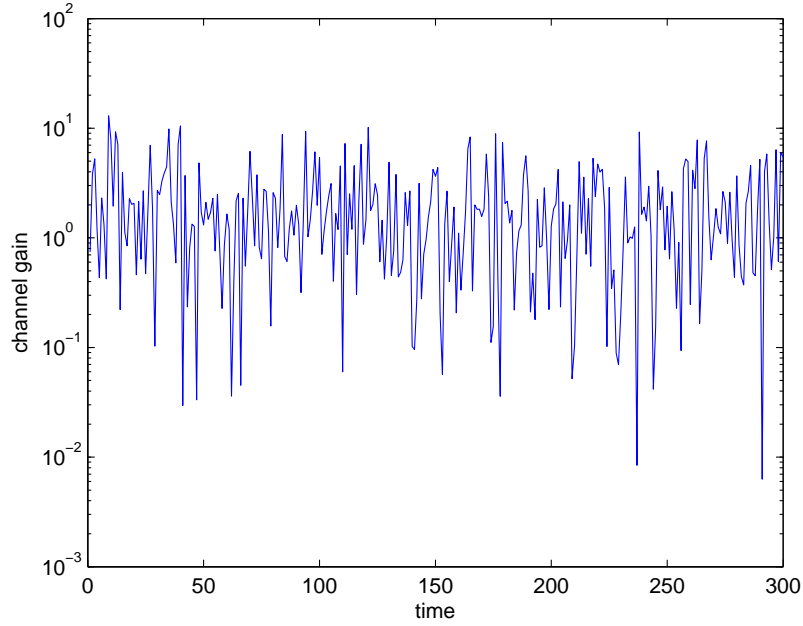
Figure 2.2. An example of typical fading profile

response associated with a given multipath delay and t represents the time variation. The autocorrelation is given by

$$\Phi_{hh}(\tau, t) = E\{h(\tau; t)h^*(\tau; t + \Delta t)\} \tag{2.2.1}$$

when $\Delta t = 0$, $\Phi_{hh}(\tau, 0) = \Phi_{hh}(\tau)$ is called multipath intensity profile or power delay profile when $\Delta t = 0$. The range of values of τ for which $\Phi_{hh}(\tau)$ is non zero is called the multipath spread of the channel, denoted by T_m.

The time-varying multipath channel can be characterized in the frequency domain by taking the Fourier transform of $h(\tau, t)$ with respect to τ, denoted by $H(f, t)$. The autocorrelation in the frequency domain is given by:

$$\varphi_{hh}(\Delta f, \Delta t) = E\{H(f; t)H^*(f + \Delta f; t + \Delta t)\} \tag{2.2.2}$$

and it is called the spaced-frequency, spaced-time correlation function of the channel [26].

If $\Delta t = 0$ is set to 0 in (2.2.2) the autocorrelation function in the frequency variable $\varphi_{hh}(\Delta f)$ can be obtained, which provides a measure of the frequency coherence of the channel i.e. the correlation between two frequencies. The bandwidth of this function $(\Delta f)_h$ is called coherence bandwidth of the channel and it is the reciprocal of the multipath spread as described by the equation:

$$(\Delta f)_h \approx \frac{1}{T_m} \tag{2.2.3}$$

If the bandwidth of the transmitted signal is greater than $(\Delta f)_h$, the channel is called frequency-selective and the channel amplitude values for frequencies separated by more than the coherence bandwidth are independent. On the other hand, if the bandwidth of the signal is narrower than the coherence bandwidth $(\Delta f)_h$, then the fading across the entire bandwidth is highly correlated and referred as flat fading.

The frequency selective channel can be also interpreted in terms of inter symbol interference (ISI). If the period of a transmitted symbol T is assumed to be smaller than T_m, then the received signal will consist of a number of past symbols introducing ICI.

The time variations of the channel is characterized by another statistical property, namely the coherence time $(\Delta t)_h$, which is a measure of the correlation of the channel impulse response at times separated by (Δt). Mathematically, it is the range of Δt values over which $\varphi_{hh}(\Delta f, \Delta t)$ is non zero. If Fourier transform is performed on $\varphi_{hh}(\Delta f, \Delta t)$ with respect to the variable Δt and $\Delta f = 0$ is set, the function $S(\lambda)$ is called the Doppler power spectrum of the channel. The range of values λ over which $S(\lambda)$ is nonzero is called the Doppler spread B_d of the channel. The coherence time is related

to the Doppler spread of the channel as $(\Delta t)_h \approx \frac{1}{B_d}$. The coherence time is a measure of the speed of change of the channel.

All the techniques that will be described in this chapter and throughout the thesis will consider only flat channels. It is also possible to have multipath channels, known as frequency selective channels, as described in this section. However, throughout the thesis, technology based on OFDM is assumed, hence in each subcarrier, the channel is considered to be flat and may or may not fade.

2.3 Spatial Diversity Techniques

One of the favorite techniques to mitigate the effects of fading is spatial diversity method. When a particular channel goes into a deep fade, the possibility of losing information is high. However, if the same signal is transmitted possibly through many uncorrelated paths, the information could be recovered with high probability. Suppose two independent fading paths undergo deep fading with probabilities P_1 and P_2 respectively, then the probability that both paths undergo deep fading simultaneously is $P_1 P_2$, which is much less than the marginal probabilities P_1 and P_2. Such independent paths can be obtained by using multiple transmit or receive antennas when they are placed sufficiently distant. The underlying technique is called spatial diversity.

Spatial diversity techniques are widely used in wireless communications. There are various kinds of spatial diversity methods such as beamforming, space-time block coding, multiuser multiplexing and use of multiple antennas at the transmitter and the receiver, known as MIMO network. Spatial diversity techniques have the ability to enhance the SINR at the receiver, thereby facilitate increased coverage and better link quality. They have the potential

to enhance the data rate for wireless communication systems substantially. In the following subsections, various diversity techniques commonly used in wireless communications are presented.

2.3.1 Beamforming

Beamforming is a signal processing technique used in the physical layer of a communication system to control the directionality of transmission or reception of signals using an array of antennas [8].

Receiver Beamforming

A typical receiver beamforming structure is shown in Figure 2.3,
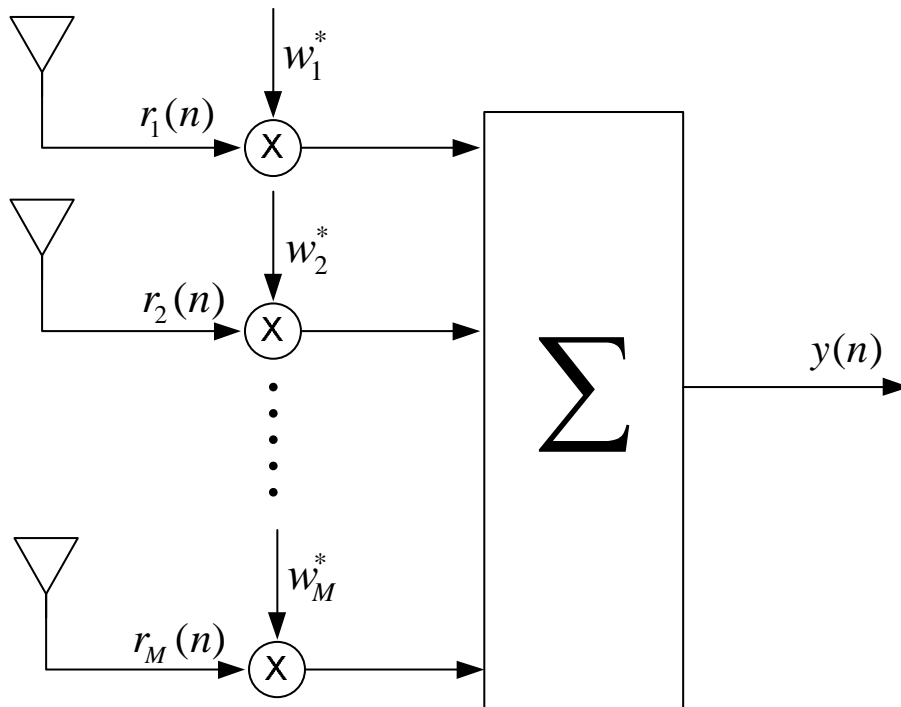
Figure 2.3. A receiver beamformer design.

The signals picked up by M antennas are linearly combined by complex coefficients, $\mathbf{w}_i^*$, known as beamformer coefficients. The output of the beamformer $y(n)$ is written in terms of the signals picked up by the anten-

nas $\mathbf{x}(n) = [x_1(n) \cdots x_M(n)]$ and the beamformer weight vector $\mathbf{w}(n) = [w_1 \ w_2 \ \cdots \ w_k]^T$ as follows:

$$y(n) = \mathbf{w}^H(n)\mathbf{x}(n) \qquad (2.3.1)$$

where n is the time index. When the arrays are placed linearly and if the source is located far away from the receiver, satisfying far field assumption, the signal impinged on the array of the antennas, $\mathbf{x}(n)$, can be written in terms of the source signal $s(n)$ and the array response vector $\mathbf{s}(\theta) \in \mathbb{C}^{M \times 1}$ as:

$$\mathbf{x}(n) = s(n)\mathbf{s}(\theta) + \mathbf{n}(n), \qquad (2.3.2)$$

where the M-by-1 steering vector is defined:

$$\mathbf{s}(\theta) = [1, e^{-j\theta_0}, \cdots, e^{-j(M-1)\theta_0}]^T, \qquad (2.3.3)$$

θ_0 is the angle of arrival of the source and $\mathbf{n}$ is the noise vector.

In the absence of interference, as we assumed in this case, the optimum beamformer vector can be written as

$$\mathbf{w} = \mathbf{s}(\theta) \qquad (2.3.4)$$

which coherently combines the signal at the receiver. It is also known otherwise as a matched filtering. However, in the presence of multiple users in the environment, the signal impinged on the array of the antennas $\mathbf{x}(n)$, can be written using the following equation:

$$\mathbf{x}(n) = \mathbf{s}(\theta_0)s_0(n) + \sum_{i=1}^{l} \mathbf{s}(\theta_i)s_i(n) + \mathbf{n}(n) \qquad (2.3.5)$$

where θ_0 is the angle of arrival of the desired source $s_0(n)$ and θ_i is the angle of arrival of the interfering source $s_i(n)$. In this case, in order to retrieve the source signal $s_0(n)$, it is important that the beamformer should suppress the undesired interference.

Various criteria can be used to design appropriate receiver beamformer in this case. A well known criterion that aims to keep the beamformer gain for the signal of interest $s_0(n)$ at 1 and minimizes the interference induced by all other sources at the output of the beamformer, is known as minimum variance distortionless response (MVDR). In this case, the MVDR beamformer can be written as in the following equation

$$\mathbf{w}_o = \frac{\mathbf{R}^{-1}\mathbf{s}(\theta_0)}{\mathbf{s}^H(\theta_0)\mathbf{R}^{-1}\mathbf{s}(\theta_0)} \tag{2.3.6}$$

where $\mathbf{R} = E\{\mathbf{x}(n)\mathbf{x}^H(n)\}\mathbb{C}^{M \times M}$ is the correlation matrix of the signal array input vector $\mathbf{x}(n)$ and $\mathbf{s}(\theta_0)$ is the array response vector for angle $\theta = \theta_0$. In order to design the above beamformer it is important to have the knowledge of θ_0. Traditionally various algorithms can be used to estimate this angle of arrival, such as Capon spectral vector estimation, Multiple Signal Classification (MUSIC) and estimation of signal parameters via rotational invariance techniques (ESPRIT) [27]. In order for these above algorithms to work properly, the antennas need to be calibrated, which is not viable or attractive in wireless communication environment. However, angle of arrival θ_0, or the array response vector $\mathbf{s}(\theta_0)$, can be estimated if the source could transmit a training signal. If the channel can be estimated in this way, the above technique can still be used. However, it is also possible to determine the optimum beamformer directly based on the training signal.

There are various criteria, such as MMSE, Least Squares (LS) and SINR, that can be used to determine the beamformers. In terms of beamformer design based on MMSE, the error signal $e(n)$ is written as the difference

between the desired response $d(n)$ and the beamformer output $y(n)$ as

$$e(n) = d(n) - y(n) = d(n) - \mathbf{w}^H \mathbf{x}(n). \tag{2.3.7}$$

The beamformer that maximizes the MSE, $J(n) = E\{|e(n)|^2\}$ can be written as

$$\mathbf{w}_{opt} = \mathbf{R}^{-1}\mathbf{P} \tag{2.3.8}$$

where $\mathbf{R} = E\{\mathbf{x}(n)\mathbf{x}^H(n)\}$ is the covariance matrix of the array signal vector and $\mathbf{P} = E\{\mathbf{x}(n)s_0(n)\} = \mathbf{s}(\theta_0)\sigma_{s_0}^2$ is the cross correlation between the array signal vector $\mathbf{x}(n)$ and the desired signal $s_0(n)$. Here the desired signal is the same as the transmitted training signal, the variance of which is $\sigma_{s_0}^2$. The above technique is also known as Wiener filter.

In order to use SINR criterion, we have to first formulate the SINR ratio in terms of the beamformer weight vector $\mathbf{w}$ as follows:

$$\text{SINR} = \frac{\mathbf{w}^H \mathbf{A} \mathbf{w}}{\mathbf{w}^H \mathbf{B} \mathbf{w}} \tag{2.3.9}$$

where $\mathbf{A} = \mathbf{s}(\theta_0)\mathbf{s}^H(\theta_0)\sigma_{s_0}^2$ and $\mathbf{B} = \sum_{i=1}^{l} \mathbf{s}(\theta_\mathbf{i})\mathbf{s}(\theta_\mathbf{i})^H \sigma_{s_i}^2 + \sigma_n^2$. In this case, the beamforming vector $\mathbf{w}$ that maximizes the SINR is the generalized eigenvector corresponding to the largest generalized eigenvalue of the matrix pair $\mathbf{A}$ and $\mathbf{B}$ given by

$$\mathbf{B}^{-1}\mathbf{A}\mathbf{w} = \lambda_{\max}\mathbf{w}, \tag{2.3.10}$$

where $\lambda_{\max}$ is the largest eigenvalue of the matrix $\mathbf{B}^{-1}\mathbf{A}$.

Transmitter Beamforming

Even though beamforming has the ability to enhance the performance substantially, employing multiple antennas at the receiver may not be so attractive due to the computational complexity. It would be better if this complexity can be transferred to the transmitter (downlink is assumed), by employing beamforming at the transmitter. Designing beamforming at the transmitter (downlink beamformer) substantially differ in several aspects from designing a beamformer at the receiver. In the latter, the design determines the performance of a specific user, whereas the transmit beamformer affects the performance of all users in its coverage area. Another difference concerns the channel knowledge. For designing the receiver beamforming, the receiver could estimate the channel coefficients using the training signal or by exploiting certain properties of the desired signal as discussed in the previous section. For transmitter beamformer, the channel knowledge could be made available to the transmitter by sending the estimates of the CSI from the receiver through a finite rate feedback channel [13, 28–31].

A typical transmitter beamformer structure with M antennas is shown in Figure 2.4. The transmitted signal can be written as

$$\mathbf{x}(n) = \mathbf{w}s(n) \qquad (2.3.11)$$

where $\mathbf{w}$ is the transmitter beamforming vector and $s(n)$ is the information symbol. For the case of a single user, if the channel between the transmitter and the receiver is denoted as $\mathbf{h}$, then the beamformer that maximizes the SNR at the receiver, is the conjugate of the channel $\mathbf{h}$. In other words, the beamformer rotates the phase of the signals, so that the signals arriving at the receiver through different paths add coherently in phase.

When there are multiple users in the environment, spatial multiplexing is possible at the transmitter, i.e., multiple beams can be steered to multiple
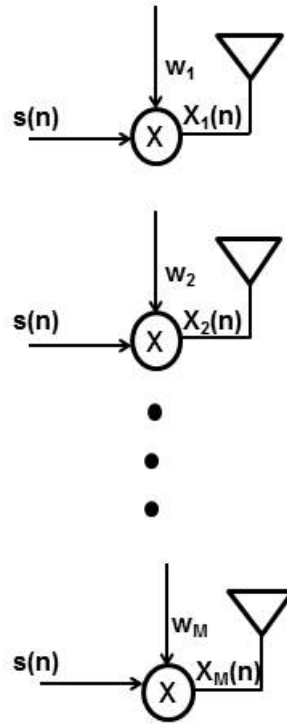
Figure 2.4. A transmitter beamformer design.

users. In this case, we can enable multiple users to access the same frequency band simultaneously, thereby enhancing the overall network throughput. A very simple mechanism to design such a spatial multiplexing technique is to use a zero forcing, also known as block diagonalization [32–36]. In this case the precoder of a particular user is determined as the one that is in the nullspace of the channels of all other users. However this scheme has a limitation as it requires substantially large number of antennas. In particular the number of antennas that are required at the transmitter should be at least greater than the number of users in the environment [37]. The received signal at the kth user is written as:

$$\mathbf{y}_k(n) = \mathbf{h}_k^H \mathbf{x}(n) + n_k(n) \tag{2.3.12}$$

where $\mathbf{h}_k = [h_1^{(k)} \cdots h_M^{(k)}]^H$ is the channel coefficient vector between the BS and the kth user and the $n_k(n)$ is the zero-mean additive white Gaussian noise with variance σ_k^2. The transmitted vector $\mathbf{x}(n) \in \mathbb{C}^{M \times 1}$ can be written as

$$\mathbf{x}(n) = \mathbf{W}\mathbf{s}(n) \qquad (2.3.13)$$

where $\mathbf{W} = [\mathbf{w}_1 \cdots \mathbf{w}_K]$, $\mathbf{w}_k \in \mathbb{C}^{M \times 1}$ is the beamforming weight vector for the kth user, $\mathbf{s}(n) = [s_1(n) \cdots s_K(n)]^T$ and $s_k(n)$ is the symbol intended for the kth user. The received signals for all users can be written into the following matrix equation:

$$\mathbf{y}(n) = \mathbf{H}\mathbf{x}(n) + \mathbf{n}(n) \qquad (2.3.14)$$

where $\mathbf{H} = [\mathbf{h}_1 \cdots \mathbf{h}_K]^H$ and assumed to be full row-rank.
$\mathbf{y}(n) = [y_1(n) \cdots y_K(n)]^T$ and $\mathbf{n}(n) = [n_1(n) \cdots n_K(n)]^T$. The SINR of the kth user can be defined as

$$\mathbf{SINR}_k = \frac{[\mathbf{HW}]_{k,k}^2}{\sum_{j \neq k}[\mathbf{HW}]_{k,j}^2 + \sigma_k^2} \qquad (2.3.15)$$

For zero forcing technique, $\mathbf{W}$ is designed such that the interference between the users is forced to zero, i.e., $[\mathbf{HW}]_{k,j} = 0$ when $k \neq j$. In addition, without loss of generality, it is assumed that $[\mathbf{HW}]_{k,k} \geq 0$ for $k = 1, \cdots, K$. If it also assumed that $E\{\mathbf{s}(n)\mathbf{s}^H(n)\} = \mathbf{I}$, then the aforementioned conditions can be stated as:

$$\mathbf{HW} = \mathrm{diag}\{\sqrt{\gamma}\} \qquad (2.3.16)$$

where $\sqrt{\gamma} = [\sqrt{\gamma}_1 \cdots \sqrt{\gamma}_K]^T$ and γ_k/σ_k^2 is the target SINR of the kth user. These restrictions yield a simple beamforming design that decouple the mul-

tiuser channel into K independent sub-channels

$$y_k(n) = \sqrt{\gamma_k}s_k(n) + n_k(n) \tag{2.3.17}$$

Other possible criteria for transmit beamformer design is to use what is known as signal to interference leakage ratio, where the precoder is designed to maximize the power at the receiver while minimizing the interference leakage to all other users in the environment. This is not very attractive criterion, because it does not guarantee a particular performance for every user in terms of BER or SINR. Therefore, a more attractive criterion is to design beamformers while ensuring the SINR achieved for every user is above a certain threshold γ_i, as shown in the following optimization [16, 38].

$$
\begin{aligned}
\min_{\mathbf{w}_i} \quad & \sum_{k=1}^{K} ||\mathbf{w}_i||_2^2 \\
\text{subject to} \quad & \frac{||\mathbf{w}_i^H \mathbf{h}_i||_2^2}{\sum_{k=1,k\neq i}^{K} ||\mathbf{w}_k^H \mathbf{h}_i||_2^2 + \sigma_i^2} \geq \gamma_i \quad i = 1, \cdots, K
\end{aligned}
\tag{2.3.18}
$$

The problem in (2.3.18) is a quadratically constrained non-convex problem. However, this problem can be converted into a semidefinite programming (SDP) problem with Lagrangian relaxation and it can be solved efficiently using convex optimization toolboxes [39–41]. In addition, it has been proven that the relaxed problem always yields an optimal rank-one solution [16].

If the transmission power is very limited or if the channels do not have enough diversity, it may not be possible for all users to attain the target SINR. In this case, other possible criteria that can be used is called SINR balancing. It is a max-min fairness approach where beamformer is designed to maximize the SINR of the worst case user [17, 42–44]. This SINR balancing

technique can be stated as

$$\max_{\mathbf{U},\mathbf{p}} \quad \min_{1 \leq k \leq K} \frac{\text{SINR}_i(\mathbf{U},\mathbf{p})}{\gamma_i}$$

$$\text{subject to} \quad \mathbf{1}^T \mathbf{p} \leq P_{max}$$

$$\mathbf{p} \geq 0 \qquad\qquad\qquad (2.3.19)$$

where $\mathbf{U} = [\mathbf{u}_1 \cdots \mathbf{u}_K]$, $||\mathbf{u}_k||_2 = 1$ and $\mathbf{p} = [p_1 \cdots p_K]^T$. $\mathbf{u}_k \in \mathbb{C}^{M \times 1}$ is the unity norm transmit beamforming weight vector and p_k is the allocated power for the kth user respectively i.e., $\mathbf{w}_k = \mathbf{u}_k \sqrt{p_k}$. In [17], an iterative algorithm has been proposed using uplink-downlink SINR duality to design the beamformers and power allocation.

Joint Transmitter-Receiver Beamforming Design

The techniques described so far considered either one antenna at the transmitter or one antenna at the receiver. A communication link with multiple antennas at the transmitter and the receiver is referred to as a MIMO system [45–47]. When multiple antennas are used at the transmitter as well as at the receiver, it is possible to obtain multiple spatial channels for the users thereby increasing the data rate directly by a factor up to the rank of the MIMO channel matrix [48].

For a point-to-point network, with M_t antennas at the transmitter and M_r antennas at the receiver, the underlying MIMO network is shown in the Figure 2.5, where the received signal $\mathbf{y}(n) \in \mathbb{C}^{M_r \times 1}$ is written as the transmitted signal $\mathbf{x}(n) \in \mathbb{C}^{M_t \times 1}$ multiplied by the MIMO channel matrix $\mathbf{H} \in \mathbb{C}^{M_r \times M_t}$ plus the noise $\mathbf{n}(n) \in \mathbb{C}^{M_r \times 1}$ as

$$\mathbf{y}(n) = \mathbf{H}\mathbf{x}(n) + \mathbf{n}(n). \qquad\qquad (2.3.20)$$
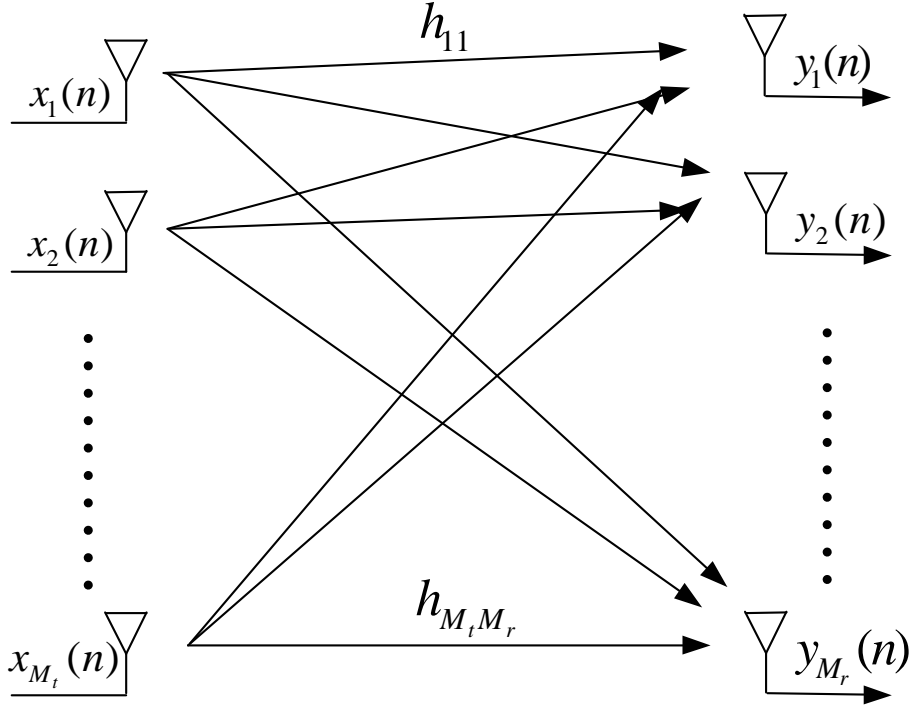
Figure 2.5. A MIMO system with M_t transmit antennas and M_r receive antennas.

It is assumed that the channel matrix $\mathbf{H}$ is known both at the transmitter and at the receiver. The matrix $\mathbf{H}$ can be decomposed using singular value decomposition (SVD) as [49]

$$\mathbf{H}(n) = \tilde{\mathbf{U}}\boldsymbol{\Sigma}\tilde{\mathbf{V}}^H \tag{2.3.21}$$

where $\tilde{\mathbf{U}} \in \mathbb{C}^{M_r \times M_r}$ and $\tilde{\mathbf{V}} \in \mathbb{C}^{M_t \times M_t}$ are unitary matrices, $\boldsymbol{\Sigma} \in \mathbb{R}^{M_r \times M_t}$ is a diagonal matrix of singular values $\{c_i\}$ of $\mathbf{H}$. Let $R_{\mathbf{H}}$ denote the rank of the matrix $\mathbf{H}$. These singular values have the property that $c_i = \sqrt{\lambda_i}$ where λ_i is the ith eigenvalue of $\mathbf{H}\mathbf{H}^H$. The parallel decomposition of the channel is obtained using linear transformation of the input and output signal through transmit precoding and receiver shaping. In transmit precoding the input $\mathbf{x}$ to the antennas is generated as

$$\mathbf{x} = \tilde{\mathbf{V}}\tilde{\mathbf{x}} \tag{2.3.22}$$

where $\tilde{\mathbf{x}}$ is the modulated symbol stream. Similarly, the received signal is shaped as

$$\tilde{\mathbf{y}} = \tilde{\mathbf{U}}^{\mathbf{H}}\mathbf{y} \qquad (2.3.23)$$

as shown in Figure 2.6. The transmit precoding and receiver shaping transform the MIMO channel into $R_{\mathbf{H}}$ parallel single-input single-output (SISO) channels as follows:

$$
\begin{aligned}
\tilde{\mathbf{y}} \; &= \tilde{\mathbf{U}}^{\mathbf{H}}(\mathbf{H}\mathbf{x} + \mathbf{n}) \\
&= \tilde{\mathbf{U}}^{\mathbf{H}}\tilde{\mathbf{U}}\boldsymbol{\Sigma}\tilde{\mathbf{V}}^{\mathbf{H}}\tilde{\mathbf{V}}\tilde{\mathbf{x}} + \tilde{\mathbf{U}}^{\mathbf{H}}\mathbf{n} \\
&= \boldsymbol{\Sigma}\tilde{\mathbf{x}} + \tilde{\mathbf{n}}
\end{aligned}
$$

$$(2.3.24)$$

where $\tilde{\mathbf{n}} = \tilde{\mathbf{U}}^{\mathbf{H}}\mathbf{n}$.
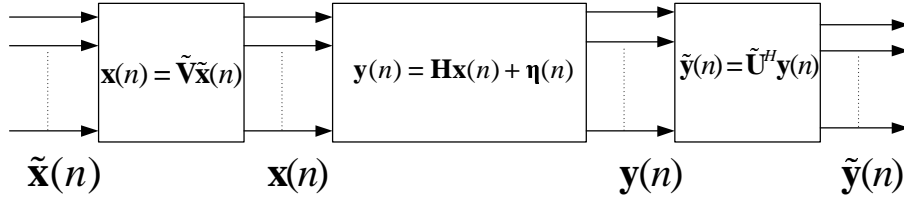
Figure 2.6. Transmit precoding and receiver shaping.

Since the parallel channels do not interfere with each other, the MIMO channel can support up to $R_{\mathbf{H}}$ times the data rate of a SISO channel. The performance of each channel depends on its gain c_i.

The channel capacity of this system is equal to the sum of capacities of each independent parallel sub-channels.

$$C = \underset{p_i:\sum_{i=1}^{R_H} p_i \leq P}{\text{maximize}} \quad \sum_{i=1}^{R_H} B\log_2\left(1 + \frac{v_i^2 P_i}{\sigma_n^2}\right), \qquad (2.3.25)$$
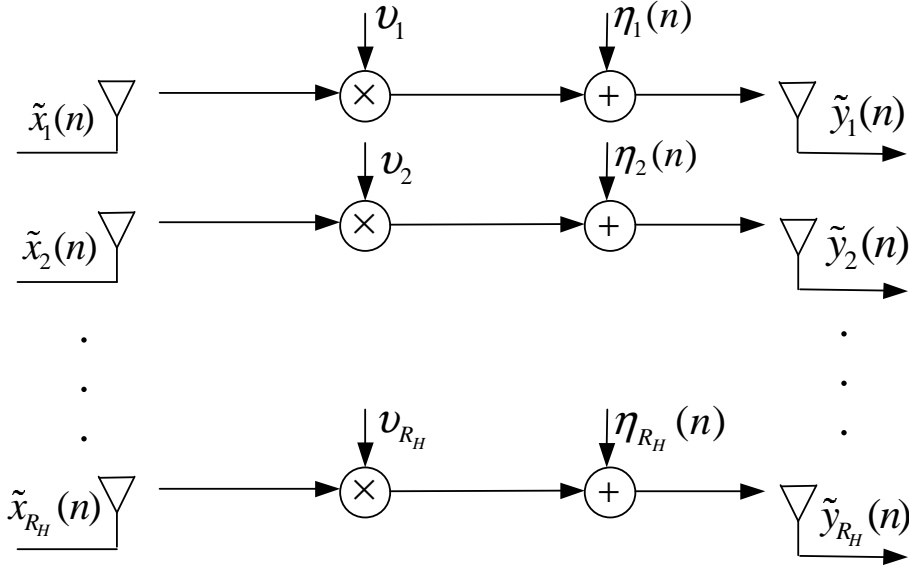
Figure 2.7. Parallel decomposition of the MIMO channel.

where P and P_i are the total transmit power and power allocated to the ith independent channel respectively. B, v_i and σ_n^2 are the bandwidth, i^{th} independent channel gain and the noise power at the receiver respectively. The power allocation problem can be formulated into a convex optimization framework as

$$
\begin{aligned}
\text{maximize} \quad & \sum_{i=1}^{R_H} B\log_2\left(1 + \frac{v_i^2 p_i}{\sigma_n^2}\right) \\
\text{subject to} \quad & \mathbf{1}^T\mathbf{p} \leq P, \\
& p_i \geq 0,
\end{aligned}
\tag{2.3.26}
$$

where $\mathbf{1} \in \mathbb{R}^{R_H \times 1}$ is a vector with all elements equal to one and $\mathbf{p} = [p_1 \; \cdots \; p_{R_H}]^T$. From KarushKuhnTucker (KKT) conditions, the following are obtained [40].

$$
\lambda_i \geq 0 \quad \forall \; i,
\tag{2.3.27}
$$

$$
\lambda_i p_i = 0 \quad \forall \; i,
\tag{2.3.28}
$$

$$
\left(\frac{B}{1 + \frac{v_i^2 p_i}{\sigma_n^2}}\right)\frac{v_i^2}{\sigma_n^2} + \lambda_i = \mu \quad \forall \; i,
\tag{2.3.29}
$$

where λ_i [1]and μ are the Lagrangian variables. From complementary slackness as in (2.3.28) and (2.3.29), the power allocations are obtained as

$$
p_i = \begin{cases} \dfrac{B}{\mu} - \dfrac{\sigma_n^2}{v_i^2}, & \lambda_i = 0; \\[2mm] 0, & \lambda_i \neq 0. \end{cases}
\tag{2.3.30}
$$

The optimal value of μ is given by

$$
\mathbf{1}^T\mathbf{p} = \sum_{i=1}^{R_H} \max\left\{0, \left(\frac{B}{\mu} - \frac{\sigma_n^2}{v_i^2}\right)\right\} = P.
\tag{2.3.31}
$$

This is called the water-filling solution where the water level is equal to $\frac{B}{\mu^*}$ as shown in Figure 2.8. The parameter μ^* is the optimal value of μ.   It is
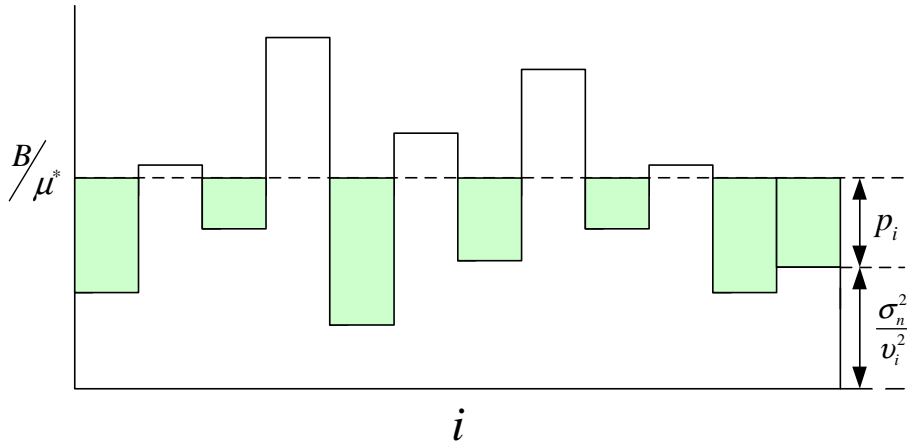
Figure 2.8. Water filling power allocation

also possible to have multiple users with multiple antennas at the transmitter and at the receiver. In this case more sophisticated techniques are required to design the precoders and the receivers. All the techniques described so far considered frequency flat channels. It is also possible to have frequency selective channels. However, throughout the thesis, technology based on

[1]Please note that the same notation has been used for eigenvalues earlier

OFDM, is assumed, hence in each subcarrier, frequency flat channel, either fading or non fading, is considered.

2.4   Relay Techniques

Since communication over a wireless channel is limited by fading, multi-path, path loss and shadowing, direct communication between nodes requires high transmission power and consequently causes increased interference. To achieve reduced power consumption, information can be conveyed to a destination through multiple intermediate nodes. These nodes are known as relays. In this section, an overview of commonly used relay architectures is presented with a focus on regenerative and non regenerative relays as well as one-way and two-way relays.

2.4.1   Non Regenerative Relays

Non regenerative relays perform some linear or non-linear operation in the analog domain that does not modify the information represented by a chosen waveform, before retransmitting the signal. Very simple operations, such as simple amplification, phase rotation, etc are usually performed [50]. One important class of wireless non regenerative relay, is known as amplify-and-forward (AF) relay. Though inherently affected by noise propagation effects, the AF relay is attractive because it provides a reasonable trade-off between performance and practical implementation costs [51].

A typical relay network consists of a source, a destination and N relay nodes is shown in Figure 2.9. During the first phase of transmission, the source transmits the signal $\sqrt{P}s$ to the relays, where s is the information symbol and P is the transmit power. The received signal at the ith relay is given by

$$x_i = \sqrt{P}f_i s + \nu_i, \qquad (2.4.1)$$
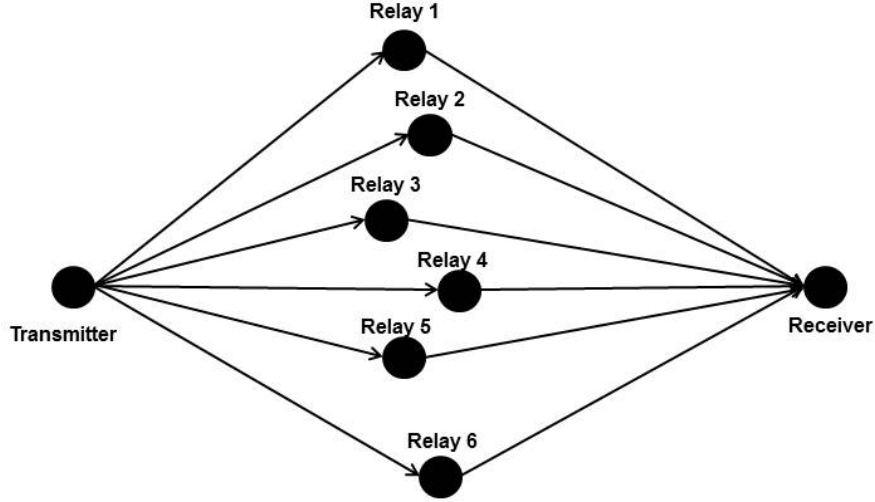
Figure 2.9. Relay network

where ν_i is the noise at the ith relay whose variance σ_ν^2. In the second phase of transmission the ith relay transmits y_i, $y_i = w_i x_i$, where w_i is the complex relay beamforming weight that is used by the ith relay to adjust the phase and the amplitude of the signal. Relay beamforming is a combination of receive and transmit beamforming as the same weights are used for the signal reception and transmission. Moreover, each relay knows only its own received signal and does not know the signals received by other relay nodes. The signal received by the destination is given by:

$$z = \sum_{i=1}^{r} g_i y_i + n. \tag{2.4.2}$$

Using (2.4.1), (2.4.2) is rewritten as

$$z = \sqrt{P} \sum_{i=1}^{r} w_i f_i g_i s + \sum_{i=1}^{r} w_i g_i \nu_i + n. \tag{2.4.3}$$

The following optimization problem can be employed to determine the relay beamforming weights w_i [52].

$$\min \ P_T \quad \text{subject to} \ \ \text{SNR} \geq \gamma \tag{2.4.4}$$

where P_T is the total relay transmit power and γ is the target SINR. In [52], it has been shown that the SNR is given by $(\mathbf{w}^H \mathbf{R} \mathbf{w})/(\sigma_n^2 + \mathbf{w}^H \mathbf{Q} \mathbf{w})$ and $P_T = \mathbf{w}^H \mathbf{D} \mathbf{w}$, where $\mathbf{w} \triangleq [w_1, \cdots, w_N]^T$ is the vector that contains the complex beamforming weights used by the relays, $\mathbf{R}$ is the correlation matrix of the vector $\mathbf{h} = \mathbf{f} \odot \mathbf{g}$, that is $\mathbf{R} \triangleq PE\{\mathbf{h}\mathbf{h}^H\}$, $\mathbf{Q} \triangleq \sigma_\nu^2 E\{\mathbf{g}\mathbf{g}^H\}$ and $\mathbf{D} \triangleq P\mathrm{diag}([E\{|\mathbf{f}_1^2|\}\ E\{|\mathbf{f}_2^2|\}\cdots E\{|\mathbf{f}_r^2|\}]) + \sigma_\nu^2 \mathbf{I}$. Finally, the vectors $\mathbf{f} = [f_1 f_2 \cdots f_r]^T$ and $\mathbf{g} = [g_1 g_2 \cdots g_r]^T$ include the channel coefficients from the transmitter to the relays and the channel coefficients from the relays to the receiver respectively. By defining $\tilde{\mathbf{w}} = \mathbf{D}^{1/2}\mathbf{w}$, the optimization problem in (2.4.4) can be written as:

$$\min_{\tilde{\mathbf{w}}} \quad ||\tilde{\mathbf{w}}||^2$$
$$\text{subject to} \quad \tilde{\mathbf{w}}^H \mathbf{D}^{-1/2}(\mathbf{R} - \gamma\mathbf{Q})\mathbf{D}^{-1/2}\tilde{\mathbf{w}} \geq \gamma\sigma_n^2 \qquad (2.4.5)$$

The inequality constraint in (2.4.5) should be satisfied with equality at the optimum. Hence a closed form solution can be obtained using Lagrangian formulation.

Another possible optimization criterion is to maximize the SNR subject to a power constraint at each relay node as

$$\max_{\mathbf{w}} \quad \frac{\mathbf{w}^H \mathbf{R} \mathbf{w}}{\sigma_n^2 + \mathbf{w}^H \mathbf{Q} \mathbf{w}}$$
$$\text{subject to} \quad \mathbf{D}_{ii}|w_i|^2 \leq P_i, \text{ for } i = 1, 2, \cdots, r \qquad (2.4.6)$$

where P_i is the maximum allowable transmit power of the ith relay and $\mathbf{D}_{ii}$ is the ith diagonal entry of the matrix $\mathbf{D}$. Defining $\mathbf{X} \triangleq \mathbf{w}\mathbf{w}^H$, this

optimization problem can be rewritten as

$$\max_{\mathbf{X}} \quad \frac{\text{Tr}(\mathbf{RX})}{\sigma_n^2 + \text{Tr}(\mathbf{QX})}$$

$$\text{subject to} \quad D_{ii}X_{ii} \leq P_i, \forall\, i = 1, \cdots, r, \text{rank}(\mathbf{X}) = 1,\ \mathbf{X} \succeq 0,$$

$$(2.4.7)$$

Using the idea of semidefinite relaxation [40] and dropping the non convex rank-one constraint, (2.4.7) can be converted into a convex form.

The results of [52] are applicable when the relays are synchronized at the symbol level and when the source-to-relay and relay-to-destination channels are frequency flat. When the channels are frequency selective or the time synchronization between the relays is not perfect, then the signal replicas passed through different relays and/or channel paths will arrive at the destination node with different delays, which will result in ISI. In [53] an asynchronous flat-fading relay network is viewed as an artificial multipath channel and orthogonal frequency division multiplexing (OFDM) scheme is used at the source and destination nodes to deal with this artificial multipath channel.

The above optimization can also be applied in multiuser networks. In [54], a network of relays is used to establish communication between multiple source destination pairs. The relays amplify and adjust the phase of the signal they receive from all transmitting sources by multiplying it with a complex beamforming weight. To obtain optimal beamforming weights, the total relay transmit power is minimized subject to QoS constraints on the received SINRs at the destinations. Using semidefinite relaxation, this power minimization problem is converted into a convex SDP problem and solved.

The authors in [55] considered the same problem but use additional constraints. They enforced the signals received by the destinations to be all in-phase. This turns the aforementioned power minimization problem into a

second-order-cone programming (SOCP) problem, which is computationally less expensive as compared to the SDP problem.

The relay networks described so far, considered only one antenna at the
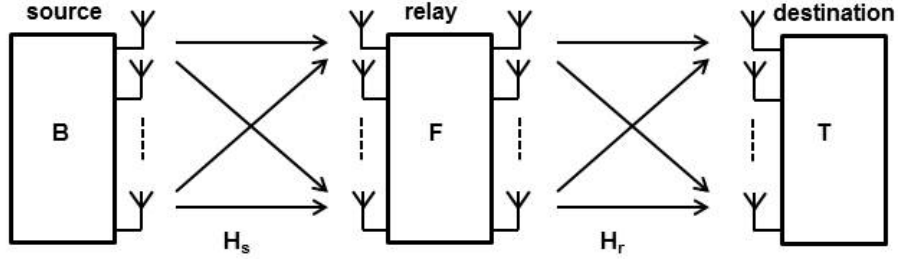
Figure 2.10. A two-hop MIMO AF relay network

transmitters, receivers and the relays. However, as mentioned earlier, the use of multiple antennas at wireless transmitters and receivers has significant advantages for high-rate multimedia transmissions over wireless channels. A block diagram of a one-way two-hop MIMO relay network in which linear processing is employed at all nodes is shown in Figure 2.10. The source vector $\mathbf{s} = [s_1, s_2, \cdots, s_K]^T$ is first linearly processed by a matrix $\mathbf{B} \in \mathbb{C}^{N_S \times K}$ and then transmitted over the source-relay link in the first phase. At the relay, the received signal is first processed by $\mathbf{F} \in \mathbb{C}^{N_R \times N_R}$ and then forwarded to the destination during the second phase [56]. Vector $\mathbf{y}$ at the input of the decision device is given by:

$$\mathbf{y} = \mathbf{THBs} + \mathbf{Tn} \tag{2.4.8}$$

where $\mathbf{T} \in \mathbb{C}^{K \times N_s}$ is the processing matrix at the destination and $\mathbf{H} = \mathbf{H}_r \mathbf{FH}_s$ is the equivalent channel matrix, $\mathbf{H}_s \in \mathbb{C}^{N_R \times N_S}$ and $\mathbf{H}_r \in \mathbb{C}^{N_S \times N_R}$ are the source-relay and the relay-destination channel matrices respectively. In addition, $\mathbf{n} \in \mathbb{C}^{N_s \times 1}$ is a zero-mean complex Gaussian vector whose co-variance matrix is $\rho \mathbf{R}_n$ with $\rho > 0$ accounting for the noise variance over

both links and

$$\mathbf{R}_n = \mathbf{H}_r \mathbf{F} \mathbf{F}^H \mathbf{H}_r^H + \mathbf{I}_{N_S} \qquad (2.4.9)$$

A popular approach in the design of AF MIMO relay systems is to maximize the capacity between the source and destination [57, 58]. Another optimization problem is the minimization of the sum of the MSEs. The MSE matrix is given by :

$$\mathbf{E} = (\mathbf{THB} - \mathbf{I}_k)(\mathbf{THB} - \mathbf{I}_k)^H + \rho \mathbf{TR}_n \mathbf{T}^H \qquad (2.4.10)$$

and the minimization problem is written as:

$$\begin{aligned}
\min_{\mathbf{B},\mathbf{T},\mathbf{F}} \quad & \sum_{k=1}^{K} [\mathbf{E}]_{k,k} \\
\text{subject to} \quad & \text{Tr}\{\mathbf{BB}^H\} \le P_S \\
& \text{Tr}\{\mathbf{F}(\mathbf{H}_s \mathbf{BB}^H \mathbf{H}_s^H + \rho \mathbf{I}_{N_R})\mathbf{F}^H\} \le P_R \qquad (2.4.11)
\end{aligned}$$

where P_S and P_R denote the power available for transmission at the source and the relay, respectively. Since $\mathbf{E}$ is increasing in each argument, the optimal $\mathbf{T}$ in (2.4.11) must be such that each $[\mathbf{E}]_{\mathbf{k},\mathbf{k}}$ is minimized for any given $(\mathbf{B}, \mathbf{F})$ [56]. This is achieved by choosing $\mathbf{T}$ equal to Wiener filter

$$\mathbf{T} = \mathbf{B}^H \mathbf{H}^H (\mathbf{HBB}^H \mathbf{H}^H + \rho \mathbf{R}_n)^{-1} \qquad (2.4.12)$$

The optimal $\mathbf{F}$ and $\mathbf{B}$ match the singular vectors of the corresponding channel matrices. In this way, the strongest spatial channels of the source-relay and relay-destination links are matched together. As a result the overall channel matrix becomes diagonal and the AF MIMO relay system becomes equivalent to a set of parallel single-input single-output (SISO) channels. A similar result was obtained for single hop systems [59]. However, as opposed

to [59], where the optimal power distribution over the parallel SISO channels can be easily found by means of water-filling inspired algorithms, (2.4.11) is not convex. To overcome this problem, [60] proposed an alternative approach in which the power is allocated separately at the source and the relay by means of a water-filling algorithm. This leads to a suboptimal procedure whose solution is shown in [60] to be close to the optimal one.

The block diagram of a linear one-way two hop MIMO system with mul-
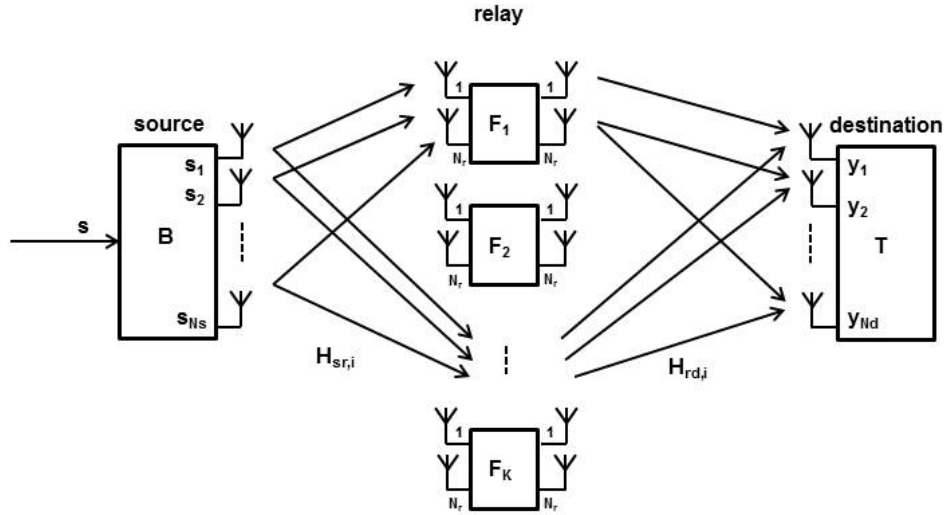
Figure 2.11. A two-hop MIMO AF relay network

tiple parallel relays is shown in Figure 2.11 in which the number of active relays is denoted by K and the matrix that contains the channel gains between the source and the kth relay for $k = 1, 2, \cdots, K$ is called $\mathbf{H}_{sr_i} \in \mathbb{C}^{N_r \times N_s}$. The vector $\mathbf{y} \in \mathbb{C}^{K \times 1}$ can be expressed as in (2.4.8) with the only difference that the matrices $\mathbf{H}_s \in \mathbb{C}^{N_r K \times N_s}$ and $\mathbf{H}_r \in \mathbb{C}^{N_s \times N_r K}$ now take the form $\mathbf{H}_{sr} = [\mathbf{H}_{sr_1}^H \mathbf{H}_{sr_2}^H \cdots \mathbf{H}_{sr_K}^H]$ while $\mathbf{F} \in \mathbb{C}^{N_r K \times N_r K}$ is block diagonal and given by $\mathbf{F} = \text{diag}\{\mathbf{F}_1, \mathbf{F}_2, \cdots, \mathbf{F}_K\}$. In [61], the authors set $\mathbf{W}$ and $\mathbf{B}$ equal to the identity matrix and found the optimal structure of $\mathbf{F}$ by minimizing the sum of the MSEs subject to a global power constraint. The solution is found in closed-form only for the simple case in which relays are equipped with single

antenna i.e. $N_r = 1$, while the multiple antenna case is addressed without imposing any power constraint at the relays. The design of $\mathbf{F}$ that minimizes the total power consumption while satisfying a set of SNR constraints is studied in [62] and a power efficient solution is derived in closed-form after solving a two-step optimization problem. Another optimization problem of the minimization of the sum of the MSEs was studied in [63]. It was found that the optimal $\mathbf{W}$ matrix is the Wiener filter and the diagonalization of the overall channel matrix is achieved up to a unitary matrix. The evaluation of the diagonal matrices requires to solve a non-convex power allocation problem, which is solved using the same arguments outlined before.

2.4.2   Regenerative Relays

Another important relay protocol is regenerative relaying, where the information bits or waveforms are modified and regenerated at the relay prior to transmission towards the destination. The most prominent examples of regenerative relays is the Decode and Forward (DF) relay, that decodes the received signals and forwards the re-encoded signals to the destination node [64,65]. This scheme is attractive when the channel between the source and the relay is of good quality, otherwise the decoding will result into errors due to low SINR and the errors will propagate to the destination. For this reason more complex coding schemes are required at the relays. The earlier implementations were based on convolutional codes [66]. Implementations based on Turbo-codes and low-density parity check (LDPC) codes have also been proposed [67–71].

Optimal precoder design for DF relay systems have been extensively studied based on various optimization criteria. In [72] a two-hop system is considered with and without direct link between the source and the destination. The aim is to find the optimal power allocation for the source and the relay while satisfying specific outage performance. In [73] a distributed sum

transmit power minimization with QoS constraints is considered for a network consisting of a single source and destination communicating through multiple relays. The algorithm in [74] attempted to minimize the uplink sum transmit power with relay selection, where two-stage transmission over parallel channels with user cooperation during the latter stage was used to improve the system performance. The sum transmit power minimization with BER constraints was addressed in [75] for a multi-hop system with one source and destination. The DF relaying was modified in such a way that perfect decoding was not required at each relay. In [76, 77] throughput maximization for DF relay assisted cellular orthogonal frequency-division multiple access (OFDMA) systems has been considered. In [76] the problem of downlink outage probability minimization and throughput maximization have been addressed. The BS is assumed to have local subcarrier CSI knowledge so that it could decide whether it is more beneficial to serve users directly or via relays. A different approach to multi-cell OFDMA throughput maximization was provided in [77] where the problem was approximated to convex and decoupled between the interfering cells via dual decomposition. In [78] relay beamforming for two-hop multi-relay system was considered with multiple antennas at the relay and single-antenna at the source and destination.

2.4.3 Two-way Relaying

The work described so far considered one-way relay networks that cooperate with each other to deliver information symbols from one or more sources to one or several destinations. In two-way relaying schemes, bidirectional connection is established between two transceivers. Two-way relay was first introduced in [79] where an achievable rate region and an outer bound for the case in which nodes operate in full-duplex were obtained. The capacity

and achievable region for two-way relay channels were studied in [80, 81].

The most straightforward approach for two or more transceivers to exchange information is to deploy two successive one-way relaying schemes as shown in Figure 2.12a. This approach demands 4 time slots to accomplish the exchange of symbols s_1, s_2 between the two transceivers, thus it is bandwidth inefficient. To improve efficiency, a time division broadcast (TDBC) approach was introduced in [82] for a single relay where a network coding was introduced to reduce the number of time slots to three as shown in Figure 2.12b. The relaying scheme of [82] requires the relay to decode the signals it receives during the first two time slots and then broadcast the XOR of the decoded signals in the third time slot. Each transceiver can retrieve its signal of interest by performing an XOR on its transmitted symbol and its received signal. In [83] the performance of TDBC was studied in detail and is compared with that of the four time slot scheme. It was concluded that the TDBC approach provides a throughput which is significantly higher than that of the traditional four time slot scheme shown in Figure 2.12a.

In [84, 85] another scheme has been introduced, which is called multiple access broadcast (MABC) scheme as shown in Figure 2.13. The two transceivers transmit their information simultaneously and a relay node retransmits a processed version of its received signal. This was extended to multi-antenna relaying MABC scheme in [85, 86]. A hybrid TDBC-MABC scheme has been proposed in [87]. Designing optimal MABC based relaying schemes for a network with multiple relays has been considered in [88]. More specifically, a wireless network which consists of two transceivers and n_r single-antenna two-way relay nodes was considered. Each relay adjusts the phase and the amplitude of the mixture it receives in the first time slot by multiplying it with a complex weight and then each relay retransmits the so-obtained signal in the second time slot.

During the first time slot, both transceivers simultaneously transmit
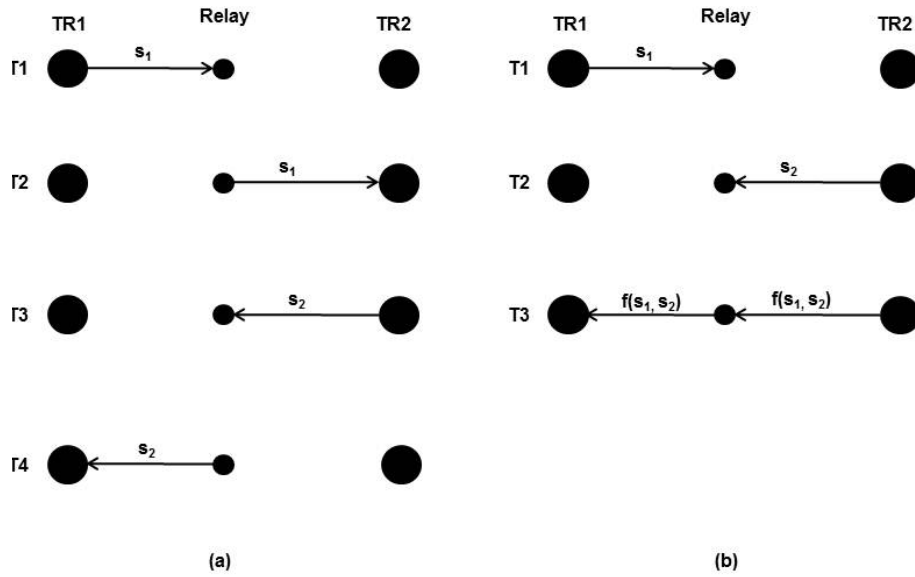
Figure 2.12. (a) Four-time slot two-way relaying scheme. (b) Three-time slot TDBC scheme
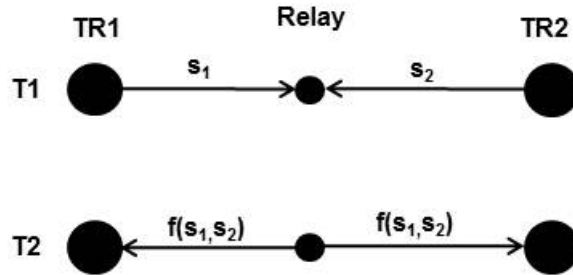
Figure 2.13. Two-time slot MABC approach

their data to the relays. The signals received at the relays $\mathbf{x} \in \mathbb{C}^{n_r \times 1}$ can be represented in a vector form as

$$\mathbf{x} = \sqrt{P_1}\mathbf{f}_1 s_1 + \sqrt{P_2}\mathbf{f}_2 s_2 + \nu \qquad (2.4.13)$$

where, P_1, P_2 are the transmit powers of transceivers TR1 and TR2, respectively and s_1, s_2 are the corresponding transmitted information symbols, $\nu \in \mathbb{C}^{n_r \times 1}$ is a vector consisting of noise at the relays and $\mathbf{f}_1$, $\mathbf{f}_2$ are the vectors of the channel coefficients from the transceivers to the relays. In

the second time slot the ith relay multiplies its received signal by a complex weight w_i^* and broadcasts the signal. The signal vector $\mathbf{t} \in \mathbb{C}^{n_r \times 1}$ transmitted by the relays can be expressed as

$$\mathbf{t} = \mathbf{W}\mathbf{x} \tag{2.4.14}$$

where $\mathbf{W} = \mathrm{diag}([w_1^* w_2^* \cdots w_{n_r}^*])$. The signals y_1 and y_2 received at the two transceivers can be written as:

$$y_1 = \sqrt{P_1}\mathbf{w}^H \mathbf{F}_1 \mathbf{f}_1 s_1 + \sqrt{P_2}\mathbf{w}^H \mathbf{F}_1 \mathbf{f}_2 s_2 + \mathbf{w}^H \mathbf{F}_1 \nu + n_1 \tag{2.4.15}$$

$$y_2 = \sqrt{P_1}\mathbf{w}^H \mathbf{F}_2 \mathbf{f}_1 s_1 + \sqrt{P_2}\mathbf{w}^H \mathbf{F}_2 \mathbf{f}_2 s_2 + \mathbf{w}^H \mathbf{F}_2 \nu + n_2 \tag{2.4.16}$$

where $\mathbf{F}_k = \mathrm{diag}(\mathbf{f}_k)$ for $k = 1, 2$ and $\mathbf{w} = \mathrm{diag}(\mathbf{W})$. The first-term in (2.4.15) depends on the signal s_1 transmitted by TR1 during the first time slot. Also, $\sqrt{P_1}\mathbf{F}_1\mathbf{f}_1$ and $\mathbf{W}$ are known to TR1. Therefore, the first term in (2.4.15) is known at TR1. Hence, this term can be subtracted from y_1 and the residual signal is processed at TR1 to extract the information symbol s_2. Similar procedure is followed by TR2 to extract s_1. The residual signals $\tilde{y}_1$ and $\tilde{y}_2$ can be written as

$$\tilde{y}_1 = \sqrt{P_1}\mathbf{w}^H \mathbf{F}_2 \mathbf{f}_1 s_1 + \mathbf{w}^H \mathbf{F}_1 \nu + n_1 \tag{2.4.17}$$

$$\tilde{y}_2 = \sqrt{P_1}\mathbf{w}^H \mathbf{F}_2 \mathbf{f}_1 s_1 + \mathbf{w}^H \mathbf{F}_2 \nu + n_2 \tag{2.4.18}$$

In order to design $\mathbf{w}$, various optimization criteria can be used. One common approach is to minimize the total transmit power subject to constraints on the received SNRs at the two transceivers. It is proven that this power minimization problem has a unique solution and the optimal beamforming

weight vector (relay coefficients) can be calculated using a gradient-based steepest descent algorithm [88]. Another optimization approach is called SNR balancing technique, which uses a max-min fairness approach, where the smaller of the two transceivers' received SINRs is maximized subject to a constraint on the total transmit power [88]. This problem is shown to have a unique solution which can be obtained through a gradient-based iterative procedure.

The two-way relay channels have been studied jointly with other physical layer transmission techniques, for example OFDM [89, 90] and multiple transmit and/or receive antennas [91, 92]. For the multi-antenna two-way relay channel, the DF relay strategy was studied in [91, 93] and the AF relay strategy or analogue network coding was studied in [94]. A distributed space-time coding strategy was also studied in [92].

2.5   Cooperative Diversity Schemes

One commonly used method for achieving diversity employs multiple antennas at the transmitter and the receiver. However the benefits of the MIMO systems in terms of combating channel fading and increasing system throughput are limited in practical systems mainly due to two reasons [95]. In a collocated MIMO system, due to space limitation, antennas have to be placed close to each other. Thus, radio signals in the collocated antennas experience similar scattering environment and this results in loss of diversity [96, 97]. Also due to the terminal size limitation, the nodes cannot be equipped with many antennas. This issue can be solved using the concept of virtual array of antennas where multiple user terminals share their antennas to receive signals for each other in a coordinated way. However, this approach is not perused in this thesis.

An important problem in cellular networks arises from the fact that mul-

tiple users share common resources and the frequency reuse among adjacent cells brings co-channel interference (CCI) [98]. Cooperative communications in wireless networks aim to combat the aforementioned problems. Two scenarios are considered for cooperative communications. The first one assumes a coordination at the BS side via the exchange of global CSI and possibly exchange of user data information over backhaul links among several cells. Another possibility is cooperation between BS and relays.

In conventional non-cooperative approach, users placed at the edge of the cell suffer from interference. This ICI is treated as noise at the receiver side and is handled by improved point to point communications between the BS and the mobile station [99]. The performance of conventional networks can be significantly improved if joint signal processing is enabled across different base-stations. There are mainly two types of coordination among different BSs. In the first one, all antennas of the relevant multi-cell BSs can be jointly employed for data processing and transmission and all BS are effectively connected to a central processing site. The BS share not only the CSI, but also the full data signals of their respective users. This way of transmission is equivalent to a single giant BS or processing center. The uplink channel can then be modeled as a multiple access channel with multiple transmitters and a single multi-antenna receiver. The downlink channel can be modeled as a broadcast channel with a single multi-antenna transmitter and multiple receivers.

This kind of full BS coordination was investigated in [100–102] for the uplink. The information-theoretic capacity of the uplink of cellular networks with full BS cooperation was performed in [103,104]. In these works, it was shown that with full BS cooperation, the traditional approach of frequency re-use is suboptimal and that full BS cooperation reduces the ICI penalty to zero. A network of M fully connected t-antennas BS can serve a total of Mt terminals in an interference-free manner simultaneously, by employ-
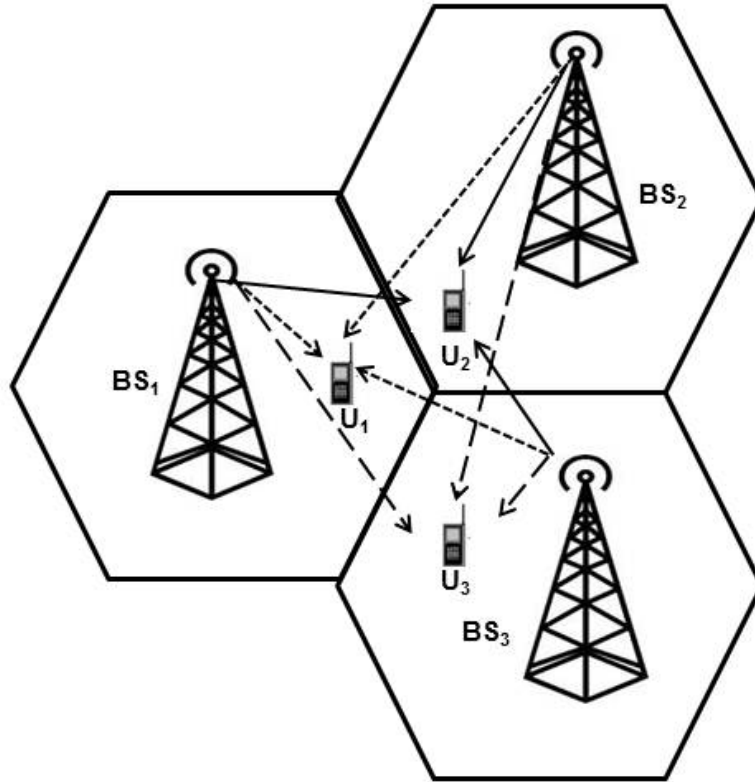
Figure 2.14. Joint Transmission

ing multi-user spatial precoding and decoding techniques similar to the ones used for the MU-MIMO channel [105]. Moreover, the key difference from the single-cell scenario with co-located antennas at one BS are the widely distributed antennas and independent large-scale fading experienced in for each link between a mobile-BS pair and its potential in co-channel cancelation [106–110]. Many coding strategies have been proposed in the literature for this setting [111–115]. This mode of operation is called multi-BS MIMO joint transmission and illustrated in Figure 2.14. The objective is to coordinate the BS transmissions so that the signals from multiple base antennas can be coherently received to improve signal quality. Similar techniques as used in multi-antenna based single BS can be used for decoding optimum beamformers for the transmission and the reception.

The multi-BS MIMO joint processing comes at a cost of signal-level coordination. The data intended for different mobile users belonging to different cells need to be shared among the BSs. The amount of backhaul communications required to achieve joint processing is substantially high. This motivated a second mode of coordination between BS of different cells, namely coordinated beamforming (CB) as illustrated in Figure 2.15. In this mode, a basic level of coordination is employed where only the CSI of the direct and interfering channels are shared among the BSs. There is no need for sharing the information symbols among different BSs. This information allows the transmission strategies across different cells to adapt to the channel state jointly. Transmission strategies include scheduling, power control, beamforming and advanced coding methods as well.

Coordinated power control and scheduling across the multiple BSs in order to adapt to the channel condition of the entire network brings improvement over traditional per cell power control. When the BSs are equipped with multiple antennas, the availability of additional spatial dimensions allows the coordination of beamforming vector across the BSs. This idea has been explored in [116–120].

An optimization problem associated with coordinated beamforming is the minimization of the transmit power across the BSs subject to a set of SINR constraints. This optimization is applicable to constant bit-rate applications with fixed QoS constraints.

Let $\mathbf{w}_{l,k}$ denote the downlink transmit beamforming vector for the kth user in the lth cell. For users with single antenna, the downlink SINR for the kth user in the lth cell can be expressed as:

$$\mathbf{\Gamma} = \frac{|\mathbf{h}_{l,l,k}^{\dagger}\mathbf{w}_{l,k}|^2}{\sum_{n \neq k}|\mathbf{h}_{l,l,k}^{\dagger}\mathbf{w}_{l,n}|^2 + \sum_{j \neq l}|\mathbf{h}_{j,l,k}^{\dagger}\mathbf{w}_{j,n}|^2 + 1} \qquad (2.5.1)$$
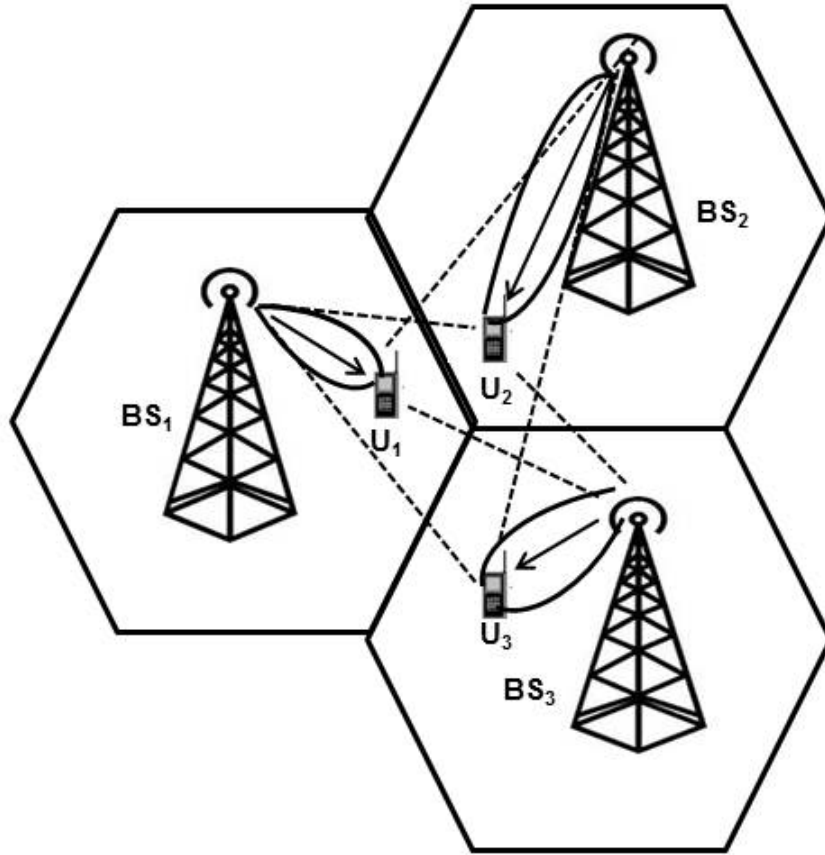
Figure 2.15. Coordinated Beamforming

where $\mathbf{h}_{j,l,k}$ is the vector channel from the jth BS to the kth user in the lth cell. Let $\gamma_{l,k}$ be the SINR target for the kth user in the lth cell. Thus, the total downlink transmit power minimization problem is formulated as [118]:

$$\min \quad \sum_{l=1}^{M}\sum_{k=1}^{K}||\mathbf{w}_{l,k}||^2$$
$$\text{subject to} \quad \Gamma_{l,k} \leq \gamma_{l,k}, \ \forall \, l = 1, \cdots, M, \ k = 1, \cdots, K$$

$$(2.5.2)$$

The minimization is over $\mathbf{w}_{l,k}$ which implies both the power and beamformers are optimized. When the number of antennas at the BS is larger than the users in each cell, the BS has adequate spatial dimension for inter-

ference mitigation. In this case, a per-cell ZF solution is also possible [121].

The transmit downlink problem for the single-cell system was first considered in [122], where an iterative algorithm was proposed for optimizing the beamforming vectors and power allocations. The key idea is to consider the virtual uplink network where it is easy to find the optimal uplink receiver beamformers and then to iterate between the beamformer update step and the power update step to satisfy the target SINRs. The optimality of this algorithm was established for single-cell network using several convex optimization based techniques [17, 123–126]. Moreover, in [127] it was shown that the uplink-downlink duality is an example of Lagrangian duality in optimization.

The duality result is also used to solve the optimization problem of minimizing the weighted per-BS or per-antenna powers. The objective is to minimize the weighted sum power. The weights are adjusted to tradeoff the powers among different BS antennas [127, 128]. When the users are equipped with multiple receive antennas duality holds as well [129]. However, the iterative updating of transmit, receiver beamformer and the power is no longer guaranteed to provide globally optimal solution. The single-cell uplink-downlink duality has been extended for multi-cell and multi-user setting in [118].

Regardless of potential advantages of coordinated multi-cell processing, there are many practical issues that need to be addressed carefully. First of all, downlink MIMO cooperation across multiple BSs requires tight synchronization so that there is ideally no carrier frequency offset between the local oscillators and the BSs. Moreover, sufficient resources must be allocated to pilot signals to ensure reliable estimation of the channel state for network coordination with spatially distributed BSs. Finally cooperative networks need an enhanced backhaul network connecting BSs with each other and

with a central processor. Hence, this is a challenging area for further research.

2.6   Conclusion

As discussed in this chapter, coordinated processing of signals and spatial diversity techniques have significant potential to meet ever increasing demand for radio spectrum. The techniques proposed in this thesis aim to address this issue through BS and relay cooperation and spatial diversity techniques.

Chapter 3

WIRELESS PEER-TO-PEER RELAY NETWORKS

Ad hoc networking capabilities are expected to become key components of overall next-generation wireless networks, because of their flexibility and the absence of a fixed infrastructure. This distinguishes them from traditional mobile wireless networks and makes them interesting for a wide range of applications, such as vehicle-to-vehicle communications, emergency services etc. These types of networks may also provide solutions on future multi-hop wireless communication to a subscriber over wireless pico BSs or relay stations that are connected wirelessly among each other having access to the wired network by some access points. The simplest ad hoc networks that will be demonstrated in this chapter are peer-to-peer networks formed by a set of stations that communicate through intermediate nodes. Hence, our focus in this chapter will be on fixed wireless peer-to-peer relay networks. Initially, beamforming design and power allocation techniques are proposed for peer-to-peer wireless networks consisting of multiple source-destination pairs communicating through a number of amplify-and-forward (AF) one way relay nodes. This framework is then extended to AF two-way-relay nodes. An underlay cognitive peer-to-peer relay network is also studied.

3.1   Introduction

The wireless peer-to-peer relay network, despite its advantages, such as enhancing link reliability, coverage and combat fading, has a problem of multiple access interference due to multiple transmission links in the network. This chapter proposes various optimization techniques to minimize interference and to enhance power efficiency. In order to maintain fairness among users, the resources are allocated such that each user's SINR is balanced and maximized. The associated optimization problem is commonly referred to as "SINR Balancing". The SINR balancing technique has been introduced for a downlink network without relays in [17], where the principle of "downlink-uplink duality" has been applied to transform the original downlink problem into an equivalent uplink problem that is more easily solved. Another approach to solve this problem is to apply conic convex optimization [124].

One can find a number of articles on AF relay beamforming. In [130] and [131] a distributed beamforming technique has been developed for a relay network which consists of a number of source-destinations pairs and relay nodes. The total relay transmit power dissipated by all relays is minimized, while the quality of the services at all destinations is guaranteed to be above certain pre-defined thresholds. Many algorithms have also been developed for the case of a single source-destination pair [52] and [132]. These techniques maximize the receiver SNR under the relay transmitted power-constraints. In [133] and [130], algorithms have been formulated such that the relay power is minimized to satisfy SINR constraints for each user. Two-way AF relays have also been studied in [85] and [134]. However, the SNR balancing problem for two-way relay network has only been studied in [88], where one pair of terminals is considered. The solution is based on the assumption that the phase of the beamforming vector has to match the aggregated phase of the channel coefficients from the relay to the two

transceivers.

In this chapter peer-to-peer networks consisting of a number of source destination pairs are considered. Due to poor quality of the channel between the source and the destination no direct communication is assumed. For this purpose, the destination nodes cooperate with a number of single or multiple antenna AF relay nodes, which receive information transmitted by the source nodes in the first phase and retransmit it to the destination nodes in the second phase. By employing multiple relays, spatial diversity gains are exploited without the need of multiple antennas at each node. In addition, the focus is on AF half duplex relays because they have many advantages from an implementation viewpoint. The same set up but for a two-way multi-user AF relay is also investigated.

Later in the chapter, motivated by the observation that the licensing based conventional radio spectrum allocation is inefficient as the spectrum is under-utilized most of the time at various geographical locations, an underlay CR approach is investigated for relay networks. The SUs access the spectrum occupied by the PUs, given that the interference they cause to the PUs is less than a specific threshold such that the Quality-of-Service (QoS) of the PU can be ensured. More specifically, an AF based relay network with multiple peer-to-peer cognitive radio users is considered. Both the transceivers and the relay nodes are cognitive terminals. It is assumed that all transceivers are equipped with single antenna, while relay nodes consist of multiple antennas. In the context of CR networks, cooperative transmission between SUs aims to increase the secondary throughput, while ensuring the PU terminals are not affected harmfully. Both a one-way and a two-way AF relaying scenario has been considered. The aim is to maximize the worst-case SU SINR subject to satisfying the total transmission power and ensuring that the interference leakage to PUs is below a specific threshold. Similar work for cognitive radio networks has been proposed in [135],

where the relay transceiver matrices have been designed to achieve a specific target SINR for every users while ensuring the interference leakage towards a primary receiver is below a certain threshold. These methods have an inherent problem that for a limited transmission power, the SINR targets may not be achieved all the time and the optimization problem may turn out to be infeasible. As a remedy to this kind of problems and also to introduce fairness among users, the approach considered in this chapter is based on an SINR balancing criterion under a CR environment.

3.2   An SINR Balancing Based Multiuser Relaying Scheme

In this section, a relay network with multiple users is considered. The relay coefficients and the transmission powers are optimized to maximize the worst-case user SINR. This problem is solved through an iterative procedure that uses SDP and GP along with bisection search method. The performance of the proposed scheme in terms of the mean SINR is evaluated for different number of relays and transmission power at the relays.

3.2.1   System Model

As shown in Figure 3.1, a network with d source-destination pairs and R relays is considered. Each node in the network has been assumed to have a single antenna to transmit and receive data. The ith relay received signal x_i is given by

$$x_i = \sum_{p=1}^{d} f_{ip} s_p + \nu_i, \qquad i = 1, \cdots, R \tag{3.2.1}$$

where s_p is the information symbol transmitted by the pth source, ν_i is the additive zero-mean noise at the ith relay node, f_{ip} represents the channel coefficient from the pth source to the ith relay and g_{ip} stands for the channel coefficient from the ith relay to the pth destination. The following assumptions are made:
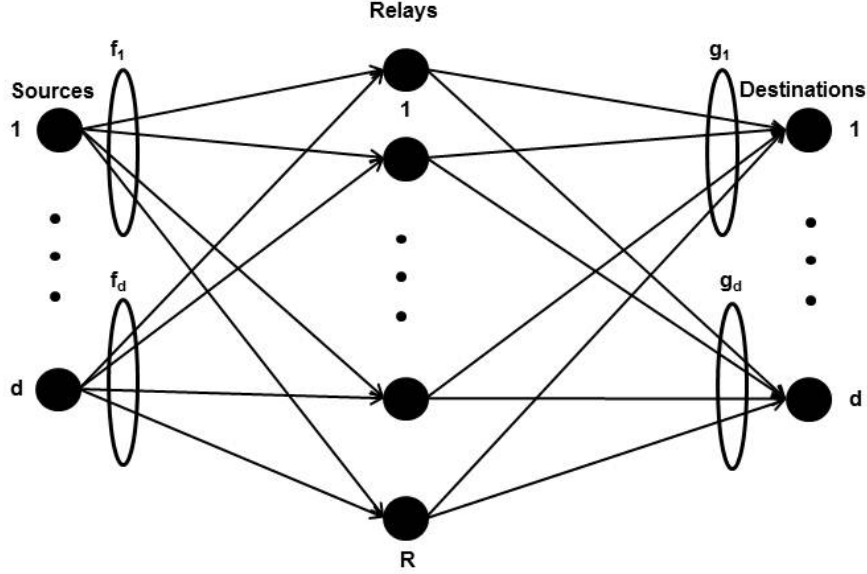
Figure 3.1. A relay network of d pairs of nodes and R relay nodes.

$A1$: The relay noises are assumed to be spatially white, i.e., $E\{\nu_i\nu_{i'}^*\} = \sigma_\nu^2\delta_{ii'}$, where σ_ν^2 is the relay noise variance and $\delta_{ii'}$ is Kronecker's delta function.

$A2$: The transmission power of the pth source is $E\{|s_p|^2\} = P_p$ for $p = 1, \cdots, d$.

$A3$: The information symbols transmitted by the different sources are un-correlated,i.e., $E\{s_p s_q^*\} = P_p\delta_{pq}$.

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

The above assumptions hold for the rest of Chapter 3 as well. Using vector notations, (3.2.1) can be rewritten as

$$\mathbf{x} = \sum_{p=1}^d \mathbf{f}_p s_p + \boldsymbol{\nu} \qquad\qquad (3.2.2)$$

where the following definitions are used: $\mathbf{x} \triangleq [x_1\ x_2\ \cdots\ x_R]^T$, $\boldsymbol{\nu} \triangleq [\nu_1\ \nu_2\ \cdots\ \nu_R]^T$, $\mathbf{f}_p \triangleq [f_{1p}\ f_{2p}\ \cdots\ f_{Rp}]^T$. The ith relay multiplies its

received signal by a complex weight coefficient w_i^*. As a result, the vector of the signals transmitted by the relays is given by

$$\mathbf{t} = \mathbf{W}^H \mathbf{x} \tag{3.2.3}$$

where $\mathbf{W} \triangleq \mathrm{diag}([w_1, w_2, \cdots, w_R])$ and $\mathbf{t}$ is an $R \times 1$ vector whose ith entry is the signal transmitted by the ith relay. The notation $\mathrm{diag}\{a_1, \cdots, a_M\}$ denotes a diagonal matrix whose diagonal elements are $a_1, \cdots, a_M$.

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

$$
\begin{aligned}
y_k &= \mathbf{g}_k^T \mathbf{t} + n_k = \mathbf{g}_k^T \mathbf{W}^H \sum_{p=1}^{d} \mathbf{f}_p s_p + \mathbf{g}_k^T \mathbf{W}^H \boldsymbol{\nu} + n_k \\
&= \mathbf{g}_k^T \mathbf{W}^H \mathbf{f}_k s_k + \mathbf{g}_k^T \mathbf{W}^H \sum_{p=1, p \neq k}^{d} \mathbf{f}_p s_p + \mathbf{g}_k^T \mathbf{W}^H \boldsymbol{\nu} + n_k
\end{aligned}
\tag{3.2.4}
$$

where n_k is the zero-mean noise at the kth destination with a variance of σ_n^2. It is also assumed 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 $\boldsymbol{\nu}$, kth destination noise n_k are statistically independent.

3.2.2   SINR Balancing Technique

The goal is to optimally calculate the relay weights $\{w_i\}_{i=1}^{R}$ with the aim of maximizing the minimum SINR perceived by the user subject to the relay power constraint. Hence, the optimization problem is stated as:

$$
\begin{aligned}
&\max_{\mathbf{w}} \min_{k} \ \mathrm{SINR}_k \\
&\text{s.t.} \quad P_R \leq P_{Rmax}, \quad k = 1, \cdots, d
\end{aligned}
\tag{3.2.5}
$$

where P_R and P_{Rmax} are, respectively, the actual and maximum allowable total transmit power of the relays and SINR_k is the SINR at the kth destination which is defined as

$$\text{SINR}_k = \frac{P_s^k}{P_i^k + P_n^k} \tag{3.2.6}$$

Here P_s^k, P_i^k, and P_n^k are the desired signal component power, the interference power, and the noise power at the kth destination, respectively. Using (3.2.3), the total transmit power at the relays is expressed as

$$P_R = \text{E}\{\mathbf{t}^H\mathbf{t}\} = \text{E}\{\mathbf{x}^H\mathbf{W}\mathbf{W}^H\mathbf{x}\} = \text{Tr}(\mathbf{W}^H\text{E}\{\mathbf{x}\mathbf{x}^H\}\mathbf{W}) \tag{3.2.7}$$

If the correlation matrix of the relay received signals is denoted by $\mathbf{R_x} \triangleq \text{E}\{\mathbf{x}\mathbf{x}^H\}$, the total transmit power can be written as

$$P_R = \text{Tr}(\mathbf{W}^H\mathbf{R_x}\mathbf{W}) = \sum_{r=1}^{R}|w_r|^2[\mathbf{R}_x]_{i,i} = \mathbf{w}^H\mathbf{D}\mathbf{w} \tag{3.2.8}$$

where $\mathbf{D} \triangleq \text{diag}([\mathbf{R}_x]_{1,1}, [\mathbf{R}_x]_{2,2}, \cdots, [\mathbf{R}_x]_{R,R})$ and $\mathbf{w} \triangleq \text{diag}(\mathbf{W})$. Note that using (3.2.2) as well as Assumptions A1-A4, the matrix $\mathbf{R}_x$ is expressed as

$$
\begin{aligned}
\mathbf{R}_x &= \sum_{p,q=1}^{d} \text{E}\{\mathbf{f}_p\mathbf{f}_q^H\}\text{E}\{s_p s_q^*\} + \sigma_\nu^2\mathbf{I} \\
&= \sum_{p=1}^{d} P_p\text{E}\{\mathbf{f}_p\mathbf{f}_p^H\} + \sigma_\nu^2\mathbf{I} = \sum_{p=1}^{d} P_p\mathbf{R}_f^p + \sigma_\nu^2\mathbf{I}
\end{aligned}
\tag{3.2.9}
$$

where $\mathbf{R}_f^p \triangleq \text{E}\{\mathbf{f}_p\mathbf{f}_p^H\}$. It follows from (3.2.9) that the relay transmit power P_R depends not only on the variances of the source-relay channel coefficients but also on the relay noise powers. To express the SINR at the kth destination, the 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 $\{w_i\}_{i=1}^{R}$ are derived. Using (3.2.4) and Assumption $A5$, the noise power at the kth destination

can be written as

$$
\begin{aligned}
P_n^k &= \mathrm{E}\{\mathbf{v}^H \mathbf{W} \mathbf{g}_k^* \mathbf{g}_k^T \mathbf{W}^H \mathbf{v}\} + \sigma_n^2 \\
&= \mathrm{Tr}\{\mathbf{W}^H \mathrm{E}\{\mathbf{v}\mathbf{v}^H\} \mathbf{W} \mathrm{E}\{\mathbf{g}_k^* \mathbf{g}_k^T\}\} + \sigma_n^2 \\
&= \sigma_\nu^2 Tr\{\mathbf{W}^H \mathbf{R}_g^k \mathbf{W}\} + \sigma_n^2
\end{aligned}
\tag{3.2.10}
$$

where $\mathbf{R}_g^k = \mathrm{E}\{\mathbf{g}_k \mathbf{g}_k^H\}$. As a result, the noise power P_n^k is given by

$$
P_n^k = \sigma_\nu^2 \sum_{r=1}^R |w_r|^2 [\mathbf{R}_g^k]_{ii} + \sigma_n^2 = \mathbf{w}^H \mathbf{D}_k \mathbf{w} + \sigma_n^2
\tag{3.2.11}
$$

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 obtained as

$$
\begin{aligned}
P_s^k &= \mathrm{E}\{\mathbf{g}_k^T \mathbf{W}^H \mathbf{f}_k \mathbf{f}_k^H \mathbf{W} \mathbf{g}_k^*\} \mathrm{E}\{|s_k|^2\} \\
&= P_k \mathrm{E}\{\mathbf{w}^H \mathrm{diag}(\mathbf{g}_k) \mathbf{f}_k \mathbf{f}_k^H \mathrm{diag}(\mathbf{g}_k^*) \mathbf{w}\} \\
&= P_k \mathrm{E}\{\mathbf{w}^H (\mathbf{g}_k \odot \mathbf{f}_k)(\mathbf{f}_k^H \odot \mathbf{g}_k^H) \mathbf{w}\} \\
&= P_k \mathbf{w}^H \mathrm{E}\{\mathbf{h}_k \mathbf{h}_k^H\} \mathbf{w} = P_k \mathbf{w}^H \mathbf{R}_h^k \mathbf{w}
\end{aligned}
\tag{3.2.12}
$$

where $\odot$ stands for Schur-Hadamard (element-wise) multiplication, $\mathbf{h}_k \triangleq (\mathbf{g}_k \odot \mathbf{f}_k) = [f_{1k}g_{1k}\ f_{2k}g_{2k} \cdots f_{Rk}g_{Rk}]^T$, $\mathbf{R}_h^k \triangleq \mathrm{E}\{\mathbf{h}_k \mathbf{h}_k^H\}$. The vector $\mathbf{h}_k$ contains the total path coefficients between the kth source and its corresponding destination via different relays. Using (3.2.4), the interference power at the kth destination is given by

$$
\begin{aligned}
P_i^k &= \mathrm{E}\left\{ \mathbf{g}_k^T \mathbf{W}^H \left(\sum_{p,q \epsilon D_k} \mathbf{f}_p \mathbf{f}_q^H s_p s_q^* \right) \mathbf{W} \mathbf{g}_k^* \right\} \\
&= \mathrm{E}\left\{ \mathbf{w}^H \mathrm{diag}(\mathbf{g}_k) \left(\sum_{p \epsilon D_k} P_p \mathbf{f}_p \mathbf{f}_p^H \right) \mathrm{diag}(\mathbf{g}_k^*) \mathbf{w} \right\} \\
&= \mathrm{E}\left\{ \mathbf{w}^H \left(\sum_{p \epsilon D_k} P_p (\mathbf{g}_k \odot \mathbf{f}_k)(\mathbf{g}_k^H \odot \mathbf{f}_k^H) \right) \mathbf{w} \right\}
\end{aligned}
\tag{3.2.13}
$$

$$= \mathbf{w}^H \sum_{p \epsilon D_k} P_p \mathbf{Q}_k \mathbf{w}$$

where $D_k = \{1, 2, \ldots, d\} - \{k\}$ and $\mathbf{h}_k^p$, $\mathbf{Q}_k$ are defined as $\mathbf{h}_k^p \triangleq \mathbf{g}_k \odot \mathbf{f}_p$,

$$\mathbf{Q}_k \triangleq \sum_{p \epsilon D_k} \mathbf{h}_k^p (\mathbf{h}_k^p)^H.$$

Summarizing (3.2.6), (3.2.8), (3.2.11), (3.2.12) and (3.2.13), the optimization problem in (3.2.5) can be rewritten as

$$\max_{\mathbf{w}} \min_{k} \frac{\mathbf{w}^H P_k \mathbf{R}_h^k \mathbf{w}}{\mathbf{w}^H (\sum_{p \epsilon D_k} P_p \mathbf{Q}_k + \mathbf{D}_k) \mathbf{w} + \sigma_n^2} \tag{3.2.14}$$

$$\text{s.t } \mathbf{w}^H \mathbf{D} \mathbf{w} \leq P_{Rmax} \quad \text{for} \quad k = 1, 2, \cdots, d.$$

Let us define $\mathbf{X} \triangleq \mathbf{w}\mathbf{w}^H$. In this case the problem in (3.2.14) becomes

$$\max_{\mathbf{X}} \min_{k} \frac{\mathrm{Tr}(P_k \mathbf{R}_h^k \mathbf{X})}{\mathrm{Tr}((\sum_{p \epsilon D_k} P_p \mathbf{Q}_k + \mathbf{D}_k) \mathbf{X}) + \sigma_n^2}$$

$$\text{s.t. } \mathrm{Tr}(\mathbf{D}\mathbf{X}) \leq P_{Rmax} \quad \text{for} \quad k = 1, 2, \cdots, d \tag{3.2.15}$$

$$\mathrm{Rank}(\mathbf{X}) = 1, \ \mathbf{X} \succeq 0$$

where $\mathbf{X} \succeq 0$ means that $\mathbf{X}$ is a positive semidefinite matrix. Following the semidefinite relaxation approach and dropping the non-convex rank-one constraint, the optimization problem in (3.2.15) is rewritten as

$$\max_{\mathbf{X}, t} \ \xi$$

$$\text{s.t. } \mathrm{Tr}(\mathbf{X}(P_k \mathbf{R}_h^k - \xi(\sum_{p \epsilon D_k} P_p \mathbf{Q}_k + \mathbf{D}_k))) \geq \sigma_n^2 \xi, \tag{3.2.16}$$

$$\mathrm{Tr}(\mathbf{D}\mathbf{X}) \leq P_{Rmax} \text{ for } k = 1, 2, \cdots, d; \ \mathbf{X} \succeq 0$$

For any given value of ξ, the feasible set in (3.2.16) is convex and as a result the optimization problem in (3.2.16) is quasiconvex.

Due to the relaxation, the matrix $\mathbf{X}_{opt}$ obtained by solving the optimization problem in (3.2.16) will not be of rank one in general. If $\mathbf{X}_{opt}$ happens to be rank one, then its principal component will be the optimal solution to

the original problem in (3.2.15). Otherwise, one has to resort to randomization techniques developed in [136] to obtain a rank-one solution from $\mathbf{X}_{opt}$.

Using the bisection method and solving (3.2.16), the maximum possible value of ξ that satisfies the constraints in (3.2.16) and the corresponding value of $\mathbf{w}$ are obtained. Let ξ_{max} be the maximum possible value of ξ. Based on [40], the following assumption is derived. If, for any given value of ξ, the convex feasibility problem

$$\text{find } \mathbf{X}$$

$$\text{s.t. } \text{Tr}(\mathbf{X}(P_k\mathbf{R}_h^k - \xi(\sum_{p \epsilon D_k} P_p\mathbf{Q}_k + \mathbf{D}_k))) \geq \sigma_n^2\xi, \qquad (3.2.17)$$

$$\mathbf{w}^H\mathbf{D}\mathbf{w} \leq P_{Rmax} \text{ for } k = 1, 2, \cdots, d; \ \mathbf{X} \succeq 0$$

is feasible, then $\xi_{max} \geq \xi$. If (3.2.17) is not feasible, then respectively $\xi_{max} < \xi$. Therefore, one can check whether the optimal value ξ_{max} of the quasiconvex optimization problem in (3.2.16) is smaller than or greater than a given value ξ by solving the convex feasibility problem (3.2.17).

Based on this observation, the quasiconvex optimization problem (3.2.16) can be solved by using a bisection algorithm, where (3.2.17) has to be solved at each step. We begin with an interval $[\xi_l \quad \xi_u]$ known to contain the optimal value ξ_{max}. The convex feasibility problem can then be solved at its midpoint $\xi = (\xi_l + \xi_u)/2$, to determine whether the optimal value is larger or smaller than ξ. If (3.2.17) is feasible for this value of ξ, then $\xi_l = \xi$ is set, otherwise $\xi_u = \xi$ is chosen and the convex feasibility problem in (3.2.17) is solved again. This procedure is repeated until the difference between ξ_u and ξ_l is smaller than some preselected threshold δ.

The optimization in (3.2.17) maximises the worst case user SINR. If the number of relays is substantially higher than the number of users, all users' SINR will tend to be equal to the worst case user SINR t_{max}. However, in general, this optimization at the relay will not be able to ensure SINR bal-

ancing. This is because, unlike transmit beamforming techniques, the relay transceiver is unable to control the power usage for each user separately at the relays. Therefore, in order to balance the SINR of all users, the transmission power should be controlled at the source level using the following GP approach.

The optimization of SINR with regards to the power at the transmitters P_d is considered, when the beamforming vector at the relay is fixed. In this case, the following optimization problem has to be solved:

$$\max_{\mathbf{p}} \min_{k} \ \text{SINR}_k$$
$$\text{s.t.} \quad \sum_{k=1}^{d} P_{dk} \leq P_{dmax} \tag{3.2.18}$$

where P_{dk} and P_{dmax} are, respectively, the actual and maximum allowable total transmit power at the transmitters and $\mathbf{p} = [p_{1k} \ p_{2k} \cdots p_{dk}]$.

Using (3.2.6), (3.2.12), (3.2.13), and considering that the relay weights $\{w_i\}_{i=1}^{R}$ are fixed, the optimization problem in (3.2.18) becomes:

$$\max_{\mathbf{p}} \min_{k} \frac{P_k \mathbf{G}_h^k}{\sum_{p \in D_k} P_p \mathbf{E}_k + \mathbf{N}_k + \sigma_n^2}$$
$$\text{s.t.} \ \sum_{k=1}^{d} P_{dk} \leq \text{P}_{dmax}.$$

where $\mathbf{G}_h^k \triangleq \text{Tr}(\mathbf{R}_h^k \mathbf{X})$, $\mathbf{E}_k \triangleq \text{Tr}(\mathbf{Q}_k \mathbf{X})$ and $\mathbf{N}_k \triangleq \text{Tr}(\mathbf{D}_k \mathbf{X})$.

The above problem can be rewritten as:

$$\max_{\mathbf{p_d}, \widetilde{t}} \ \widetilde{t}$$
$$\text{s.t.} \ \mathbf{E}_k \sum_{p \in D_k} P_p P_k^{-1} \widetilde{t} + \sigma_n^2 \mathbf{N}_k P_k^{-1} \widetilde{t} + P_k^{-1} \widetilde{t} < \mathbf{G}_h^k \tag{3.2.19}$$
$$\sum_{k=1}^{d} P_{dk} \leq P_{dmax}.$$

The problem (3.2.19) is convex and belongs to the class of GP [137]. As a

Table 3.1. Algorithmic solution of the SINR balancing problem

1) **Initialize** $\mathbf{P}_d = \mathbf{P}_d^{(0)}$
2) **Repeat**
3)   Solve the problem (3.2.17) of maximizing the minimum SINR,
       with fixed $\mathbf{P}_d$ and obtain an updated $\mathbf{X}$.
4)   Solve the SINR balancing problem (3.2.19) with fixed $\mathbf{X}$
       and obtain an updated $\mathbf{P}_d$.
5) **Until** the SINR converges.

result, it can be efficiently solved using numerical methods such as cvx [138].

Hence, the iterative SINR-balancing algorithm is described as follows:

3.2.3   Simulation Results

In order to assess the performance of the proposed scheme three numerical

examples are considered. A network of two source-destination pairs is as-

sumed. The noise power at the relays and at the receivers are assumed to be

equal to each other and the initial signal power at each source is equal to 1.

The channel coefficients between the sources and the relays as well as those

between the relays and the receivers have been generated using zero-mean

circularly symmetric i.i.d Gaussian random variables.

In the first case, a network with $R = 3$ relays, 2 users and maximum

allowable relay transmit power $P_{max} = 2$ is considered. Figure 3.2 displays

the histogram of the achievable SINR. In Figure 3.3 the maximum allowable

transmit power at the relays versus the mean SINR is plotted. The SINR

shown in Figure 3.2 is the worst case user SINR and not the balanced SINR.

This means that the simulations have been performed without controlling

the power usage at the transmitters. In Figure 3.4 the same simulation has

been performed but this time the balanced SINR is calculated. The maxi-

mum allowable power at the transmitters is $P_d = 2$. From these two Figures,

it can be noticed that the performance of the users is improved as the maxi-
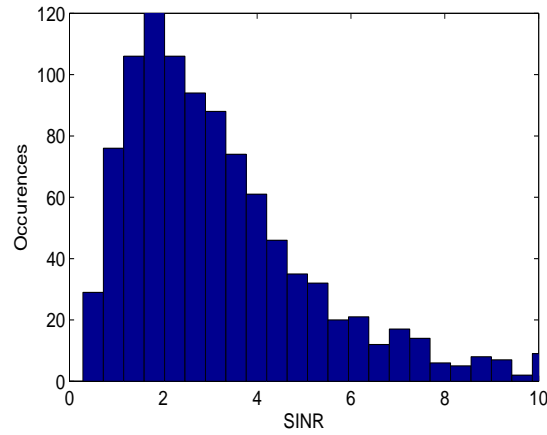
Figure 3.2. SINR histogram. SINR is shown in linear scale
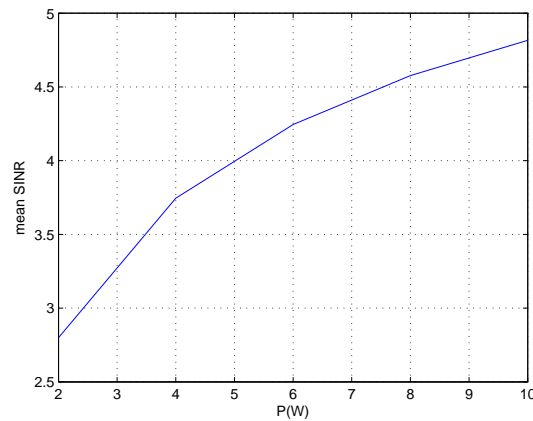
Figure 3.3. The mean SINR versus the maximum allowable transmit power P_{max}(Watts) at the relays for a network with 3 relays without controlling the power usage for each user.

mum allowable transmit power at the relays is increased and that the second Figure has better performance than the first one. In Figure 3.5 the mean SINR versus different number of relays is plotted, while the maximum allowable transmit power at the relays is fixed to $P_{max} = 2$. Again, it can be seen that the performance at the receivers is substantially improved when the number of the relays is increased.
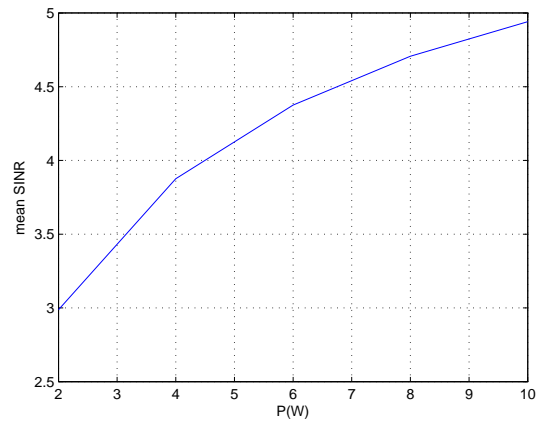
Figure 3.4. The mean balanced SINR versus the maximum allowable transmit power P_{max} at the relays for a network with 3 relays.
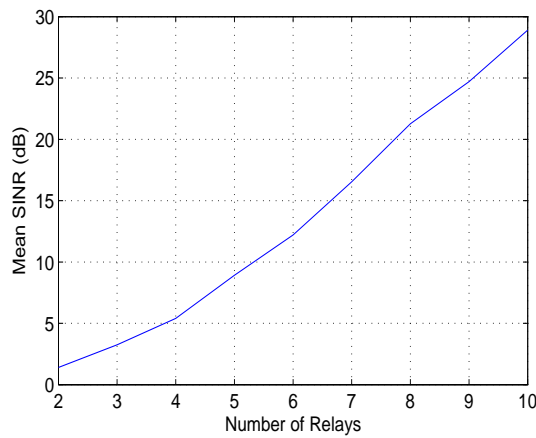
Figure 3.5. The mean SINR versus different number of relays for a network with maximum allowable transmit power P_{max}=2 Watts.

3.3 An iterative semidefinite and Geometric programming technique for the SINR balancing in two-way relay network

In this section a solution to the problem of designing distributed beamformer and allocating power for a two-way multi-relay network is proposed. In such a network, the relay nodes use AF relaying protocol to help two transceivers exchange information in a bidirectional manner. Two-way AF relays have been studied in [85, 88, 134]. However the SINR balancing problem for two-way relay network has only been studied in [88], where one pair of terminals is considered. The solution is based on the assumption that the phase of the beamforming vector has to match the aggregated phase of the channel coefficients from the relay to the two transceivers. In this section, a two-way, multi-user AF relay is considered and an SINR balancing technique is proposed, which uses a max-min fair design approach. The smaller of the transceivers' received SINRs is maximized subject to two constraints, the total transceivers transmit power and the total relay transmit power.

3.3.1 System Model

The model shown in Figure 3.6 consisting of r relays and d source-destination pairs is considered. All nodes of the network are single-antenna units and can either transmit or receive information. For simplification purposes, the sources-destinations on the left side of the Figure will be referred as Transceivers 1 and the sources-destinations on the right side of the Figure as Transceivers 2. The $r \times 1$ complex vector $\mathbf{x}$ of the received signals at the relays can be written as:

$$\mathbf{x} = \sum_{p=1}^{d} \mathbf{f}_p s_p + \sum_{q=1}^{d} \mathbf{g}_q s_q + \boldsymbol{\nu} \tag{3.3.1}$$
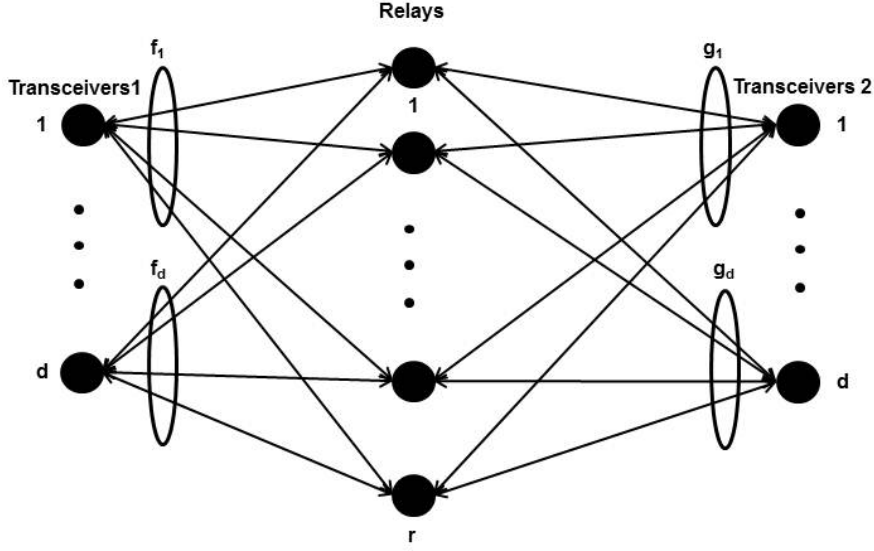
Figure 3.6. A relay network of d pairs of nodes and r relay nodes, all single-antenna units.

where f_{ip}, represent the flat fading channel coefficients from the pth source of Transceivers 1 to the ith relay and g_{iq} represent the flat fading channel coefficients from the ith relay to the qth source of Transceivers 2 respectively. Additionally, the following definitions are used: $\mathbf{x} \triangleq [x_1 \ x_2 \ \cdots \ x_r]^T$, $\boldsymbol{\nu} \triangleq [\nu_1 \ \nu_2 \ \cdots \ \nu_r]^T$, $\mathbf{f}_p \triangleq [f_{1p} \ f_{2p} \ \cdots \ f_{rp}]^T$, $\mathbf{g}_q \triangleq [g_{1q} \ g_{2q} \ \cdots \ g_{rq}]^T$, s_p, s_q are the information symbols transmitted by the pth and qth sources of Transceivers 1 and 2 respectively and $\boldsymbol{\nu}$ is the $r \times 1$ complex vector of the noise at the relays. The ith relay multiplies its received signal by a complex weight coefficient w_i^*. As a result, the vector of the signals transmitted by the relays is given by:

$$\mathbf{t} = \mathbf{W}^* \mathbf{x} \qquad\qquad (3.3.2)$$

where $\mathbf{W} \triangleq \operatorname{diag}([w_1, w_2, \cdots, w_r])$. The received signals at the kth destination of Transceivers 1 and 2 are given by

$$y_{1k} = \mathbf{f}_k^T \mathbf{t} + n_k = \mathbf{f}_k^T \mathbf{W} (\sum_{p=1}^{d} \mathbf{f}_p s_p + \sum_{q=1}^{d} \mathbf{g}_q s_q + \boldsymbol{\nu}) + n_{1k}$$

$$y_{2k} = \mathbf{g}_k^T \mathbf{t} + n_k = \mathbf{g}_k^T \mathbf{W} (\sum_{p=1}^{d} \mathbf{f}_p s_p + \sum_{q=1}^{d} \mathbf{g}_q s_q + \boldsymbol{\nu}) + n_{2k}$$

$$(3.3.3)$$

n_{1k}, n_{2k} is the receiver noises at the kth transceiver for Transceivers 1 and 2 respectively. Since $\mathbf{a}^T \operatorname{diag}(\mathbf{b}) = \mathbf{b}^T \operatorname{diag}(\mathbf{a})$, we can rewrite (3.3.3) as:

$$y_{1k} = \mathbf{w}^H \mathbf{F}_k \mathbf{g}_k s_{2k} + \mathbf{w}^H \mathbf{F}_k \sum_{q=1, q \neq k}^{d} \mathbf{g}_q s_q + \mathbf{w}^H \mathbf{F}_k \mathbf{f}_k s_{1k}$$

$$+ \mathbf{w}^H \mathbf{F}_k \sum_{p=1, p \neq k}^{d} \mathbf{f}_p s_p + \mathbf{w}^H \mathbf{F}_k \boldsymbol{\nu} + n_{1k} \qquad (3.3.4)$$

$$y_{2k} = \mathbf{w}^H \mathbf{G}_k \mathbf{f}_k s_{1k} + \mathbf{w}^H \mathbf{G}_k \sum_{p=1, p \neq k}^{d} \mathbf{f}_p s_p + \mathbf{w}^H \mathbf{G}_k \mathbf{g}_k s_{2k}$$

$$+ \mathbf{w}^H \mathbf{G}_k \sum_{q=1, q \neq k}^{d} \mathbf{g}_q s_q + \mathbf{w}^H \mathbf{G}_k \boldsymbol{\nu} + n_{2k} \qquad (3.3.5)$$

where $\mathbf{F}_k \triangleq \operatorname{diag}(\mathbf{f}_k)$, $\mathbf{G}_k \triangleq \operatorname{diag}(\mathbf{g}_k)$. The third term in (3.3.4) depends on the signal s_k transmitted by transceiver k during the first time slot. As $\mathbf{F}_k \mathbf{f}_k s_k$ is known at transceiver k, the third term in (3.3.4) can be subtracted from $\mathbf{y}_{1k}$ and the residual signal can be processed at the kth transceiver to extract the symbol s_k. Similarly, the third term in (3.3.5) can be subtracted from $\mathbf{y}_{2k}$ to extract the symbol s_k. That is, the residual signals $\widetilde{y}_{1k} \triangleq y_{1k} - \mathbf{w}^H \mathbf{F}_k \mathbf{f}_k s_k$

and $\widetilde{y}_{2k} \triangleq y_{2k} - \mathbf{w}^H \mathbf{G}_k \mathbf{g}_k s_k$ are expressed as:

$$\widetilde{y}_{1k} = \mathbf{w}^H \mathbf{F}_k \mathbf{g}_k s_{2k} + \mathbf{w}^H \mathbf{F}_k \sum_{q=1,q\neq1}^{d} \mathbf{g}_q s_q$$

$$+ \mathbf{w}^H \mathbf{F}_k \sum_{p=1,p\neq1}^{d} \mathbf{f}_p s_p + \mathbf{w}^H \mathbf{F}_k \boldsymbol{\nu} + n_{1k} \qquad (3.3.6)$$

$$\widetilde{y}_{2k} = \mathbf{w}^H \mathbf{G}_k \mathbf{f}_k s_{1k} + \mathbf{w}^H \mathbf{G}_k \sum_{p=1,p\neq1}^{d} \mathbf{f}_p s_p +$$

$$+ \mathbf{w}^H \mathbf{G}_k \sum_{q=1,q\neq1}^{d} \mathbf{g}_q s_q + \mathbf{w}^H \mathbf{G}_k \boldsymbol{\nu} + n_{2k} \qquad (3.3.7)$$

and can be used at the corresponding transceivers to extract the desired information symbols.

3.3.2   SINR Balancing

The optimal powers P_{1k}, P_{2k} of the Transceivers 1 and 2 should be obtained, as well as the relay weight vector $\mathbf{w}$ through maximizing the worst-case user SINR under a total power constraint. Mathematically, the optimization problem is formulated as:

$$\max_{\mathbf{p_1}\succeq0,\mathbf{p_2}\succeq0,\mathbf{w}} \; \min_{k}(\mathrm{SINR}_{1k}, \mathrm{SINR}_{2k})$$

$$\text{s.t.} \quad \sum_{k=1}^{d} P_{1k} + \sum_{k=1}^{d} P_{2k} \leq P_{dmax}, \; P_R \leq P_{Rmax} \qquad (3.3.8)$$

where SINR_{1k}, SINR_{2k}, is the receive SINRs at the kth transceiver for Transceivers 1 and 2 respectively, P_R, P_{Rmax} is the actual and the given maximum transmit power at the relays and P_{dmax} is the given maximum total allowable transmit power at the users, $\mathbf{p_1} \triangleq [P_{11} \; P_{12} \cdots P_{1d}]$, $\mathbf{p_2} \triangleq [P_{21} \; P_{22} \cdots P_{2d}]$, and $\mathbf{p_1} \succeq 0$, $\mathbf{p_2} \succeq 0$ means that all entries of the vectors $\mathbf{p_1}$, $\mathbf{p_2}$ are non-negative.

Using (3.3.2) the total transmit power at the relays can be expressed as:

$$P_R = \mathrm{E}\{\mathbf{t}^H \mathbf{t}\} = \mathbf{w}^H \mathbf{D} \mathbf{w} \tag{3.3.9}$$

where $\mathbf{D} \triangleq \mathbf{E}\{\mathbf{X}^H \mathbf{X}\}$, $\mathbf{X} \triangleq \mathrm{diag}(\mathbf{x})$.

Using (3.3.1) the matrix $\mathbf{D}$ can be written as:

$$\mathbf{D} = \sum_{p=1}^{d} P_{1p} \mathbf{F_p} \mathbf{F_p}^H + \sum_{q=1}^{d} P_{2q} \mathbf{G_q} \mathbf{G_q}^H + \sigma^2 \mathbf{I} \tag{3.3.10}$$

In order to derive the expressions for SINR_{1k}, SINR_{2k}, that represent the desired signal component power P_s^k, the interference power P_i^k and the noise power P_n^k at the kth destination of the Transceivers 1 and 2 are calculated. The kth desired signal power at the Transceiver 1 can be obtained as:

$$P_s^{1k} = \mathrm{E}\{\mathbf{w}^H \mathbf{F}_k \mathbf{g}_k \mathbf{g}_k^H \mathbf{F}_k^H \mathbf{w}\} \mathrm{E}\{|s_{2k}|^2\} = P_{2k} \mathrm{E}\{\mathbf{w}^H (\mathbf{f}_k \odot \mathbf{g}_k)(\mathbf{f}_k^H \odot \mathbf{g}_k^H)\mathbf{w}\}$$

$$= P_{2k} \mathbf{w}^H \mathrm{E}\{\mathbf{h}_k \mathbf{h}_k^H\}\mathbf{w} = P_{2k} \mathbf{w}^H \mathbf{R}_h^k \mathbf{w}$$

$$\tag{3.3.11}$$

and at the Transceiver 2 as

$$P_s^{2k} = P_{1k} \mathbf{w}^H \mathbf{R}_h^k \mathbf{w} \tag{3.3.12}$$

where $\mathbf{h}_k \triangleq (\mathbf{f}_k \odot \mathbf{g}_k) = [f_{1k}g_{1k}\ f_{2k}g_{2k} \cdots f_{Rk}g_{Rk}]^T$ and $\mathbf{R}_h^k \triangleq \mathrm{E}\{\mathbf{h}_k \mathbf{h}_k^H\}$. The vector $\mathbf{h}_k$ contains the total path coefficients between the kth source and its corresponding destination via different relays.

Denoting $D_k = \{1, 2, \ldots, d\} - \{k\}$ and using (3.3.6), the interference power at the kth destination of Transceiver 1 is given by:

$$P_i^{1k} = \mathrm{E}\left\{\mathbf{w}^H \mathbf{F}_k \left(\sum_{q \in D_k} P_q \mathbf{g}_q \mathbf{g}_q^H + \sum_{p \in D_k} P_p \mathbf{f}_p \mathbf{f}_p^H\right) \mathbf{F}_k^H \mathbf{w}\right\}$$

$$= \mathrm{E}\left\{\mathbf{w}^H \left(\sum_{q \in D_k} P_q (\mathbf{f}_k \odot \mathbf{g}_q)(\mathbf{f}_k^H \odot \mathbf{g}_q^H) + \sum_{p \in D_k} P_p (\mathbf{f}_k \odot \mathbf{f}_p)(\mathbf{f}_k^H \odot \mathbf{f}_p^H)\right) \mathbf{w}\right\}$$

$$= \mathbf{w}^H (\sum_{q \epsilon D_k} P_q \mathbf{Q}_{1k} + \sum_{p \epsilon D_k} P_p \mathbf{D}_{1k}) \mathbf{w} = \mathbf{w}^H \mathbf{T_{1k}} \mathbf{w} \qquad (3.3.13)$$

where

$$\mathbf{h}_k^q \triangleq \mathbf{f}_k \odot \mathbf{g}_q, \ \ \mathbf{Q}_{1k} \triangleq \sum_{q \epsilon D_k} \mathbf{h}_k^q (\mathbf{h}_k^q)^H \text{ and } \mathbf{D}_{1k} \triangleq \sum_{p \epsilon D_k} (\mathbf{f}_k \odot \mathbf{f}_p)(\mathbf{f}_k^H \odot \mathbf{f}_p^H).$$

The interference power at the kth destination of Transceiver 2 is given by:

$$P_i^{2k} = \mathbf{w}^H (\sum_{p \epsilon D_k} P_p \mathbf{Q}_{2k} + \sum_{q \epsilon D_k} P_q \mathbf{D}_{2k}) \mathbf{w} = \mathbf{w}^H \mathbf{T_{2k}} \mathbf{w} \qquad (3.3.14)$$

where

$$\mathbf{h}_k^p \triangleq \mathbf{f}_k \odot \mathbf{g}_p, \ \ \mathbf{Q}_{2k} \triangleq \sum_{p \epsilon D_k} \mathbf{h}_k^p (\mathbf{h}_k^p)^H \text{ and } \mathbf{D}_{2k} \triangleq \sum_{q \epsilon D_k} (\mathbf{g}_k \odot \mathbf{g}_q)(\mathbf{g}_k^H \odot \mathbf{g}_q^H).$$

Using (3.3.6) and Assumption A4, the noise power at the kth destination of

Transceiver 1 is written as:

$$P_n^{1k} = \mathrm{E}\{\mathbf{w}^H \mathbf{F}_k \boldsymbol{\nu} \boldsymbol{\nu}^H \mathbf{F}_k^H \mathbf{w}\} + \sigma_n^2 = \mathbf{w}^H \mathbf{R}_f^k \mathbf{w} + \sigma_n^2 \qquad (3.3.15)$$

and Transceiver 2 as:

$$P_n^{2k} = \mathbf{w}^H \mathbf{R}_g^k \mathbf{w} + \sigma_n^2 \qquad (3.3.16)$$

where $\mathbf{R}_f^k \triangleq \sigma_\nu^2 \mathbf{F}_k \mathbf{F}_k^H$ and $\mathbf{R}_g^k \triangleq \sigma_\nu^2 \mathbf{G}_k \mathbf{G}_k^H$.

Before using the above derivations to rewrite the optimization problem, we note that (3.3.8) is equivalent to two subproblems. The first one, calculates the relay weights $\{w_i\}_{i=1}^r$ and maximizes the worst case user SINR. Hence, it is formulated as:

$$\max_{\mathbf{w}} \min_k (\mathrm{SINR}_{1k}, \mathrm{SINR}_{2k})$$
$$\text{s.t.} \quad P_R \leq P_{Rmax}, \quad k = 1, \cdots, d \qquad (3.3.17)$$

where P_{Rmax} is the maximum allowable total transmit power of the relays.

Using (3.3.10) to (3.3.16) and defining $\mathbf{W} \triangleq \mathbf{w}\mathbf{w}^H$, the optimization

problem in (3.3.17) is written as

$$\max_{\mathbf{W}} \min_{k} \left(\frac{\text{Tr}(P_{2k}\mathbf{R}_h^k\mathbf{W})}{\text{Tr}((\mathbf{T}_{1k} + \mathbf{R}_f^k)\mathbf{W}) + \sigma_n^2}, \frac{\text{Tr}(P_{1k}\mathbf{R}_h^k\mathbf{W})}{\text{Tr}((\mathbf{T}_{2k} + \mathbf{R}_g^k)\mathbf{W}) + \sigma_n^2} \right)$$

$$\text{s.t. } \text{Tr}(\mathbf{DW}) \leq P_{Rmax} \quad \text{for} \quad k = 1, 2, \cdots, d \tag{3.3.18}$$

$$\text{Rank}(\mathbf{W}) = 1, \ \mathbf{W} \succeq 0$$

The rank constraint in (3.3.18) is not convex. Using a semidefinite relaxation [38], optimization problem is formulated as follows:

$$\max_{\mathbf{W},t} t$$

$$\text{s.t. } \text{Tr}(\mathbf{W}(P_{2k}\mathbf{R}_h^k - t(\mathbf{T}_{1k} + \mathbf{R}_f^k))) \geq \sigma_n^2 t,$$

$$\text{Tr}(\mathbf{W}(P_{1k}\mathbf{R}_h^k - t(\mathbf{T}_{2k} + \mathbf{R}_g^k))) \geq \sigma_n^2 t, \tag{3.3.19}$$

$$\text{Tr}(\mathbf{DW}) \leq P_{Rmax} \quad \text{for} \quad k = 1, 2, \cdots, d$$

Due to the relaxation, the matrix $\mathbf{W}^*$ obtained by solving the optimization problem in (3.3.19) will not be of rank one all the time. If $\mathbf{W}^*$ happens to be rank one, then its principal eigenvector yields the optimal solution to the original problem. Otherwise alternative techniques such as randomization techniques have to be used (e.g. [136, 139, 140]), to obtain a suboptimal rank-one solution from $\mathbf{W}^*$. The probability of obtaining higher than rank-one solution at varying users is summarized later in the Table 3.3. The simulation parameters for this Table are mentioned at the Simulations subsection.

For any fixed value of t the set of feasible W in (3.3.19) is convex. It follows that the optimization problem in (3.3.19) is quasi convex [40]. To solve (3.3.19), the following observation has to been used. Let t_{max} be the maximum value of t that is obtained by solving the optimization problem

(3.3.19). If for any given t, the convex feasibility problem

$$
\begin{aligned}
&\text{find}\ \ \mathbf{W} \\
&\text{s.t.}\ \ \mathrm{Tr}(\mathbf{W}(P_{2k}\mathbf{R}_h^k - t(\mathbf{T}_{1k} + \mathbf{R}_f^k))) \geq \sigma_n^2 t, \\
&\quad\quad \mathrm{Tr}(\mathbf{W}(P_{1k}\mathbf{R}_h^k - t(\mathbf{T}_{2k} + \mathbf{R}_g^k))) \geq \sigma_n^2 t, \\
&\quad\quad \mathrm{Tr}(\mathbf{DW}) \leq P_{Rmax}\ \ \text{for}\ \ k = 1, 2, \cdots, d\ \ \ \mathbf{W} \succeq 0
\end{aligned}
\tag{3.3.20}
$$

is feasible, then $t_{max} \geq t$. Conversely, if (3.3.20) is not feasible, then $t_{max} \leq t$. Based on this observation, one can check whether the optimal value t_{max} of the quasi-convex problem (3.3.19) is smaller or greater than any given value t using a bisection technique and solving a convex feasibility problem at each step. Starting with a preselected interval $[l, u]$, that contains the optimal value t_{max}, the convex feasibility problem at the midpoint $t = (l + u)/2$ is then solved, to determine whether the optimal value is larger or smaller than t. If (3.3.20) is feasible for this value of t, then $l = t$ is set, otherwise, $u = t$ is chosen and the convex feasibility problem in (3.3.20) is solved again. This procedure is repeated until the difference between u and l is smaller than some preselected threshold δ.

It should be noted that the optimization in (3.3.19) maximizes the worst case user SINR. If the number of relays is substantially higher than the number of users, all users SINR will tend to be equal to the worst case user SINR t_{max}. However, in general, this optimization at the relay will not be able to ensure SINR balancing. This is because, unlike transmit beamforming techniques [17], the relay transceiver has the inability to control the power usage for each user separately at the relays. Therefore, in order to balance the SINR of all users, transmission power control at the source level is essential, using the following GP approach.

Optimization of SINR with regards to the power at the transmitters P_{1k}, P_{2k} is considered, when the relay coefficient vector $\mathbf{w}$ is fixed. In this case,

the following optimization problem has to be solved:

$$\max_{\mathbf{p_1} \succeq 0, \mathbf{p_2} \succeq 0} \min_k (\text{SINR}_{1k}, \text{SINR}_{2k})$$

$$\text{s.t.} \quad \sum_{k=1}^{d} P_{1k} + \sum_{k=1}^{d} P_{2k} \leq P_{dmax} \tag{3.3.21}$$

where P_{dmax} is the maximum allowable total transmit power at the transmitters.

Using (3.3.11) to (3.3.16) and assuming that the relay weights $\{w_i\}_{i=1}^{r}$ are fixed, the optimization problem in (3.3.21) is rewritten as:

$$\max_{\mathbf{p_1} \succeq 0} \min_k \left(\frac{P_{2k} \mathbf{G}_h^k}{\sum\limits_{q \epsilon D_k} P_q \mathbf{E}_h^{1k} + \sum\limits_{p \epsilon D_k} P_p \mathbf{F}_h^{1k} + \mathbf{N}_{1k}} \right)$$

$$\max_{\mathbf{p_2} \succeq 0} \min_k \left(\frac{P_{1k} \mathbf{G}_h^k}{\sum\limits_{p \epsilon D_k} P_p \mathbf{E}_h^{2k} + \sum\limits_{q \epsilon D_k} P_q \mathbf{F}_h^{2k} + \mathbf{N}_{2k}} \right)$$

$$\text{s.t.} \quad \sum_{k=1}^{d} P_{1k} + \sum_{k=1}^{d} P_{2k} \leq P_{dmax} \tag{3.3.22}$$

where $\mathbf{G}_h^k \triangleq \text{Tr}(\mathbf{R}_h^k \mathbf{W})$, $\mathbf{E}_h^{1k} \triangleq \text{Tr}(\mathbf{Q}_{1k} \mathbf{W})$, $\mathbf{E}_h^{2k} \triangleq \text{Tr}(\mathbf{Q}_{2k} \mathbf{W})$, $\mathbf{F}_h^{1k} \triangleq$ $\text{Tr}(\mathbf{D}_{1k} \mathbf{W})$, $\mathbf{F}_h^{2k} \triangleq \text{Tr}(\mathbf{D}_{2k} \mathbf{W})$, $\mathbf{N}_{1k} \triangleq \text{Tr}(\mathbf{R}_f^k \mathbf{W})$ and $\mathbf{N}_{2k} \triangleq \text{Tr}(\mathbf{R}_g^k \mathbf{W})$.

The above problem can be rewritten as:

$$\max_{\mathbf{p_1}, \mathbf{p_2}, \widetilde{t}} \widetilde{t}$$

$$\text{s.t.} \quad \sum_{k=1}^{d} P_{1k} + \sum_{k=1}^{d} P_{2k} \leq P_{dmax}.$$

$$\mathbf{E}_h^{1k} \sum_{q \epsilon D_k} P_q P_{2k}^{-1} \widetilde{t} + \mathbf{F}_h^{1k} \sum_{p \epsilon D_k} P_p P_{2k}^{-1} \widetilde{t} + \mathbf{N}_{1k} P_{2k}^{-1} \widetilde{t} < \mathbf{G}_h^k \tag{3.3.23}$$

$$\mathbf{E}_h^{2k} \sum_{p \epsilon D_k} P_p P_{1k}^{-1} \widetilde{t} + \mathbf{F}_h^{2k} \sum_{q \epsilon D_k} P_q P_{1k}^{-1} \widetilde{t} + \mathbf{N}_{2k} P_{1k}^{-1} \widetilde{t} < \mathbf{G}_h^k$$

The problem (3.3.23) is convex and belongs to the class of GP [137]. As a result, it can be efficiently solved using numerical methods such as cvx [138].

The iterative SINR-balancing algorithm is summarized in Table 3.2.

Table 3.2. Algorithmic solution of the SINR balancing problem

1) **Initialize** $\mathbf{P}_1^{(0)} = [1, ..., 1]^T$, $\mathbf{P}_2^{(0)} = [1, ..., 1]^T$.
2) **Repeat**
3)   Solve (3.3.20) with fixed $\mathbf{P}_{1d}$, $\mathbf{P}_{2d}$ to obtain an updated $\mathbf{W}$.
4)   Solve (3.3.23) with fixed $\mathbf{W}$, to obtain an updated $\mathbf{P}_{1d}$, $\mathbf{P}_{2d}$.
5) **Until** the SINR converges.

3.3.3   Simulation Results

In order to assess the performance of the proposed SINR balancing technique, two numerical examples are considered. A network of 2 source-destination pairs is assumed. The noise power at the relays and at the receivers is assumed to be equal to 0.1. The initial signal power at each source is equal to 1 and the maximum allowable sum power at the transmitters is set to $P_{dmax} = 4$. In Figure 3.7 the maximum allowable transmit power at the relays versus the mean SINR is plotted for a network with 5 relays. From this Figure, we see that the performance of the users is improved as the maximum allowable transmit power at the relays is increased. In Figure 3.8 the mean SINR versus different number of relays is plotted, while the maximum allowable transmit power at the relays is fixed to $P_{max} = 5$. Again, it can be seen that the performance at the receivers is substantially improved when the number of the relays is increased. Finally, for producing the results of Table 3.3 a network of 5 relays with $P_{max} = 5$ is assumed.

Table 3.3. The probability of non-one rank solution for various numbers of users

Users:	4	6	8	10
Probability :	0.005	0.11	0.23	0.4
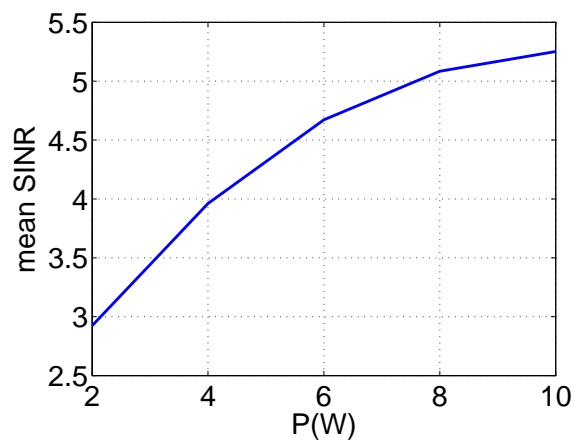
Figure 3.7. The mean SINR versus the maximum allowable relay transmit power P_{max} at the relays for a network with 5 relays.
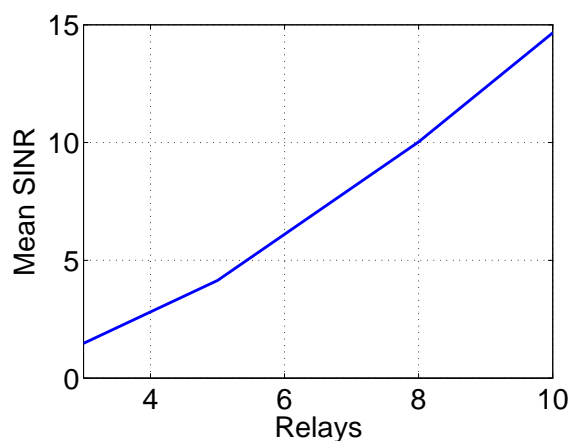
Figure 3.8. The mean SINR versus different number of relays for a network with maximum allowable transmit power $P_{max}=5$

3.4 SINR balancing techniques for a cognitive radio relay network

with multiple peer-to-peer users

In this section, an SINR balancing technique for a CR based peer-to-peer relay network is proposed. The main motivation behind this, is that the radio spectrum is becoming a scarce resource due to ever increasing growth in wireless devices and emerging high data rate interactive and multimedia

applications. Most of the useful spectrum with good propagation characteristics is already very crowded and not much spectrum is available for emerging and future generations of wireless networks. CR [25] is a disruptive technology that has the potential to provide solutions for spectrum scarcity. The idea of CR technology originated from the observation that the licensing based conventional radio spectrum allocation is inefficient as the spectrum is under utilized most of the time at various geographical locations. Hence a dynamic spectrum allocation based on CR principle is actively considered for future generations of wireless networks. The basic idea behind CR technology is that an opportunistic user (namely secondary user) could access the spectrum of the licensed user, namely PU, provided that the SU transmissions do not affect the PU network harmfully. There are generally three types of CR networks namely overlay, underlay and interweaved networks [141]. In the interweave approach, cognitive transmitters are required to sense availability of spectrum and transmit signals only when frequency holes are available. This is also known as white space filling. In the overlay approach, the secondary users help the primary users to offset the interference caused by the secondary transmission by assisting and relaying PU signals. For the underlay approach, the SUs could coexist with the PU network; however they need to ensure the interference caused to the primary receivers is below a predefined threshold to ensure that the primary network is not affected harmfully.

In this section, resource allocation techniques for an underlay CR relay network with multiple peer-to-peer users are considered. All SUs have single antenna at their transmitter and the receiver, hence interference free simultaneous transmission of signals is not possible for the CR transmitters. However, multiple MIMO relays placed between the paths of the transmitters and the receivers can perform spatial multiplexing of the transmitted multiuser signals and forward them to the corresponding receiver destina-
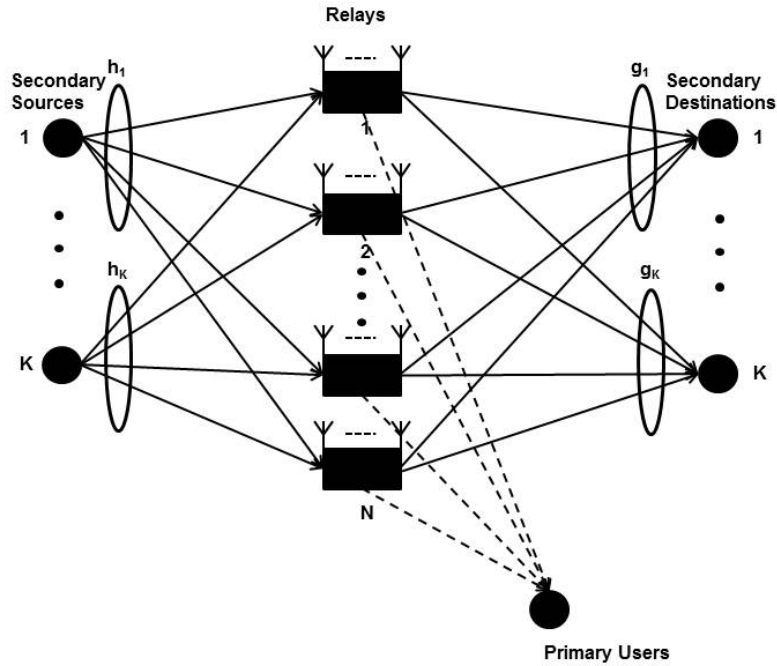
Figure 3.9. The two-way multi-antenna relay channel with multiple users

tions by appropriately processing the received signal at the relay through a set of transceiver matrices.

3.4.1   System Model

As shown in Figure 3.9, a collaborative relay network consisting of one PU, N relays and K number of secondary peer-to-peer users is considered. No direct communication link between the transmitter and the receiver of the secondary users is assumed. The channels are assumed to be quasi-static flat-fading. In this scenario, the SUs use relays to enhance their spectrum efficiency and to increase coverage, while the interference leakage to the PU must be below a certain threshold. During the first time-slot, the K users transmit their data simultaneously to the relay node. In the second time-slot, the relay retransmits the received data towards the receivers of the secondary users. Our model consists of a set of N cooperative relays each equipped with

---

M antennas. The transmitter and the receiver of each secondary user consist of a single antenna. Let $\mathbf{h}_{ip}$ represent the flat fading channel coefficient vector from the ith source to the pth relay and $\mathbf{g}_{jp}$ represent the flat fading channel coefficient vector from the pth relay to the jth destination. The received signal at the relays $\mathbf{y}_R \in \mathbb{C}^{MN \times 1}$ can be written as:

$$\mathbf{y}_R = \sum_{i=1}^{K} \mathbf{h}_i \sqrt{P_i} s_i + \mathbf{n_R}. \tag{3.4.1}$$

where the following definitions are used: $\mathbf{h}_i \triangleq [\mathbf{h}_{i1}^T \ \mathbf{h}_{i2}^T \ \cdots \ \mathbf{h}_{iN}^T]^T$, s_i is the ith source transmitted signal with the corresponding power P_i and $\mathbf{n_R}$ is the $MN \times 1$ complex Gaussian noise vector at the relays. The noise power at the relays is σ_R^2. In the second time slot, the relays process the signals and retransmit the received data towards the secondary users' destinations. The linear operation at the relays can be represented as:

$$\mathbf{x}_R = \mathbf{W} \mathbf{y}_R, \tag{3.4.2}$$

where $\mathbf{x}_R \in \mathbb{C}^{MN \times 1}$ is the transmitted signal at the relays and $\mathbf{W} \in \mathbb{C}^{(MN) \times (MN)}$ is the relay processing matrix, $\mathbf{W} = \text{blkdiag}\{\mathbf{W_1}, \ldots, \mathbf{W_N}\}$, where $\mathbf{W_k}$ is the relay amplification matrix at the kth relay. The received signal at the kth secondary user destination is given by:

$$y_k = \mathbf{g}_k^T \mathbf{x}_R + n_k \mathbf{g}_k^T \mathbf{W} \sum_{i=1}^{K} \mathbf{h}_i \sqrt{P_i} s_i + \mathbf{g}_k^T \mathbf{W} \mathbf{n}_R + n_k, \tag{3.4.3}$$

where $\mathbf{g}_j \triangleq [\mathbf{g}_{j1}^T \ \mathbf{g}_{j2}^T \ \cdots \ \mathbf{g}_{jN}^T]^T$ and n_k is the receiver noise at the kth secondary user destination. The signal v received at the PU is written as:

$$v = \mathbf{z}^T \mathbf{W} \mathbf{y}_R + n_p = \mathbf{z}^T \mathbf{W} \sum_{i=1}^{K} \mathbf{h}_i \sqrt{P_i} s_i + \mathbf{z}^T \mathbf{W} \mathbf{n}_R + n_p$$

where $\mathbf{z} = [\mathbf{z_1}^T \dots \mathbf{z_N}^T]^T$ is the vector of channel coefficients between the PU and the cognitive relays, and n_p is the noise at the PU receiver. The channel state information between the relays and PU can be obtained based on cooperation. The required protocols to obtain these channel state information have been provided in [142], [143].

The transmitted power at the relays is equal to:

$$P_R = \text{Tr}(\text{E}\{\mathbf{x}_R \mathbf{x}_R^H\}) = \sum_{i=1}^{K} ||\mathbf{W}\mathbf{h}_i||^2 P_i + \text{Tr}(\mathbf{W}\mathbf{W}^H)\sigma_R^2. \qquad (3.4.4)$$

The interference and the relay noise leakage to the PU is formulated as:

$$P_I = \sum_{i=1}^{K} |\mathbf{z}^T \mathbf{W}\mathbf{h}_i|^2 P_i + ||\mathbf{W}^H \mathbf{z}^*||^2 \sigma_R^2. \qquad (3.4.5)$$

For later analysis, we have to formulate the SINR cost function. Thus, the received signal power at the *kth* secondary receiver is written as:

$$P_s^k = |\mathbf{g}_k^T \mathbf{W}\mathbf{h}_k|^2 P_k. \qquad (3.4.6)$$

The received interference power at the *kth* secondary receiver is given as:

$$P_i^k = \sum_{j=1, j \neq k}^{K} |\mathbf{g}_k^T \mathbf{W}\mathbf{h}_j|^2 P_j \qquad (3.4.7)$$

The thermal noise power at the *kth* secondary user destination can be written as:

$$P_n^k = ||\mathbf{W}^H \mathbf{g}_k^*||^2 \sigma_R^2 + \sigma_k^2. \qquad (3.4.8)$$

3.4.2   Optimization Problem

Our goal is to jointly adjust the relay coefficient weights and the transmission powers to maximize the worst case user SINR (i.e. SINR balancing), while keeping the interference leakage P_I to the PU below a certain threshold.

Thus, the problem is formulated as P1:

$$\text{P1}: \max_{\mathbf{p} \succeq 0, \mathbf{W}} \quad \min_{k} \text{SINR}_k$$

$$\text{s.t.} \quad \sum_{i=1}^{K} P_k \leq P_{Kmax},$$

$$P_I \leq P_{Imax}, \ P_R \leq P_{Rmax} \qquad (3.4.9)$$

where $\text{SINR}_k = \frac{P_s^k}{P_i^k + P_n^k}$ is the receive SINR at the kth secondary user receiver, P_{Rmax} and P_{Kmax} are the given maximum total allowable transmit power at the relays and the secondary source transmitters, P_{Imax} is the maximum allowable interference leakage to the PU receiver, while $\mathbf{p} \triangleq [P_1 \ldots P_K]$ and $\mathbf{p} \succeq 0$ means that all entries of the vectors $\mathbf{p}$ are non-negative. Next, we break the above problem into two subproblems P1(a) and P1(b). The first P1(a), aims to maximize the worst-case user SINR over all possible relay beamforming weights.

$$\text{P1}(a): \max_{\mathbf{W}} \quad \min_{k} \text{SINR}_k$$

$$\text{s.t.} \quad P_R \leq P_{Rmax}, \ P_I \leq P_{Imax}, \qquad (3.4.10)$$

while the second P1(b), aims to balance all the users' SINR over all possible transmission powers.

$$\text{P1}(b): \max_{\mathbf{p} \succeq 0} \quad \min_{k} \text{SINR}_k$$

$$\text{s.t.} \quad \sum_{k=1}^{K} P_k \leq P_{Kmax},$$

$$P_I \leq P_{Imax}. \qquad (3.4.11)$$

The first subproblem P1(a) can be written equivalently as:

$$\max_{\mathbf{W},t} \quad t \tag{3.4.12}$$

$$\text{s.t.} \quad \text{SINR}_k \geq t$$

$$P_R \leq P_{Rmax}, \ P_I \leq P_{Imax}.$$

Next, we focus on how to solve problem (3.4.12). For the convenience of analysis, we modify P_i^k, P_s^k and P_n^k as follows.

Using the Kronecker product identity $\text{Vec}(\mathbf{AXB}) = (\mathbf{B}^T \otimes \mathbf{A})\text{Vec}(\mathbf{X})$ and by defining $\mathbf{w}_k = \text{Vec}(\mathbf{W}_k)$, the equations (3.4.6), (3.4.7) can be rewritten as:

$$P_s^k = |(\mathbf{h}_{k1}^T \otimes \mathbf{g}_{k1}^T)\mathbf{w_1} + \ldots + (\mathbf{h}_{kN}^T \otimes \mathbf{g}_{kN}^T\mathbf{w_N})|^2 P_k \tag{3.4.13}$$

$$P_i^k = \sum_{\substack{j=1 \\ j \neq k}}^{K} |(\mathbf{h}_{j1}^T \otimes \mathbf{g}_{k1}^T)\mathbf{w}_1 + \ldots + (\mathbf{h}_{jN}^T \otimes \mathbf{g}_{kN}^T)\mathbf{w}_N|^2 P_j \tag{3.4.14}$$

Let $\mathbf{f}_k = [(\mathbf{h}_{k1}^T \otimes \mathbf{g}_{k1}^T) \ldots (\mathbf{h}_{kN}^T \otimes \mathbf{g}_{kN}^T)]$, $\mathbf{q}_{jk} = [(\mathbf{h}_{j1}^T \otimes \mathbf{g}_{k1}^T) \ldots (\mathbf{h}_{jN}^T \otimes \mathbf{g}_{kN}^T)]$ and $\mathbf{x} = [\mathbf{w}_1^T \ldots \mathbf{w}_N^T]^T$.

Then it follows from (3.4.13) and (3.4.14) that:

$$P_s^k = |\mathbf{f}_k\mathbf{x}|^2 P_k$$

$$P_i^k = \sum_{\substack{j=1 \\ j \neq k}}^{K} |\mathbf{q}_{jk}\mathbf{x}|^2 P_j. \tag{3.4.15}$$

Furthermore, by defining:

$$\widetilde{\mathbf{G}}_k = \text{blkdiag}\{\mathbf{G}_{k1}, \ldots, \mathbf{G}_{kN}\}, \ \text{for } k = 1, \ldots K, \tag{3.4.16}$$

$$\mathbf{G}_{kn} = \begin{bmatrix} \mathbf{g}_{kn}(1,1) & \mathbf{0} & \ldots & \mathbf{g}_{kn}(M,1) & \mathbf{0} \\ \mathbf{0} & \mathbf{g}_{kn}(1,1) & \ldots & \mathbf{0} & \mathbf{g}_{kn}(M,1) \end{bmatrix} \tag{3.4.17}$$

for $n = 1, \ldots N$, we have $||\mathbf{W}^H \mathbf{g}_k^*||^2 = ||\widetilde{\mathbf{G}}_k \mathbf{x}||^2$ and (3.4.8) becomes:

$$P_n^k = ||\widetilde{\mathbf{G}}_k \mathbf{x}||^2 \sigma_R^2 + \sigma_k^2. \tag{3.4.18}$$

Using above transformations, (3.4.5) can be rewritten as:

$$P_I = \sum_{j=1}^{K} |\mathbf{t}_j \mathbf{x}|^2 P_j + ||\widetilde{\mathbf{Z}}^H \mathbf{x}^*||^2 \sigma_R^2 + \sigma_p^2, \tag{3.4.19}$$

because $||\widetilde{\mathbf{Z}}^H \mathbf{x}^*||^2 = ||\mathbf{W}^H \mathbf{z}^*||^2$. The matrix $\widetilde{\mathbf{Z}}$ has been generated from $\mathbf{z}$ in a similar way that $\widetilde{\mathbf{G}}_k$ has been from $\mathbf{g}_k^*$. Moreover, $\mathbf{t}_j = [(\mathbf{h}_{j1}^T \otimes \mathbf{z}^T) \ldots (\mathbf{h}_{jN}^T \otimes \mathbf{z}^T)]$. Next we can express P_R in (3.4.4) as $P_R = \sum_{k=1}^{K} ||\widehat{\mathbf{H}}_{\mathbf{k}} \mathbf{x}||^2 P_k + \mathrm{Tr}(\mathbf{x}\mathbf{x}^H)\sigma_R^2$, where

$$\widehat{\mathbf{H}}_{\mathbf{k}} = \mathrm{blkdiag}\{\widehat{\mathbf{H}}_{\mathbf{k1}}, \ldots, \widehat{\mathbf{H}}_{\mathbf{kN}}\}, \text{ for } k = 1, \ldots K \tag{3.4.20}$$

$$\widehat{\mathbf{H}}_{\mathbf{kn}} = \mathrm{blkdiag}\{\mathbf{h}_{\mathbf{k1}}, \ldots, \mathbf{h}_{\mathbf{kM}}\}, \text{ for } n = 1, \ldots N \tag{3.4.21}$$

Using the above transformations, (3.4.12) can be rewritten as

$$\begin{aligned}
&\max_{\mathbf{x},t} \quad t \\
&\text{subject to} \\
&\frac{|\mathbf{f}_k \mathbf{x}|^2 P_k}{\sum_{\substack{j=1 \\ j \neq k}}^{K} |\mathbf{q}_{j,k} \mathbf{x}|^2 P_j + ||\widetilde{\mathbf{G}}_k \mathbf{x}||^2 \sigma_R^2 + \sigma_k^2} \geq t, \\
&\quad \text{for } k = 1, \ldots K \\
&\sum_{j=1}^{K} |\mathbf{t}_{\mathbf{j}} \mathbf{x}|^2 P_j + ||\widetilde{\mathbf{Z}}^H \mathbf{x}^*||^2 \sigma_R^2 \leq P_{Imax}, \\
&\sum_{j=1}^{K} ||\widehat{\mathbf{H}}_{\mathbf{j}} \mathbf{x}||^2 P_j + \mathrm{Tr}(\mathbf{x}\mathbf{x}^H)\sigma_R^2 \leq P_R.
\end{aligned}$$

$$\tag{3.4.22}$$

The above problem is still non-convex. However, the optimal solution could

be obtained via a relaxed SDP.

We first define: $\mathbf{X} = \mathbf{x}\mathbf{x}^H$, $\mathbf{\Theta} = \sum_{k=1}^{K} \widehat{\mathbf{H}}_{\mathbf{k}}^H \widehat{\mathbf{H}}_{\mathbf{k}} P_k + \sigma_R^2 \mathbf{I}$, $\mathbf{E_k} = \mathbf{f}_k \mathbf{f}_k^H$, $\mathbf{D_{j,k}} = \mathbf{q}_{j,k}\mathbf{q}_{j,k}^H$, $\mathbf{N_k} = \widetilde{\mathbf{G}}_{k1}^H \widetilde{\mathbf{G}}_k \sigma_R^2$, $\mathbf{C_j} = \mathbf{t_j}\mathbf{t_j^H}$.

Then problem (3.4.22) can be equivalently rewritten as:

$$
\begin{aligned}
\max_{\mathbf{x},t} \quad & t \\
\text{s.t.} \quad & \frac{\text{Tr}(P_k \mathbf{E_k} \mathbf{X})}{\sum_{\substack{j=1 \\ j \neq k}} \text{Tr}(\mathbf{D_j} \mathbf{X}) P_j + \text{Tr}(\mathbf{N}_k \mathbf{X}) + \sigma_k^2} \geq t, \\
& \text{for } k = 1, \dots K \\
& \sum_{j=1}^{K} \text{Tr}(\mathbf{C_j} \mathbf{X} P_j) + \widetilde{\mathbf{Z}}^H \widetilde{\mathbf{Z}} \sigma_R^2 \leq P_{Imax}, \\
& \text{Tr}(\mathbf{\Theta_k} \mathbf{X}) \leq P_R \\
& \text{Rank}(\mathbf{X}) = 1, \ \mathbf{X} \succeq 0. \quad\quad (3.4.23)
\end{aligned}
$$

The above problem is still not convex due to the rank-one constraint. However, if we remove this constraint, the problem is relaxed into a SDP problem formulated as:

$$
\begin{aligned}
\max_{\mathbf{x},t} \quad & t \\
\text{s.t.} \quad & \frac{\text{Tr}(P_k \mathbf{E_k} \mathbf{X})}{\sum_{\substack{j=1 \\ j \neq k}} \text{Tr}(\mathbf{D_j} \mathbf{X}) P_j + \text{Tr}(\mathbf{N}_k \mathbf{X}) + \sigma_k^2} \geq t, \\
& \text{for } k = 1, \dots K \\
& \sum_{j=1}^{K} \text{Tr}(\mathbf{C_j} \mathbf{X} P_j) + \widetilde{\mathbf{Z}}^H \widetilde{\mathbf{Z}} \sigma_R^2 \leq P_{Imax} \\
& \text{Tr}(\mathbf{\Theta_k} \mathbf{X}) \leq P_R, \ \mathbf{X} \succeq 0. \quad\quad (3.4.24)
\end{aligned}
$$

The above relaxed SDP can be solved using standard convex optimization toolbox. Note that the above subproblem only maximizes the worst case user SINR, but may not balance the SINR of all users. This is because, unlike transmit beamforming techniques [17], the relay transceiver has the inability to control the power usage for each user separately at the relays.

Therefore, in order to balance the SINR of all users, we need to control the

transmission power at the source level.

For this purpose, we solve problem P1(b):

$$
\max_{\mathbf{p} \succeq 0} \quad \min_{k} \frac{\mathrm{Tr}(P_k \mathbf{E_k X})}{\sum_{\substack{j=1 \\ j \neq k}} \mathrm{Tr}(\mathbf{D_j X}) P_j + \mathrm{Tr}(\mathbf{N}_k \mathbf{X}) + \sigma_k^2}
$$

$$
\text{for } k = 1, \ldots K
$$

$$
\text{s.t.} \quad \sum_{k=1}^{K} P_k \leq P_{Kmax},
$$

$$
\sum_{j=1}^{K} \mathrm{Tr}(\mathbf{C_j X} P_j) + \widetilde{\mathbf{Z}}^H \widetilde{\mathbf{Z}} \sigma_R^2 \leq P_{Imax}. \tag{3.4.25}
$$

The above problem can be written as:

$$
\max_{\mathbf{p_1}, \widetilde{t}} \widetilde{t} \text{ s.t. to}
$$

$$
\sum_{\substack{j=1 \\ j \neq k}} \mathrm{Tr}(\mathbf{D_j X}) P_j + \mathrm{Tr}(\mathbf{N}_k \mathbf{X}) + \sigma_k^2 \leq \mathrm{Tr}(\mathbf{E_k X}) P_k \widetilde{t}^{-1},
$$

$$
\text{for } k = 1, \ldots K
$$

$$
\sum_{k=1}^{K} P_k \leq P_{Kmax},
$$

$$
\sum_{j=1}^{K} \mathrm{Tr}(\mathbf{C_j X} P_j) + \widetilde{\mathbf{Z}}^H \widetilde{\mathbf{Z}} \sigma_R^2 \leq P_{Imax}. \tag{3.4.26}
$$

The problem (3.4.26) is convex and belongs to the class of GP [137] . As a

result, it can be efficiently solved using interior point methods [138].

Based on the above we get the following algorithm for P1:

3.4.3   Simulation Results

The performance of the proposed scheme is assessed with the help of two

numerical examples. The elements of channels $\mathbf{H}$, $\mathbf{G}$ and $\mathbf{Z}$ are assumed to

be circularly symmetric complex Gaussian (CSCG) variables with zero mean

and unity variance. The noise variance at the destination nodes and at the

Table 3.4. Algorithmic solution of the SINR balancing problem

1) **Initialize p**.
2) **Repeat**
3)   Solve problem P1(a) to obtain optimal value of **W**
4)   Solve problem P1(b) to obtain optimal value of **p**
5)   Update the old values of **p** before solving P1(a)
6) **Until** $\text{SINR}_{current} - \text{SINR}_{previous} \leq \varepsilon$,
where ε is a small positive constant for controlling the accuracy
The converged value of SINR is the optimal solution of SINR
in (3.4.11).

relay was set to 0.1W and the stopping criterion for the iterative method was set to 10^{-3}.

In the first example, we study the effect of the different maximum allowable relay transmit power P_{Rmax} to the mean SINR for a two-user network with four relays, each equipped with two antennas. The maximum transmit power at the users is P_{Kmax}=10W and the maximum allowable interference leakage to the PU is P_{Imax}=0.2W. In Figure 3.10 it is observed that the mean SINR increases as the maximum allowable relay power increases.

In the second example, a two-user network with different number of relays is simulated, each of them equipped with two antennas. The maximum allowable relay transmit power P_{Rmax}=5W and the maximum transmit power at the users is P_{Kmax}=10W. The results were generated by using a Monte Carlo experiment with 100 independent realization of the channel gains. As can been seen in Figure 3.11, the average mean SINR increases as the number of relays in the network increases.
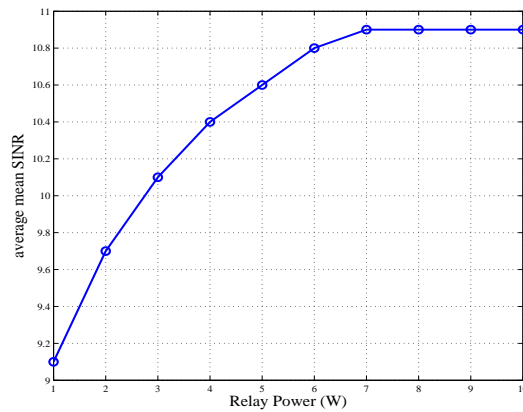
Figure 3.10. The mean SINR versus the maximum allowable relay transmit power P_{Rmax} at the relays for a network with 2 users on both sides, 4 relays, equipped with 2 antennas each and P_{Imax}=0.2W.
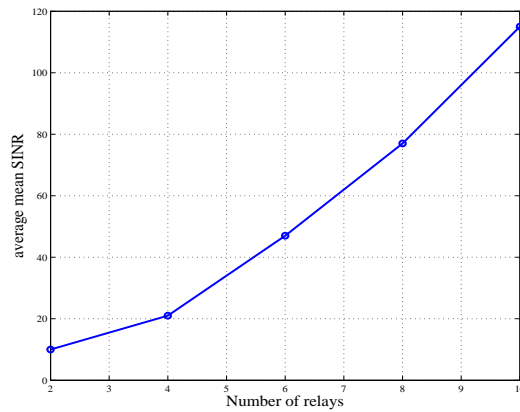
Figure 3.11. The mean SINR versus different number of relays for a network with maximum allowable transmit power P_{Rmax}=5W and P_{Imax}=0.2W.

3.5   An SINR Balancing Technique for a Cognitive Two-Way Relay Network

In this section the SINR balancing technique proposed in section 3.4, is extended to a two-way AF relaying scheme with multiple source-destination pairs of SUs under a CR environment. The aim is the design of relay beamforming complex coefficients in order to maximize the worst-case SU SINR subject to total transmit power constraint, while ensuring that the interfer-
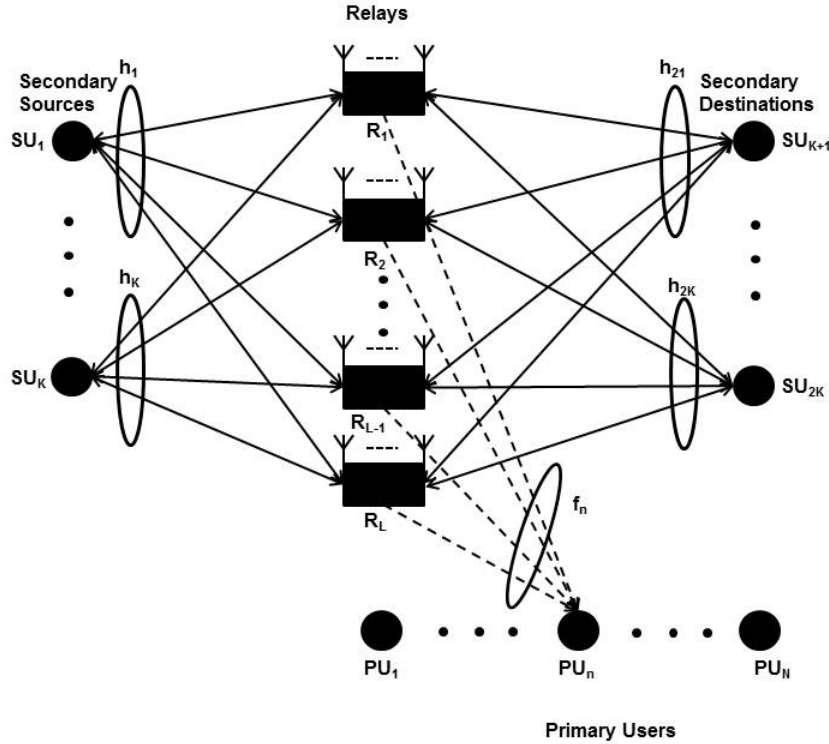
Figure 3.12. The two-way multi-antenna relay channel with multiple users

ence leakage to PUs is below specific thresholds.

3.5.1   System Model

A wireless two-way cognitive relay network is considered, which consists of K single-antenna source-destination pairs and L multiple-antenna relay nodes, each equipped with M antennas, all operating in the same frequency band allocated to N single-antenna PUs PU_n $n = 1, \ldots, N$. The pair of sources k and $K + k$ exchange messages through the set of L relays. All the channels are assumed to be independent, frequency-flat Rayleigh block-fading and the channel links are established through a two-step AF cooperative scheme. In the first step, the $2K$ sources transmit simultaneously their source information to the L relays. In the second step, the L relays amplify their respective faded mixtures of their received signals and relay them to the destination receivers. Let s_k be the information symbol transmitted by the kth source,

which transmits with power p_k, i.e. $E\{|s_k|^2\} = p_k$ for $k = 1, \ldots, 2K$. It should hold that

$$\mathbf{1}^T \mathbf{p} \le P_1 \tag{3.5.1}$$

where $\mathbf{p} \triangleq [p_1 \ldots p_{2K}]^T$ and P_1 is the total power available at the sources. Let $\mathbf{h}_{lk} \triangleq [h_{1lk}, \ldots, h_{Mlk}]^T$ stand for the channel between the kth source and the lth relay, then we introduce $\mathbf{h}_k \triangleq [\mathbf{h}_{1k}^T, \ldots, \mathbf{h}_{Lk}^T]^T$. Thus, the $(ML \times 1)$ received signal vector at the relays is given by

$$\mathbf{y}_R = \sum_{k=1}^{2K} \mathbf{h}_k s_k + \mathbf{n}_R \tag{3.5.2}$$

where $\mathbf{n}_R \triangleq [\mathbf{n}_1^T \ldots \mathbf{n}_L^T]^T$ contains the noise components present at the relay and the noise components are assumed to be zero-mean, spatially uncorrelated and unity variance. The lth relay multiplies its received signal by a processing matrix $\mathbf{W}_l$. Thus, the vector of signals transmitted by all relays is given by

$$\mathbf{x}_R = \mathbf{W}\mathbf{y}_R \tag{3.5.3}$$

where $\mathbf{W} = \text{blkdiag}\{\mathbf{W}_1 \ldots \mathbf{W}_L\}$. From (3.5.2) and (3.5.3) the relays' total transmit power is given by

$$
\begin{aligned}
P_r = E\{\mathbf{x}_R^H \mathbf{x}_R\} &= \sum_{k=1}^{2K} ||\mathbf{W}\mathbf{h}_k||^2 p_k + \text{Tr}(\mathbf{W}\mathbf{W}^H) \\
&= \sum_{k=1}^{2K} ||\mathbf{A}_k \mathbf{b}||^2 p_k + \text{Tr}(\mathbf{b}\mathbf{b}^H) \tag{3.5.4}
\end{aligned}
$$

where

$\mathbf{A}_k = \text{blkdiag}\{\mathbf{A}_{k1} \ldots \mathbf{A}_{kL}\}$, for $k = 1, \ldots, 2K$,

$\mathbf{A}_{kl} = \text{blkdiag} \overbrace{\{\mathbf{h}_{kl} \ldots \mathbf{h}_{kl}\}}^{M times}$, $\mathbf{b} = [\mathbf{w}_1^T \ldots \mathbf{w}_L^T]^T$ and $\mathbf{w}_l = \text{Vec}(\mathbf{W}_l)$. It is also required that

$$P_r \le P_2 \tag{3.5.5}$$

where P_2 is the upper-bound on the relays' total transmit power.

Let $\mathbf{f}_{nl} = [f_{1nl} \ldots f_{Mnl}]^T$ denote the channel gains between the lth relay and the nth PU, $\mathbf{f}_n = [\mathbf{f}_{n1}^T \ldots \mathbf{f}_{nl}^T]^T$ the channel gain between all the relays and the nth PU and I_n^R is the acceptable interference power threshold caused by the relays on PU_n, then it should hold that

$$
\begin{aligned}
E\{|\mathbf{f}_n^T \mathbf{x}_R|^2\} &= \sum_{k=1}^{2K} |\mathbf{f}_n^T \mathbf{W} \mathbf{h}_k|^2 p_k + ||\mathbf{W} \mathbf{f}_n||^2 + \sigma_p^2 \\
&= \sum_{k=1}^{2K} |\mathbf{t}_k^T \mathbf{b}|^2 p_k + ||\mathbf{F}_n \mathbf{b}||^2 + \sigma_p^2 \leq I_n^R \\
&\quad \text{for } n = 1, 2, \ldots, N \qquad\qquad (3.5.6)
\end{aligned}
$$

where $\mathbf{t}_k = [(\mathbf{h}_{1k}^T \otimes \mathbf{f}_n^T) \ldots (\mathbf{h}_{Lk}^T \otimes \mathbf{f}_n^T)]$ has been obtained by applying the Kronecker product identity $\text{Vec}(\mathbf{AXB}) = (\mathbf{B}^T \otimes \mathbf{A})\text{Vec}(\mathbf{X})$ and

$$
\mathbf{F}_n = \text{blkdiag}\{\mathbf{F}_{n1} \ldots \mathbf{F}_{nL}\}, \text{ for } n = 1, \ldots, N, \ \mathbf{F}_{nl} = \text{blkdiag} \overbrace{\{\mathbf{f}_{nl} \ldots \mathbf{f}_{nl}\}}^{M times}.
$$

The received signal at the kth destination is given by

$$
y_k = \mathbf{h}_k^T \mathbf{x}_R + n_k = \mathbf{h}_k^T \mathbf{W} \sum_{i=1}^{2K} \mathbf{h}_i s_i + \mathbf{h}_k^T \mathbf{W} \mathbf{n}_R + n_k
$$

$$(3.5.7)$$

where n_k is the noise at the kth destination. Note that the first term of the above equation contains the self-interference $\mathbf{h}_k s_k$. Assume that h_k is known at kth user via training-based estimation [57] before data transmission. Thus, kth receiver can subtract its self-interference from y_k. From (3.5.7), subtracting the self-interference, we obtain:

$$
\begin{aligned}
\widetilde{y}_k &= \mathbf{h}_k^T \mathbf{W} \sum_{\substack{i=1 \\ i \neq k}}^{2K} \mathbf{h}_i s_i + \mathbf{h}_k^T \mathbf{W} \mathbf{n}_R + n_k = \mathbf{h}_k^T \mathbf{W} \mathbf{h}_j s_j + \mathbf{h}_k^T \mathbf{W} \sum_{\substack{i=1 \\ i \neq k,j}}^{2K} \mathbf{h}_i s_i \\
&\quad + \mathbf{h}_k^T \mathbf{W} \mathbf{n}_R + n_k \qquad\qquad\qquad\qquad\qquad\qquad\qquad (3.5.8)
\end{aligned}
$$

where

$$
j = \begin{cases} K + k & \text{if } k \leq K \\ -K + k & \text{if } k > K \end{cases}
$$

The first term of (3.5.8) is the desired signal, the second is the interference and the last two terms describe the total noise received at the kth destination.

Thus, the received signal and interference power at the kth receiver may be represented as:

$$P_s^k = |\mathbf{h}_k^T \mathbf{W} \mathbf{h}_j|^2 p_k = |\mathbf{d}_k \mathbf{b}|^2 p_k$$
$$P_i^k = \sum_{\substack{i=1 \\ i \neq k,j}}^{2K} |\mathbf{h}_k^T \mathbf{W} \mathbf{h}_i|^2 p_i = \sum_{\substack{i=1 \\ i \neq k,j}}^{2K} |\mathbf{q}_i \mathbf{b}|^2 p_i$$

$$(3.5.9)$$

where

$\mathbf{d}_k = [(\mathbf{h}_{1j}^T \otimes \mathbf{h}_{1k}^T) \ldots (\mathbf{h}_{Lj}^T \otimes \mathbf{h}_{Lk}^T)]$ and $\mathbf{q}_i = [(\mathbf{h}_{1i}^T \otimes \mathbf{h}_k^T) \ldots (\mathbf{h}_{Li}^T \otimes \mathbf{h}_{Lk}^T)]$.

The aggregated noise power received at the kth destination is

$$P_n^k = ||\mathbf{W}\mathbf{h}_k||^2 + \sigma_k^2 = ||\mathbf{N}_k \mathbf{b}||^2 + \sigma_k^2$$

$$(3.5.10)$$

where

$\mathbf{N}_k = \text{blkdiag}\{\mathbf{N}_{k1} \ldots \mathbf{N}_{kL}\}$, for $k = 1, \ldots, 2K$ and

$\mathbf{N}_{kl} = \text{blkdiag} \overbrace{\{\mathbf{h}_{kl} \ldots \mathbf{h}_{kl}\}}^{M times}$.

Using (3.5.9), (3.5.10) the SINR at kth destination is:

$$\text{SINR}_k(\mathbf{p}, \mathbf{b}) = \frac{|\mathbf{d}_k \mathbf{b}|^2 p_k}{\sum_{\substack{i=1 \\ i \neq k,j}}^{2K} |\mathbf{q}_i \mathbf{b}|^2 p_i + ||\mathbf{N}_k \mathbf{b}||^2 + \sigma_k^2}$$

$$(3.5.11)$$

3.5.2   Optimization Problem

The design goal is to jointly adjust the relay coefficient weights $\mathbf{b}$ and the transmission power vector $\mathbf{p}$ to maximize the worst user SINR (i.e. SINR balancing), while keeping both the interference leakage to the PUs and the total transmit and relays' powers below a certain threshold. Thus, the prob-

lem is formulated as P1:

$$\text{P1}: \max_{\mathbf{p} \succeq 0, \mathbf{b}} \quad \min_{1 \leq k \leq 2K} \text{SINR}_k(\mathbf{p}, \mathbf{b})$$

$$\text{s.t.} \quad \mathbf{1}^T \mathbf{p} \leq P_1$$

$$\sum_{k=1}^{2K} ||\mathbf{A}_k \mathbf{b}||^2 p_k + tr(\mathbf{b}\mathbf{b}^H) \leq P_2$$

$$\sum_{k=1}^{2K} |\mathbf{t}_k^T \mathbf{b}|^2 p_k + ||\mathbf{F}_n \mathbf{b}||^2 + \sigma_p^2 \leq I_n^R,$$

$$\text{for } n = 1, 2 \ldots N \qquad (3.5.12)$$

The above optimization problem is not convex with respect to both design parameters $\mathbf{b}$, $\mathbf{p}$ and is thus difficult to solve jointly via standard convex optimization techniques. Therefore we need to break the above problem into two subproblems P1(a) and P1(b).

The first P1(a) aims to maximize the worst-case user SINR over all possible relay beamforming weights $\mathbf{b}$ for a given sources' transmit power.

$$\text{P1}(a): \max_{\mathbf{b}} \quad \min_{1 \leq k \leq 2K} \text{SINR}_k$$

$$\text{s.t.} \quad \sum_{k=1}^{2K} ||\mathbf{A}_k \mathbf{b}||^2 p_k + tr(\mathbf{b}\mathbf{b}^H) \leq P_2$$

$$\sum_{k=1}^{2K} |\mathbf{t}_k^T \mathbf{b}|^2 p_k + ||\mathbf{F}_n \mathbf{b}||^2 + \sigma_p^2 \leq I_n^R,$$

$$\text{for } n = 1, 2 \ldots N \qquad (3.5.13)$$

while the second P1(b), aims to balance all the users' SINR over all possible transmission powers for the given beamforming vectors.

$$\text{P1}(b): \max_{\mathbf{p} \succeq 0} \quad \min_{1 \leq k \leq 2K} \text{SINR}_k$$

$$\text{s.t.} \quad \mathbf{1}^T \mathbf{p} \leq P_1$$

$$\sum_{k=1}^{2K} ||\mathbf{A}_k \mathbf{b}||^2 p_k + tr(\mathbf{b}\mathbf{b}^H) \leq P_2$$

$$\sum_{k=1}^{2K} |\mathbf{t}_k^T \mathbf{b}|^2 p_k + ||\mathbf{F}_n \mathbf{b}||^2 + \sigma_p^2 \leq I_n^R,$$

$$\text{for } n = 1, 2 \ldots N \qquad (3.5.14)$$

Introducing $\mathbf{B} = \mathbf{b}\mathbf{b}^H$, $\mathbf{D}_k = \mathbf{d}_k\mathbf{d}_k^H$, $\mathbf{Q}_k = \mathbf{q}_i\mathbf{q}_i^H$, $\mathbf{N} = \widetilde{\mathbf{N}}_k\widetilde{\mathbf{N}}_k^H$, $\mathbf{R}_k = \mathbf{A}_k\mathbf{A}_k^H$ and $\mathbf{T}_k = \mathbf{t}_k\mathbf{t}_k^H + \mathbf{F}_n\mathbf{F}_n^H$ the first subproblem P1(a) can be written equivalently as:

$$
\begin{aligned}
\max_{\mathbf{B},t} \quad & t \\
\text{s.t.} \quad & \frac{\mathrm{Tr}(\mathbf{D}_k\mathbf{B})p_k}{\sum_{\substack{i=1\\i\neq k,j}}^{2K} p_i\mathrm{Tr}(\mathbf{Q}_k\mathbf{B}) + \mathrm{Tr}(\mathbf{N}\mathbf{B}) + \sigma_k^2} \geq t, \\
& \text{for } k = 1,\ldots,2K \\
& \sum_{k=1}^{2K} p_k\mathrm{Tr}(\mathbf{R}_k\mathbf{B}) + \mathrm{Tr}(\mathbf{B}) \leq P_2 \\
& \sum_{k=1}^{2K} p_k\mathrm{Tr}(\mathbf{T}_k\mathbf{B}) + \sigma_p^2 \leq I_n^R, \\
& \text{for } n = 1,2\ldots N \\
& \mathrm{rank}(\mathbf{B}) = 1, \; \mathbf{B} \succeq 0. \quad\quad (3.5.15)
\end{aligned}
$$

Due to the constraint rank($\mathbf{B}$)= 1, the optimization problem in (3.5.15) is not convex. Hence we remove the rank constraint so that the problem is relaxed into SDP [136], [144] as follows:

$$
\begin{aligned}
\max_{\mathbf{B},t} \quad & t \\
\text{s.t.} \quad & \frac{\mathrm{Tr}(\mathbf{D}_k\mathbf{B})p_k}{\sum_{\substack{i=1\\i\neq k,j}}^{2K} p_i\mathrm{Tr}(\mathbf{Q}_k\mathbf{B}) + \mathrm{Tr}(\mathbf{N}\mathbf{B}) + \sigma_k^2} \geq t, \\
& \text{for } k = 1,\ldots,2K \\
& \sum_{k=1}^{2K} p_k\mathrm{Tr}(\mathbf{R}_k\mathbf{B}) + \mathrm{Tr}(\mathbf{B}) \leq P_2 \\
& \sum_{k=1}^{2K} p_k\mathrm{Tr}(\mathbf{T}_k\mathbf{B}) + \sigma_p^2 \leq I_n^R, \\
& \text{for } n = 1,2\ldots N \\
& \mathbf{B} \succeq 0 \quad\quad (3.5.16)
\end{aligned}
$$

For any given $t = t_r$ the feasible set in (3.5.16) is convex. Given the convexity of the above SDP problem, the optimal solution could be efficiently

found by using bisection method. By solving (3.5.16), we obtain the maximum value of t that satisfies the constraints and the corresponding value of **B**.

SDP relaxation usually leads to an optimal B with rank one for the problem in (17). However, if rank of the matrix B turns out to be greater than one, randomization techniques [145] can be used to obtain a rank one solution.

Note that the above subproblem only maximized the worst case user SINR. However, if the number of the relays is substantially higher than the number of users, the resulting SINR will tend to be equal for all users. This is because, unlike transmit beamforming techniques [17], the relay transceivers has the inability to control the power usage for each user separately at the relays. Therefore, in order to balance the SINR of all users, we need to control the transmission power at the source level.

For this purpose, we solve problem P1(b):

$$
\begin{aligned}
\max_{\mathbf{p} \succeq 0} \quad & \min_{1 \leq k \leq 2K} \frac{\mathrm{Tr}(\mathbf{D}_k \mathbf{B}) p_k}{\sum_{\substack{i=1 \\ i \neq k, j}}^{2K} p_i \mathrm{Tr}(\mathbf{Q}_k \mathbf{B}) + \mathrm{Tr}(\mathbf{NB}) + \sigma_k^2} \\
\text{s.t.} \quad & \mathbf{1}^T \mathbf{p} \leq P_1 \\
& \sum_{k=1}^{2K} p_k \mathrm{Tr}(\mathbf{R}_k \mathbf{B}) + \mathrm{Tr}(\mathbf{B}) \leq P_2 \\
& \sum_{k=1}^{2K} p_k \mathrm{Tr}(\mathbf{T}_k \mathbf{B}) + \sigma_p^2 \leq I_n^R, \\
& \text{for } n = 1, 2 \ldots N
\end{aligned}
\tag{3.5.17}
$$

The above problem can be rewritten as:

$$
\begin{aligned}
\max_{\mathbf{p} \succeq 0, \widetilde{t}} \quad & \widetilde{t} \\
\text{s.t.} \quad & \mathbf{1}^T \mathbf{p} \le P_1 \\
& \widetilde{t}\left(\sum_{\substack{i=1 \\ i \ne k,j}}^{2K} p_i \mathrm{Tr}(\mathbf{Q}_k \mathbf{B}) + \mathrm{Tr}(\mathbf{N}\mathbf{B}) + \sigma_k^2 \right) \le \mathrm{Tr}(\mathbf{D}_k \mathbf{B}) p_k \\
& \sum_{k=1}^{2K} p_k \mathrm{Tr}(\mathbf{R}_k \mathbf{B}) + \mathrm{Tr}(\mathbf{B}) \le P_2 \\
& \sum_{k=1}^{2K} p_k \mathrm{Tr}(\mathbf{T}_k \mathbf{B}) + \sigma_p^2 \le I_n^R, \\
& \text{for } n = 1, 2 \dots N
\end{aligned}
$$

$$(3.5.18)$$

The problem (3.5.17) is convex and belongs to the class of GP [137] . As a result, it can be efficiently solved using interior point methods [138].

Based on the above we get the following algorithm that optimizes $\mathbf{p}$ and $\mathbf{b}$:

Table 3.5.   Optimization of the relays' beamforming vectors and sources' power

---

1) **Initialize p $\ge$ 0.**
2) **Repeat**
2a)   Solve problem (3.5.16) to obtain optimal value of **B**
        if **B** is rank-one, then the principal eigenvector
        of **B** is **b**
        if **B** is of higher rank we use randomization
        techniques to obtain **b**
2b)   Solve problem (3.5.18) to obtain optimal value of **p**,
      for given **b**
2c)   Update the old values of **p** before solving (3.5.16)
3) **Until** $\mathrm{SINR}_{\text{current}} - \mathrm{SINR}_{\text{previous}} \le \varepsilon$, where ε is a small positive constant for controlling the accuracy of the algorithm. The converged value of SINR is the optimal solution of SINR in (3.5.12).

3.6   Simulation Results

In order to assess the performance of the proposed scheme three numerical examples on the achievable mean SINR with different number of users, relays and relay powers are considered. For convenience, all PUs' and SUs' channel coefficients $\mathbf{h}_k$, $\mathbf{f}_n$ are modeled as Gaussian random variables. Figure 3.13 shows the mean SINR for different number of maximum allowable transmit power at the relays for a network with 2 secondary users on both sides, 2 relays, each of which is equipped with 2 antennas and 1 PU. Moreover, in all the simulations the noise power at the relays and at the transceivers is assumed to be equal to 1 and 0.1 respectively. It is observed that the achievable SINR is increased as the maximum allowable transmit power gets larger. In Figure 3.14, we show the achievable mean SINR for different number of secondary users, for a network with 2 relays, each equipped with 2 antennas and one PU. As the number of users decreases, the performance at the transceivers is improved. The maximum allowable transmit power at the relays is 10W. Finally in Figure 3.15 the mean SINR versus different number of relays is plotted. The maximum allowable transmit power at the relays is set to 5W and each of the relays is equipped with one antenna.
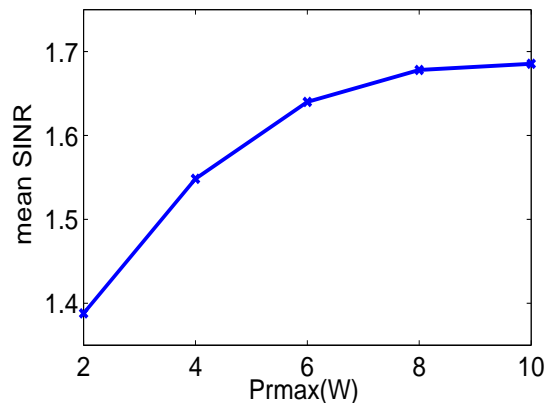
Figure 3.13. The mean SINR versus the maximum allowable relay transmit power P_2 at the relays for a network with 2 users on both sides, 2 relays, equipped with 2 antennas each and P_1=7W.
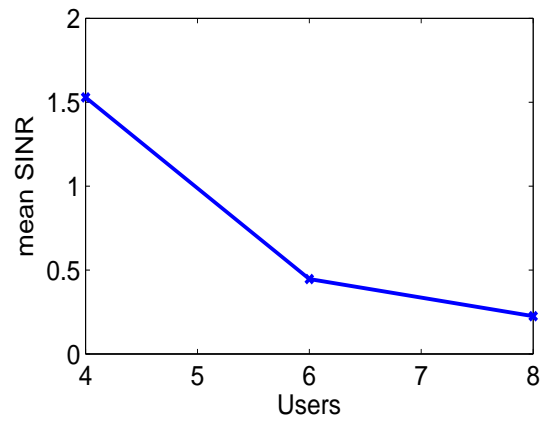
Figure 3.14. The mean SINR versus different number of users for a network with 2 relays, each with two antennas and maximum allowable transmit power P_2=10W and I_1^R=7W.
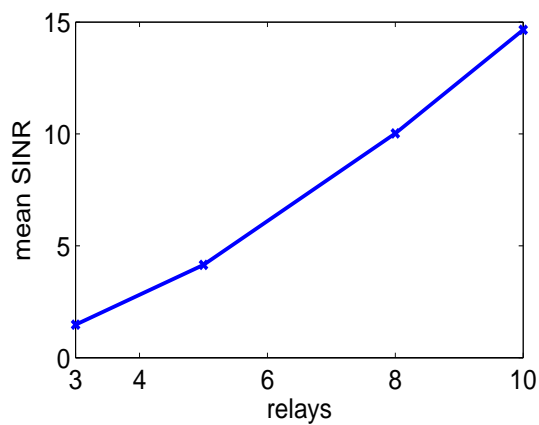
Figure 3.15. The mean SINR versus different number of relays for a network with maximum allowable transmit power P_2=5W and I_1^R=7W.

3.7 Conclusion

In this chapter, an SINR balancing based relay signal forwarding scheme was proposed for peer-to-peer networks. First, a one-way AF relay network with multiple users and single-antenna relay nodes was considered. Then, for

enhancing the total network throughput, a two-way AF relay network was considered. Additionally, in order to meet the needs for emerging and future generations of wireless networks, both sources and relays were assumed to be part of a cognitive radio network. More specifically, the cognitive terminals i.e., secondary users, are considered to access the spectrum occupied by the PUs, given that the interference they cause to the PUs is less than a certain threshold, such that PUs can be protected. This setup has also been extended for a cognitive two-way relay network and for relay nodes with multiple antennas in order to improve the quality of the received signal. It was demonstrated that the relay beamforming problem can be solved using SDP and the power allocation problem can be solved using GP. The final solution consists of an iterative procedure between the beamforming design and the power allocation. Simulation results have been provided to validate the performance of the proposed spatial multiplexing and power allocation techniques.

Chapter 4

TRANSMITTER-RECEIVER AND RELAY OPTIMIZATION FOR SPECTRUM SHARING MULTIPLE-INPUT AND MULTIPLE-OUTPUT PEER-TO-PEER USERS

In this chapter a spectrum sharing peer-to-peer relay network is considered where multiple source nodes with multiple antennas communicate with their desired destination nodes with multiple antennas through a MIMO relay. The work available in the literature on peer-to-peer network [146] considers multiple users with multiple antennas at the transmitters and receivers, however, only one data stream per user was considered. Hence beamformers are used at the transmitter and receiver. However, if multiple data streams are to be considered for each user, the transmitter and receiver processing units should be matrices instead of beamformers. Such a network with multiple data steams for each MIMO user is considered in this chapter. The focus is on optimizing the MMSE of the peer-to-peer MIMO relay network with linear transmit and receiver filters. In the first transmission phase K multiple-antenna sources transmit their signals through a spectrum shared channel and a multiple-antenna relay receives differently faded and noisy mixture of the source signals. In the second phase, the relay linearly pro-

cesses its received signals and then retransmits them to the destinations. The goal is to determine jointly the optimal precoding, receiver and relay processing matrices that minimize the weighted sum-MSE, while concurrently satisfying constraints in terms of the total transmission powers of the sources and relay.

As the first step to solve this optimization problem, it is shown that for any multipoint-to-multipoint MIMO network with linear beamformers, a dual equivalent network exists that attains the same MSE under a total power constraint. More specifically, the duality that preserves the individual users' MSEs in a peer-to-peer network is derived. Early work on this field focused on the MSE duality of the broadcast and multiple access channels [147], [148]. This property is then used to transform the optimization problem to an equivalent one that provides an iterative solution. The precoding and receiver filters for a given relaying matrix are shown to be the solutions to an alternating optimization-based algorithm, while for a given set of transmitter and receiver filters, the relaying matrix optimization is formulated using second order cone programming (SOCP).

4.1 System Model and Problem Formulation

A peer-to-peer network depicted in Figure 4.1 is considered that consists of K pairs of multiple-antenna transceivers and one MIMO relay. The source and the destination of the kth pair are equipped with $N_{s,k}$ and $N_{d,k}$ antennas respectively, whereas the relay has N_r antennas. The vector $\mathbf{s}_k \in \mathbb{C}^{M_k}$ contains the M_k uncorrelated unity variance symbols of user k. $\mathbf{U}_k \in \mathbb{C}^{N_{s,k} \times M_k}$ denotes the precoding matrix and $\mathbf{V}_k \in \mathbb{C}^{N_{d,k} \times M_k}$ is the receiver matrix of the kth user. No direct link between the sources and the destinations is assumed. In the first time slot, all sources transmit simultaneously. The
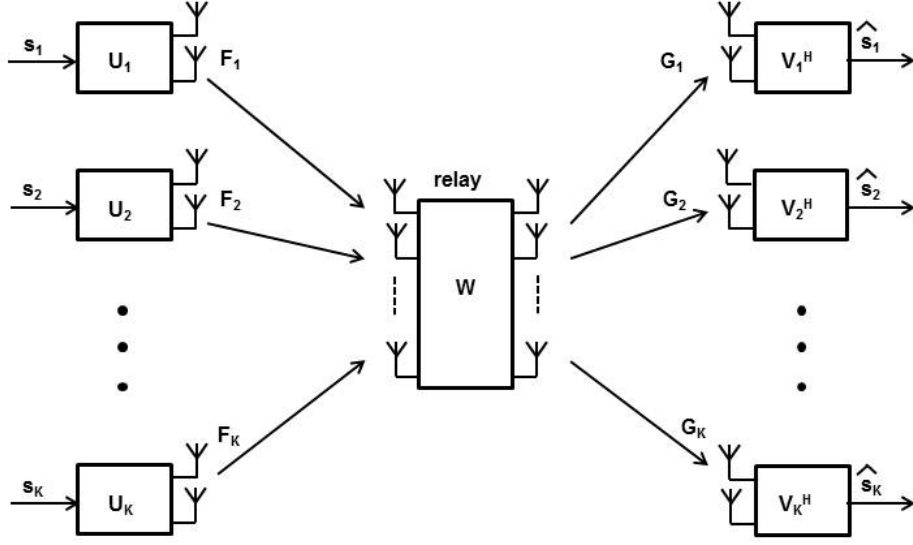
Figure 4.1. A MIMO relay with multiple peer-to-peer MIMO users

signal picked up by the relay is written as:

$$\mathbf{y}_r = \sum_{i=1}^{K} \mathbf{F}_i \mathbf{U}_i \mathbf{s}_i + \mathbf{n}_r \in \mathbb{C}^{N_r} \qquad (4.1.1)$$

where $\mathbf{F}_i \in \mathbb{C}^{N_r \times N_{s,k}}$ is the frequency flat MIMO channel of user k to the relay and $\mathbf{n}_r \in \mathbb{C}^{N_r \times 1}$ is the zero-mean additive white Gaussian noise (ZM-AWGN) vector at the relay with covariance matrix $\mathrm{E}\{\mathbf{n}_r \mathbf{n}_r^H\} = \sigma_{n_r}^2 \mathbf{I}_{N_r}$. The relay multiplies its received signal vector by a matrix $\mathbf{W} \in \mathbb{C}^{N_r \times N_r}$. Thus the signal vector transmitted by the relay can be written as:

$$\mathbf{x}_r = \mathbf{W} \mathbf{y}_r \qquad (4.1.2)$$

The received signal at the kth destination is written as

$$\mathbf{y}_k = \mathbf{G}_k \mathbf{W} \sum_{i=1}^{K} \mathbf{F}_i \mathbf{U}_i \mathbf{s}_i + \mathbf{G}_k \mathbf{W} \mathbf{n}_r + \mathbf{n}_k \qquad (4.1.3)$$

where $\mathbf{G}_k \in \mathbb{C}^{N_{d,k} \times N_r}$ is the frequency flat channel from the relay to the kth user and $\mathbf{n}_k \in \mathbb{C}^{N_{d,k} \times 1}$ is the zero mean additive white gaussian noise vector

at the kth destination with covariance matrix $\mathrm{E}\{\mathbf{n}_k\mathbf{n}_k^H\} = \sigma_{n_k}^2\mathbf{I}_{N_{d,k}}$. Since the noise at the relay is also propagated to the destination, the noise present at the receiver is coloured. This can be whitened by sending the received signal at the kth terminal through a whitening filter $(\mathbf{G}_k\mathbf{W}\mathbf{W}^H\mathbf{G}_k^H\sigma_{n_r}^2 + \sigma_{n_k}^2\mathbf{I}_{N_{d,k}})^{-\frac{1}{2}}$. Hence, the input-output relationship of the kth user can be written as

$$
\begin{aligned}
\mathbf{y}_k' &= (\mathbf{G}_k\mathbf{W}\mathbf{W}^H\mathbf{G}_k^H\sigma_{n_r}^2 + \sigma_{n_k}^2\mathbf{I}_{N_{d,k}})^{-\frac{1}{2}}\mathbf{G}_k\mathbf{W}\sum_{i=1}^{K}\mathbf{F}_i\mathbf{U}_i\mathbf{s}_i + \mathbf{n}_k' \\
&= \mathbf{G}_k'\mathbf{W}\sum_{i=1}^{K}\mathbf{F}_i\mathbf{U}_i\mathbf{s}_i + \mathbf{n}_k' \qquad\qquad (4.1.4)
\end{aligned}
$$

where $\mathbf{y}_k'$ is signal at the output of the noise whitening filter and $\mathbf{G}_k'$ and $\mathbf{n}_k'$ are the effective channel matrix and the noise vector corresponding to the kth user. Note that by formulation, $E\{\mathbf{n}_k'\mathbf{n}_k'^H\} = \mathbf{I}_{N_r}$.

The symbol estimate $\widehat{\mathbf{s}}_k \in \mathbb{C}^{M_k\times 1}$ generated by the linear filtering matrix $\mathbf{V}_k^H$ can be written as

$$
\widehat{\mathbf{s}}_k = \mathbf{V}_k^H\mathbf{G}_k'\mathbf{W}\sum_{i=1}^{K}\mathbf{F}_i\mathbf{U}_i\mathbf{s}_i + \mathbf{V}_k^H\mathbf{n}_k'. \qquad\qquad (4.1.5)
$$

Thus, the kth user MSE is written as

$$
\begin{aligned}
\varepsilon_k = \mathrm{E}\left\{||\mathbf{s}_k - \widehat{\mathbf{s}}_k||_2^2\right\} &= \mathrm{Tr}(\mathbf{I}_{M_k} - \mathbf{V}_k^H\mathbf{G}_k'\mathbf{W}\mathbf{F}_k\mathbf{U}_k - \mathbf{U}_k^H\mathbf{F}_k^H\mathbf{W}^H\mathbf{G}_k'^H\mathbf{V}_k \\
&+ \mathbf{V}_k^H\mathbf{G}_k'\mathbf{W}\sum_{i=1}^{K}\mathbf{F}_i\mathbf{U}_i\mathbf{U}_i^H\mathbf{F}_i^H\mathbf{W}^H\mathbf{G}_k'^H\mathbf{V}_k + \sigma_{k'}^2\mathbf{V}_k\mathbf{V}_k^H).
\end{aligned}
$$

Note that $\sigma_{k'}^2 = 1$, because $\mathrm{E}\{\mathbf{n}_k'\mathbf{n}_k'^H\} = \mathbf{I}_{N_r}$. We assume that the users' total transmit power is

$$
P_T = \mathrm{E}\left\{\sum_{i=1}^{K}||\mathbf{U}_i\mathbf{s}_i||_2^2\right\} = \sum_{i=1}^{K}||\mathbf{U}_i||_F^2, \qquad\qquad (4.1.6)
$$

and the relay's transmit power is

$$P_R = \mathrm{E}\{\mathbf{x}_r^H \mathbf{x}_r\} = \mathrm{Tr}\left(\mathbf{W}(\sum_{i=1}^{K} \mathbf{F}_i \mathbf{U}_i \mathbf{U}_i^H \mathbf{F}_i^H + \sigma_{n_r}^2 \mathbf{I}_{N_r})\mathbf{W}^H\right), \quad (4.1.7)$$

where we assumed that the symbol vectors $\mathbf{s}_1, \cdots, \mathbf{s}_K$ are mutually uncorrelated with identity covariance matrix, i.e., $\mathrm{E}\{\mathbf{s}_k \mathbf{s}_k^H\} = \mathbf{I}_{M_k}$ $\forall k$. The aim is to design transmit, relay and receiver filters so as to minimize the sum of the weighted user-wise MSEs of all users:

$$\min_{\mathbf{U}_1...\mathbf{U}_K, \mathbf{W}, \mathbf{V}_1...\mathbf{V}_K} \sum_{k=1}^{K} w_k \varepsilon_k$$

$$\text{s.t.} \sum_{i=1}^{K} ||\mathbf{U}_i||_F^2 \leq P_{Tmax} \text{ and } P_R \leq P_{Rmax} \quad (4.1.8)$$

where $w_1 \ldots w_K$ are certain positive weights. The power-constrained optimization problem is very difficult to solve directly, as all variables $\mathbf{U}_k$, $\mathbf{W}$, $\mathbf{V}_k^H$ are coupled in (4.1.8) and the problem is non-convex. For this reason, we use an iterative approach to obtain the precoder, receive filters and relay matrix. Specifically the transmit and receiver filters are first obtained for a given relay amplification matrix and then the relay matrix is obtained for a given set of transmit and receiver filters iteratively.

4.2 Linear Downlink-Uplink Duality in Peer-to-Peer Networks

In this subsection, in order to facilitate solving problem (4.1.8) with respect to $\mathbf{U}_k$ and $\mathbf{V}_k^H$, it will be shown that each user's MSE that is achievable in the downlink can also be achieved in the virtual uplink under the same total transmit power constraint using MSE duality.

For a fixed relaying matrix $\mathbf{W}$, the weighted user-wise MSE optimization problem (4.1.8) is a function of the transmit matrix $\mathbf{U}_k$ and the receive
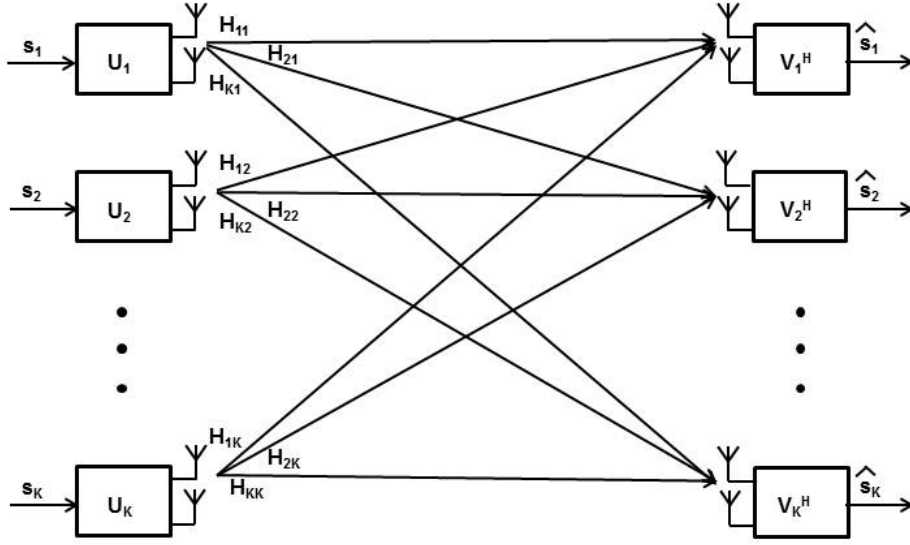
Figure 4.2. Multiple peer-to-peer MIMO users and downlink channel. The combined effect of the original channels and the relay transceiver matrix is shown with an equivalent set of channels $\mathbf{H}_{lk}$.

matrix $\mathbf{V}_k$ as follows,

$$\min_{\mathbf{U}_1...\mathbf{U}_K, \mathbf{V}_1...\mathbf{V}_K} \sum_{k=1}^{K} w_k \varepsilon_k$$

$$\text{s.t.} \sum_{i=1}^{K} ||\mathbf{U}_i||_F^2 \leq P_{Tmax} \tag{4.2.1}$$

4.2.1   Equivalent Downlink and Uplink Description

For a given relaying matrix $\mathbf{W}$, the $N_{d,k} \times N_{s,k}$ equivalent MIMO channel matrix from the ith source to the jth destination node can be denoted as $\mathbf{H}_{i,j} = \mathbf{G}_i' \mathbf{W} \mathbf{F}_j$ (see Figure 4.2). Substituting this relationship in (4.1.5), the estimate of the signal in the downlink can be written as

$$\widehat{\mathbf{s}}_k = \mathbf{V}_k^H \sum_{i=1}^{K} \mathbf{H}_{k,i} \mathbf{U}_i \mathbf{s}_i + \mathbf{V}_k^H \mathbf{n}_k' \tag{4.2.2}$$
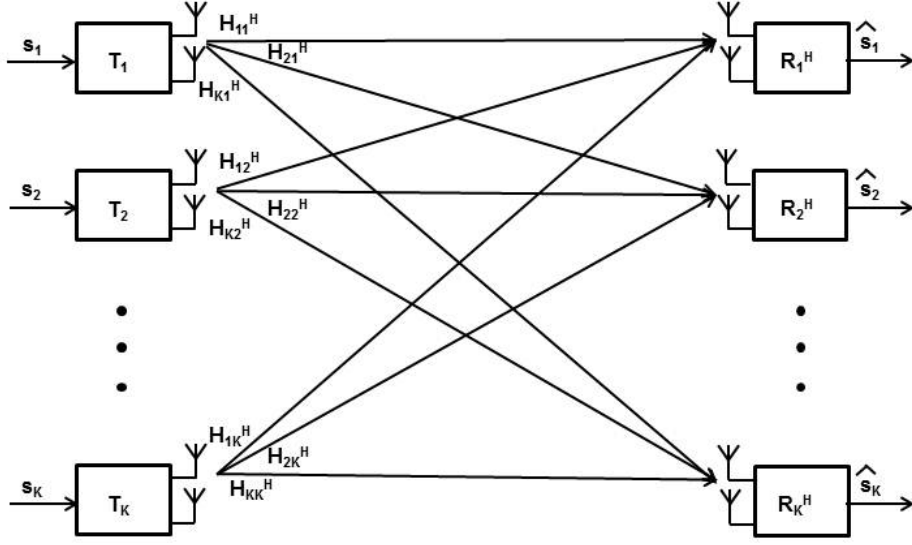
Figure 4.3. Multiple peer-to-peer MIMO users and the equivalent virtual uplink channel

We derive its downlink MSE $\varepsilon_k^{DL} = \mathrm{E}\,\{||\mathbf{s}_k - \widehat{\mathbf{s}}_k||_2^2\}$ as

$$
\begin{aligned}
\varepsilon_k^{DL} &= \mathrm{Tr}(\mathbf{I}_{M_k} - \mathbf{V}_k^H \mathbf{H}_{k,k} \mathbf{U}_k - \mathbf{U}_k^H \mathbf{H}_{k,k}^H \mathbf{V}_k + \mathbf{V}_k^H \sum_{i=1}^{K} \mathbf{H}_{k,i} \mathbf{U}_i \mathbf{U}_i^H \mathbf{H}_{k,i}^H \mathbf{V}_k \\
&\quad + \mathbf{V}_k \mathbf{V}_k^H)
\end{aligned}
\tag{4.2.3}
$$

The equivalent uplink model is obtained by reversing the direction and swapping the roles of the source and destination nodes as shown in Figure 4.3. For the virtual uplink, the precoders are denoted by $\mathbf{T}_k \in \mathbb{C}^{N_{d,k} \times M_k}$, and the channel from user k needs to be Hermitian transposed for matching the dimensions. In this case, the received signal at the kth destination can be written as:

$$
\mathbf{y}_k = \sum_{i=1}^{K} \mathbf{H}_{i,k}^H \mathbf{T}_i \mathbf{s}_i + \mathbf{n}_k'
\tag{4.2.4}
$$

The symbol estimate $\widehat{\mathbf{s}}_k$ of user k in the dual uplink is written as

$$
\widehat{\mathbf{s}}_k = \mathbf{R}_k^H \sum_{i=1}^{K} \mathbf{H}_{i,k}^H \mathbf{T}_i \mathbf{s}_i + \mathbf{R}_k^H \mathbf{n}_k'
\tag{4.2.5}
$$

where $\mathbf{R}_k \in \mathbb{C}^{N_{s,k} \times M_k}$ is the receiver filter at the destination. Thus, we obtain the uplink MSE $\varepsilon_k^{UL} = \mathrm{E}\{||\mathbf{s}_k - \widehat{\mathbf{s}}_k||_2^2\}$ of user k as

$$\varepsilon_k^{UL} = \mathrm{Tr}(\mathbf{I}_{M_k} - \mathbf{R}_k^H \mathbf{H}_{k,k}^H \mathbf{T}_k - \mathbf{T}_k^H \mathbf{H}_{k,k} \mathbf{R}_k +$$
$$\mathbf{R}_k^H \sum_{i=1}^{K} \mathbf{H}_{i,k}^H \mathbf{T}_i \mathbf{T}_i^H \mathbf{H}_{i,k} \mathbf{R}_k + \mathbf{R}_k \mathbf{R}_k^H) \qquad (4.2.6)$$

In the uplink the same average transmission power should be used, i.e.,

$$P_{Tmax} = \mathrm{E}\{\sum_{i=1}^{K} ||\mathbf{T}_i \mathbf{s}_i||_2^2\} = \sum_{i=1}^{K} ||\mathbf{T}_i||_F^2 \qquad (4.2.7)$$

4.2.2   Uplink-Downlink Conversion

The switching of the roles between the uplink and the downlink is depicted by interchanging every transmit filter in the uplink with the respective receive filter in the downlink and by choosing the transmit filters from the downlink as the respective receive filters in the uplink (see Figure 4.3). As the dual domain has to consume the same amount of transmit power, we need to weight every user's precoder with a scalar a_k and the respective receive filter with the reciprocal of this scalar, such that (4.1.6) and (4.2.7) hold simultaneously.

In this particular section, we construct an equivalent downlink setup for an uplink channel. Given arbitrary $\mathbf{R}_k^H$ and $\mathbf{T}_k$ for all users $k = 1, \ldots, K$, we derive matrices $\mathbf{U}_k$ and $\mathbf{V}_k$ such that users' MSEs do not change in the uplink and in the downlink.

In order to leave users' MSE unchanged, different scalars $a_k \in \Re_{+,0} \in \{1, \ldots, K\}$ have to be assigned to different users. For this reason, we set the respective filters in the dual downlink as:

$$\mathbf{U}_k = a_k \mathbf{R}_k \text{ and } \mathbf{V}_k = \frac{1}{a_k} \mathbf{T}_k \qquad (4.2.8)$$

Substituting (4.2.8) in the MSE ε_k^{DL} expression in (4.2.3) and equating with (4.2.6) we obtain

$$\varepsilon_k^{UL} = \varepsilon_k^{DL} \iff \sum_{i=1}^{K} ||\mathbf{R}_k^H \mathbf{H}_{i,k}^H \mathbf{T}_i||_F^2 + ||\mathbf{R}_k||_F^2$$

$$= \sum_{i=1}^{K} \frac{a_i^2}{a_k^2} ||\mathbf{R}_i^H \mathbf{H}_{k,i}^H \mathbf{T}_k||_F^2 + \frac{1}{a_k^2} ||\mathbf{T}_k||_F^2 \quad \forall k \qquad (4.2.9)$$

We can rewrite (4.2.9) as:

$$a_k^2 \left[\sum_{i=1,i\neq k}^{K} ||\mathbf{R}_k^H \mathbf{H}_{i,k}^H \mathbf{T}_i||_F^2 + ||\mathbf{R}_k||_F^2 \right]$$

$$- \sum_{i=1,i\neq k}^{K} a_i^2 ||\mathbf{R}_i^H \mathbf{H}_{k,i}^H \mathbf{T}_k||_F^2 = ||\mathbf{T}_k||_F^2 \quad \forall k \qquad (4.2.10)$$

In matrix form (4.2.9) can be rewritten as

$$\mathbf{X}\mathbf{a} = \mathbf{T} \qquad (4.2.11)$$

where

$$[\mathbf{X}]_{k,j} = \begin{cases} \sum_{i=1,i\neq k}^{K} ||\mathbf{R}_k^H \mathbf{H}_{i,k}^H \mathbf{T}_i||_F^2 + ||\mathbf{R}_k||_F^2 & k = j \\ -||\mathbf{R}_i^H \mathbf{H}_{k,i}^H \mathbf{T}_k||_F^2 & k \neq j, \end{cases}$$

$\mathbf{a} = [a_1 \ldots a_K]^T$ and $\mathbf{T} = [||\mathbf{T}_1||_F^2 \ldots ||\mathbf{T}_K||_F^2]^T$. From (4.2.11), it is shown that there exists $a_k \in \Re_+^*$, such that the same MSE target can be achieved in the downlink. This is because $\mathbf{X}$ is a Z-matrix, real-valued with off-diagonal entries less than zero, and also strictly column diagonally dominant [149]. Consequently, $\mathbf{X}$ is non-singular and and its inverse has only nonnegative entries. As a result, the desired weight vector $\mathbf{a}$ is given by:

$$\mathbf{a} = \mathbf{X}^{-1} \mathbf{T} \qquad (4.2.12)$$

Summing up all users' MSEs and by means of (4.2.8), we obtain

$$\sum_{i=1}^{K} a_i^2 ||\mathbf{R}_i||_F^2 = \sum_{i=1}^{K} ||\mathbf{T}_i||_F^2 \iff \sum_{i=1}^{K} ||\mathbf{U}_i||_F^2 = \sum_{i=1}^{K} ||\mathbf{T}_i||_F^2 \qquad (4.2.13)$$

The left-hand side (LHS) of (4.2.13) is the transmitted power in the downlink, while the right-hand side (RHS) of (4.2.13) is the transmitted power in the uplink, which means that by setting $\mathbf{U}_k = a_k \mathbf{R}_k$ and $\mathbf{V}_k = \frac{1}{a_k} \mathbf{T}_k$ and a total power threshold P_{Tmax}, we find the transmit and receiver filters in the uplink that achieve the same MSE values that were obtained with the filters $\mathbf{U}_k$ and $\mathbf{V}_k$ in the downlink.

4.2.3   Downlink-Uplink Conversion

In the previous section, we have shown that the MSE region in the uplink is also a subset of the MSE region in the downlink. In order to establish the duality, we have to show that the MSE region in the downlink is a subset of the MSE region in the uplink.

Given the precoders $\mathbf{U}_k$ and the receive filter matrices $\mathbf{V}_k$ in the downlink, the respective filters in the uplink are set to

$$\mathbf{R}_k = \frac{1}{\widetilde{a}_k} \mathbf{U}_k \text{ and } \mathbf{T}_k = \widetilde{a}_k \mathbf{V}_k \qquad (4.2.14)$$

Inserting this in the uplink expression (4.2.6) and equating with (4.2.3) we obtain

$$\varepsilon_k^{DL} = \varepsilon_k^{UL} \iff \sum_{i=1}^{K} ||\mathbf{V}_k^H \mathbf{H}_{k,i} \mathbf{U}_i||_F^2 + ||\mathbf{V}_k||_F^2$$

$$= \sum_{i=1}^{K} \frac{\widetilde{a}_i^2}{\widetilde{a}_k^2} ||\mathbf{V}_i^H \mathbf{H}_{i,k} \mathbf{U}_k||_F^2 + \frac{1}{\widetilde{a}_k^2} ||\mathbf{U}_k||_F^2 \quad \forall k \qquad (4.2.15)$$

Rearranging (4.2.15) and using $||(.)||_F = ||(.)^H||_F$, we obtain

$$\mathbf{Y}\widetilde{\mathbf{a}} = \sigma_{k'}^2 \mathbf{U} \qquad (4.2.16)$$

and similarly with (4.2.12) we have:

$$\widetilde{\mathbf{a}} = \mathbf{Y}^{-1}\mathbf{U} \qquad (4.2.17)$$

where

$$[\mathbf{Y}]_{k,j} = \begin{cases} \sum_{i=1,i\neq k}^{K} ||\mathbf{V}_k^H \mathbf{H}_{k,i}\mathbf{U}_i||_F^2 + ||\mathbf{V}_k||_F^2 & k = j \\ \\ -||\mathbf{V}_i^H \mathbf{H}_{i,k}\mathbf{U}_k||_F^2 & k \neq j, \end{cases}$$

$\widetilde{\mathbf{a}} = [\widetilde{a}_1 \ldots \widetilde{a}_K]^T$, $\mathbf{U} = [||\mathbf{U}_1||_F^2 \ldots ||\mathbf{U}_K||_F^2]^T$ and $\widetilde{\mathbf{a}} \, \epsilon \, \Re_+^*$ due to the properties of the matrix $\mathbf{Y}$.

Summing up all rows of (4.2.15), we obtain

$$\sum_{i=1}^{K} \widetilde{a}_i^2 ||\mathbf{V}_i||_F^2 = \sum_{i=1}^{K} ||\mathbf{U}_i||_F^2 \iff \sum_{i=1}^{K} ||\mathbf{T}_i||_F^2 = \sum_{i=1}^{K} ||\mathbf{U}_i||_F^2 \qquad (4.2.18)$$

This means that the total power consumption is the same during the conversion from the downlink to the uplink. It is proven that both links, downlink and uplink, have the same achievable MSE region under a sum power constraint.

4.3   Uplink and Downlink MMSE Receiver Filter Matrices

In this subsection, the uplink and downlink MMSE receive filters are obtained for a given set of transmit filters in both the links. Please note that the relay transceiver has already been absorbed into the channel $\mathbf{H}_{i,j}$.

In the downlink channel, for a fixed transmit precoder $\mathbf{U}_k$, the optimum

linear receiver $\mathbf{V}_k$, that minimizes each user's MSE ε_k is:

$$\mathbf{V}_k = \mathbf{Z}_{DLk}^{-1}\mathbf{H}_{k,k}\mathbf{U}_k \ \forall k \tag{4.3.1}$$

where $\mathbf{Z}_{DLk} = \sum_{i=1}^{K}\mathbf{H}_{k,i}\mathbf{U}_i\mathbf{U}_i^H\mathbf{H}_{k,i}^H + \mathbf{I}$. The MMSE of user k, when it uses the optimum $\mathbf{V}_k$ is then written as

$$\varepsilon_k^{DL}(\mathbf{U}) = \mathrm{Tr}(\mathbf{I} - \mathbf{U}_k^H\mathbf{H}_{k,k}^H\mathbf{Z}_{DLk}^{-1}\mathbf{H}_{k,k}\mathbf{U}_k) \tag{4.3.2}$$

Similarly, for the virtual uplink channel, for a fixed transmit filter $\mathbf{T}_k$, the MMSE receiver filter is given by

$$\mathbf{R}_k = \mathbf{Z}_{ULk}^{-1}\mathbf{H}_{k,k}^H\mathbf{T}_k \ \forall k, \tag{4.3.3}$$

where $\mathbf{Z}_{ULk} = \sum_{i=1}^{K}\mathbf{H}_{i,k}^H\mathbf{T}_i\mathbf{T}_i^H\mathbf{H}_{i,k} + \mathbf{I}$. The kth user MSE is written as

$$\varepsilon_k^{UL}(\mathbf{T}) = \mathrm{Tr}(\mathbf{I} - \mathbf{T}_k^H\mathbf{H}_{k,k}\mathbf{Z}_{ULk}^{-1}\mathbf{H}_{k,k}^H\mathbf{T}_k). \tag{4.3.4}$$

4.4   MSE Alternating Optimization Framework for Peer-to-Peer Networks

By applying the above shown duality for a peer-to-peer network, MSE optimization problem can be solved by optimizing the MSE values of the equivalent uplink system. However, in peer-to-peer networks, complexity of the equivalent uplink problem is still high, since downlink MSEs, $\varepsilon_1^{DL}, \ldots, \varepsilon_K^{DL}$ are all coupled by the choice of the transmit filters $\mathbf{U}_1, \ldots, \mathbf{U}_K$, while the uplink MSEs are coupled by the filters $\mathbf{T}_1, \ldots, \mathbf{T}_K$. For this reason, the filters using an alternating manner by switching between the equivalent uplink

and downlink channels are optimized. The algorithm in this case is summarized in Table I. Here the superscript (n) denotes the number of iteration number.

Algorithm 1 Table I: Procedure for solving the problem in (4.2.1) via alternating optimization between the uplink and the downlink

---

1. *initialize*: $\mathbf{T}_k^0 = (\sqrt{P_{max}/M_k K})\mathbf{I}, \forall k$ and choose maximal number of iterations n_{max}

2. **Repeat**

3. $n \leftarrow n + 1$

4. Uplink channel:

   (a) For a given set of $\mathbf{T}_k^{(n-1)}, \forall k$, update $\mathbf{R}_k^{(n)}, \forall k$ using (4.3.3).

   (b) Compute weight vector $\mathbf{a}^{(n)}$ using (4.2.12).

5. Downlink channel:

   (a) Update $\mathbf{U}_k^{(n)}, \forall k$ using (4.2.8).

   (b) Compute receive filters $\mathbf{V}_k^{(n)}, \forall k$ using (4.3.1).

   (c) Compute weight vector $\widetilde{\mathbf{a}}^{(n)}$ using (4.2.17).

   (d) Update transmit filters $\mathbf{T}_k^{(n)}$ using (4.2.14).

   (e) Compute the weighted sum MSE, and denote this ε_n

6. **Until** $|\varepsilon_n - \varepsilon_{n-1}| < \xi_1$, where ξ_1 is stopping threshold.

4.5   Weighted Sum-MSE Optimization

In this section, it is demonstrated how the proposed duality can be applied to the weighted sum MSE optimization problem (4.1.8). The matrices, $\mathbf{U}_k, \forall k$, $\mathbf{V}_k, \forall k$ and $\mathbf{W}$ that minimize the weighted sum-MSE as in (4.1.8), subject to a transmit sum-power constraint $P_T \leq P_{Tmax}$ and a relay-power constraint $P_R \leq P_{Rmax}$ cannot be found using a closed form. However, given the relay matrix $\mathbf{W}$, the transmit matrix $\mathbf{U}_k, \forall k$ and the receive filters $\mathbf{V}_k, \forall k$ can be calculated by using the above alternating optimization algorithm, based on

the MSE duality. Then for a given set of transmit and receive filters, the optimum relaying matrix $\mathbf{W}$ can be obtained using interior point methods, as this problem is convex.

4.5.1   Deriving Transmit and Receiver Filters for a fixed Relaying Matrix

The aim in this section is to minimize the weighted sum-MSE of a peer-to-peer relay network for a fixed relaying matrix. The transmit and receive filters cannot be optimized jointly. This leads to optimizing the filters using an alternating way by switching between the uplink and the downlink. The scaling factors from (4.2.12) and (4.2.17) need not to be calculated from the duality according to the previous section. They can be calculated using the Lagrangian function associated with (4.2.1). The motivation for this approach is to achieve lower computational complexity.

The Lagrangian function of the optimization problem (4.2.1) is:

$$
L(\mathbf{U}, \mu) = \sum_{k=1}^{K} w_k M_k - \sum_{k=1}^{K} w_k \mathrm{Tr}(\mathbf{U}_k^H \mathbf{H}_{k,k}^H \mathbf{Z}_{DLk} \mathbf{H}_{k,k} \mathbf{U}_k)
$$
$$
+ \mu \left(\sum_{k=1}^{K} ||\mathbf{U}_k||_F^2 - P_{Tmax} \right) \tag{4.5.1}
$$

where $\mu > 0$ is the Lagrange multiplier associated with the total transmit power constraint. It is apparent that the precoding filters $\mathbf{U}_k$ satisfy the Karush-Kuhn-Tucker (KKT) conditions of (4.2.1). So differentiating (4.5.1) with respect to $\mathbf{U}_k^*$ yields:

$$
\frac{\partial L}{\partial \mathbf{U}^*} = \mu \mathbf{U}_k - w_k \mathbf{H}_{k,k}^H \mathbf{Z}_{DLk}^{-1} \mathbf{H}_{k,k} \mathbf{U}_k
$$
$$
+ \sum_{i=1}^{K} w_i \mathbf{H}_{i,k}^H \mathbf{Z}_{DLi}^{-1} \mathbf{H}_{i,i} \mathbf{U}_i \mathbf{U}_i^H \mathbf{H}_{i,i}^H \mathbf{Z}_{DLi}^{-1} \mathbf{H}_{i,k} \mathbf{U}_k \tag{4.5.2}
$$

According to the first-order KKT condition:

$$\mu \mathbf{U}_k = w_k \mathbf{H}_{k,k}^H \mathbf{Z}_{DLk}^{-1} (\sum_{\substack{i=1 \\ i \neq k}}^{K} \mathbf{H}_{k,i} \mathbf{U}_i \mathbf{U}_i^H \mathbf{H}_{k,i}^H + \mathbf{I}) \mathbf{Z}_{DLk}^{-1} \mathbf{H}_{k,k} \mathbf{U}_k$$

$$- \sum_{\substack{i=1 \\ i \neq k}}^{K} w_i \mathbf{H}_{i,k}^H \mathbf{Z}_{DLi}^{-1} \mathbf{H}_{i,i} \mathbf{U}_i \mathbf{U}_i^H \mathbf{H}_{i,i}^H \mathbf{Z}_{DLi}^{-1} \mathbf{H}_{i,k} \mathbf{U}_k \qquad (4.5.3)$$

Multiplying (4.5.3) with $\mathbf{U}_k^H$ from the left and taking its trace, it can be noticed that w_k satisfies the same linear system of equations as the scalars $\widetilde{a}_k$ of the downlink-uplink conversion except from a constant $\frac{\mu}{\sigma_{k'}^2}$. Therefore, the following relation is obtained

$$\frac{\widetilde{a}_1^2}{w_1} = \frac{\widetilde{a}_2^2}{w_2} = \ldots = \frac{\widetilde{a}_K^2}{w_K} \qquad (4.5.4)$$

From (4.5.4), (4.2.16) does not need to be solved for the downlink-uplink transformation, as $\widetilde{a}_1, \widetilde{a}_2, \ldots, \widetilde{a}_k$ have already been determined by the relationship $\widetilde{a}_k = \widetilde{a}_0 \sqrt{w_k}$, where $\widetilde{a}_0$ is chosen such that $\mathbf{U}_k$ satisfies the transmit power constraint in the downlink. In the same way, using the KKT conditions of the weighted sum-MSE for the uplink and following the same approach, $a_1, a_2, \ldots, a_k$ can be determined by setting $a_k = a_0 \sqrt{w_k}$, where the coefficient a_0 is determined from the transmit power constraint in the uplink.

4.5.2   Deriving Relay Transceiver for fixed Transmit and Receiver Filters

For a given transmit filter $\mathbf{U}_k$ and receive filter $\mathbf{V}_k$, the relaying matrix $\mathbf{W}$ can be obtained by reformulating problem (4.2.1) as:

$$\min_{\mathbf{W}} \quad \sum_{k=1}^{K} w_k \Big[\mathrm{Tr}(\mathbf{I}_{M_k}) - 2\mathrm{real}\Big(\mathrm{Tr}(\mathbf{V}_k^H \mathbf{G}_k' \mathbf{W} \mathbf{F}_k \mathbf{U}_k) \Big) +$$
$$\mathrm{Tr}(\mathbf{V}_k^H \mathbf{G}_k' \mathbf{W} \mathbf{R}_r \mathbf{W}^H \mathbf{G}_k'^H \mathbf{V}_k) \Big]$$
$$\text{s.t.} \quad \mathrm{Tr}(\mathbf{W}\mathbf{R}_r\mathbf{W}^H) \leq P_{Rmax} \qquad\qquad (4.5.5)$$

where $\mathbf{R}_r = (\sum_{i=1}^{K} \mathbf{F}_i \mathbf{U}_i \mathbf{U}_i^H \mathbf{F}_i^H + \sigma_{n_r}^2 \mathbf{I})$ and the relay power is denoted as $\mathrm{P}_r = \mathrm{E}\{\mathbf{x}_r^H \mathbf{x}_r\} = \mathrm{Tr}(\mathbf{W}\mathbf{R}_r\mathbf{W}^H)$. By setting $\mathbf{w} = \mathrm{Vec}(\mathbf{W})$, where Vec is an operator that forms a vector by stacking the columns of a matrix and by using the matrix identity: $\mathrm{trace}(\mathbf{A}^T\mathbf{H}) = (\mathrm{Vec}(\mathbf{A}))^T \mathrm{Vec}(\mathbf{H})$, (4.5.5) can be equivalently expressed as:

$$\min_{\mathbf{w}} \quad \sum_{k=1}^{K} w_k \Big[M_k - 2\mathrm{real}\Big(\mathrm{Vec}(\mathbf{G}_k^T \mathbf{V}_k^* \mathbf{U}_k^T \mathbf{F}_k^T)\mathbf{w} \Big)$$
$$+ \mathbf{w}^H \Big((\mathbf{R}_r^{\frac{1}{2}} \otimes \mathbf{V}_k^T \mathbf{G}_k')^T (\mathbf{R}_r^{\frac{T}{2}} \otimes \mathbf{V}_k^H \mathbf{G}_k) \Big) \mathbf{w} \Big]$$
$$\text{s.t.} \quad \mathbf{w}^H (\mathbf{R}_r^{\frac{1}{2}} \otimes \mathbf{I})^T (\mathbf{R}_r^{\frac{T}{2}} \otimes \mathbf{I})\mathbf{w} \leq P_{Rmax} \qquad (4.5.6)$$

where $\otimes$ denotes the Kronecker product. The above problem is a SOCP and can be solved using interior point methods [138].

Now the original weighted sum-mse minimization problem (4.2.1) can be solved using an iterative algorithm as shown in Table II. Based on the above, the overall algorithm to determine $\mathbf{U}_k$, $\mathbf{V}_k$ and $\mathbf{W}_k$ is summarized as below.

Algorithm 2 Table II: Weighted Sum-MSE Algorithm for MIMO Peer-to-Peer Relay Network via Alternating Optimization between Uplink and Downlink and SOCP Optimization

1. *initialize*: $\mathbf{T}_k^0 = (\sqrt{P_{max}/M_k K})\mathbf{I}, \forall k$ and $\mathbf{W}^0 = b\mathbf{I}$, where b is chosen to satisfy the relay power constraint. Choose the maximal number of iterations n_{max}, m_{max}.

2. **Repeat**

3. $m \leftarrow m + 1$

   (a) **Repeat**

   (b) $n \leftarrow n + 1$

   (c) Uplink channel:

      i. For a given set of precoder $\mathbf{T}_k^{(n-1)}, \forall k$ and relay matrix $\mathbf{W}^{(n-1)}$, find $\mathbf{R}_k^{(n)}, \forall k$ using (4.3.3).

      ii. For Uplink/Downlink conversion compute $a_0 = \sqrt{\dfrac{P_T}{\sum_i \operatorname{Tr}(w_i \mathbf{T}_i^H \mathbf{H}_{i,i} \mathbf{Z}_{ULi}^{-2} \mathbf{H}_{i,i}^H \mathbf{T}_i)}}$

      iii. Find $a_k^{(n)} = a_0 \sqrt{w_k}, \forall k$

   (d) Downlink channel:

      i. Update $\mathbf{U}_k^{(n)}, \forall k$ using (4.2.8).

      ii. Compute receive filters $\mathbf{V}_k^{(n)}, \forall k$ using (4.3.1).

      iii. For Downlink/Uplink conversion compute $\widetilde{a}_0 = \sqrt{\dfrac{P_T}{\sum_i \operatorname{Tr}(w_i \mathbf{U}_i^H \mathbf{H}_{i,i} \mathbf{Z}_{DLi}^{-2} \mathbf{H}_{i,i}^H \mathbf{U}_i)}}$

      iv. Find $a_k^{(n)} = \widetilde{a}_0 \sqrt{w_k}, \forall k$

      v. Update transmit filters $\mathbf{T}_k^{(n)}$ using (4.2.14).

   (e) **Until** $|\varepsilon_n - \varepsilon_{n-1}| < \xi_1$

4. For a given transmit filter $U_k^{(m-1)}$ and receive filter $V_k^{(m-1)}$, compute $\mathbf{W}^{(m)}$ using (4.5.6)

5. Compute the weighted sum MMSE and denote this β_n

6. **Until** $|\beta_n - \beta_{n-1}| < \xi_2$, where ξ_2 is a stopping threshold.

4.6   Convergence Analysis

In this section, the convergence of the algorithm shown in Table II is demonstrated. Consider an arbitrary iteration number n. First consider the inner repeat loop (3) for the design of transmitter and receiver filters. Step 3c returns the receiver filters $\mathbf{R}_k^{(n)}, \forall k$, which were aimed at minimizing the weighted sum-MSE for given transmit filters $\mathbf{T}_k^{(n-1)}, \forall k$, relay matrix $\mathbf{W}^{(n-1)}$ and a total power limit P_{max}. Let the minimum weighted sum MSE obtained at this stage $\epsilon^{(n,step3c)}$.

At step 3d(i) and the design of the transmit filters $\mathbf{U}_k^{(n)}, \forall k$, the MSE duality ensures that the same weighted sum MSE $\epsilon^{(n,step3c)}$ can be achieved with the same total transmit power P_{max} in the downlink. In step 3d(ii), using these transmit filters $\mathbf{U}_k^{(n)}, \forall k$, the receiver filters $\mathbf{V}_k^{(n)}, \forall k$, have been designed so that the minimum weighted sum-MSE value should be further reduced, i.e. $\epsilon^{(n,step3d(ii))} \leq \epsilon^{(n,step3c)}$.

In step 3d(v), the transmit filters $\mathbf{T}_k^n, \forall k$ and the same receiver filters $\mathbf{R}_k^n, \forall k$ achieve the same weighted sum-MSE, $\epsilon^{(n,step3d(v))} = \epsilon^{(n,step3d(ii))}$.

For the subsequent iteration $n+1$, step3c minimizes further the weighted sum-MSE by choosing optimal $\mathbf{R}_k^{(n+1)}, \forall k$ for given $\mathbf{T}_k^{(n)}, \forall k$, $\epsilon^{(n+1,step3c)} \leq \epsilon^{(n,step3d(v))}$.

As a result, the weighted sum-MSE value monotonically decreases at each iteration. However, since there is a total power limit, the weighted sum-MSE ϵ converges to a limit as $n \to \infty$. It was shown so far that for the inner-loop, at each iteration the sum-MSE decreases monotonically and converges. Let us now look at the outer loop for the design of the relay.

At step 4, for a given set of transmit filters $\mathbf{U}_k^{(m-1)}, \forall k$ and receiver filters $\mathbf{V}_k^{(m-1)}, \forall k$, the optimal relay filter $\mathbf{W}^{(m)}, \forall k$ is chosen so that the sum-MSE value is reduced further, i.e. $\epsilon^{(m,step4)} \leq \epsilon^{(m-1,step4)}$. Since there is a relay power constraint, the weighted sum-MSE value ϵ must converge to a limit

as $m \to \infty$. This concludes the proof that the weighted sum-MSE decreases monotonically and the iterative algorithm converges.

However, since the overall problem (4.1.8) is not convex, the global optimality of Algorithm I cannot be guaranteed. In general, different initializations may affect the convergence speed and the minimum weighted sum-MSE value of Algorithm I. The initial values of transmit filters $\mathbf{T}_k, \forall k$ can be chosen as matrices containing the right singular vectors of the channels for better convergence instead of choosing random matrices.

4.7   Simulation Results

A number of simulations was performed to characterize the performance of the proposed scheme under various scenarios. In all simulations, the elements of the channel matrices $\mathbf{F}_k$, $\mathbf{H}_k$ and $\mathbf{G}_k$ are assumed to be circularly symmetric complex Gaussian random variables with zero mean and unity variance per dimension. Also, in all simulations, we have assumed two peer-to-peer users, each employing two antennas at the transmitter and receiver, hence allowing two simultaneous data streams for each user.

For the first simulation, the total power at the transmitters was set to unity and the power available at the relay to 10W. A set of random channels was generated and the achieved MMSE was computed for different number of antennas at the relay as shown in Figure 4.4. The results have been shown for two different values of noise variances 0.01W and 0.001W. Both relay and user terminals have identical noise variances. To generate results for different number of relay antennas, the simulation started by choosing random channel matrices $\mathbf{F}_k$, $\mathbf{H}_k$ and $\mathbf{G}_k$ for the case of six antennas at the relay and then the rest of the simulations has been performed by reducing the dimensions of the initial matrices according to the number of relay an-
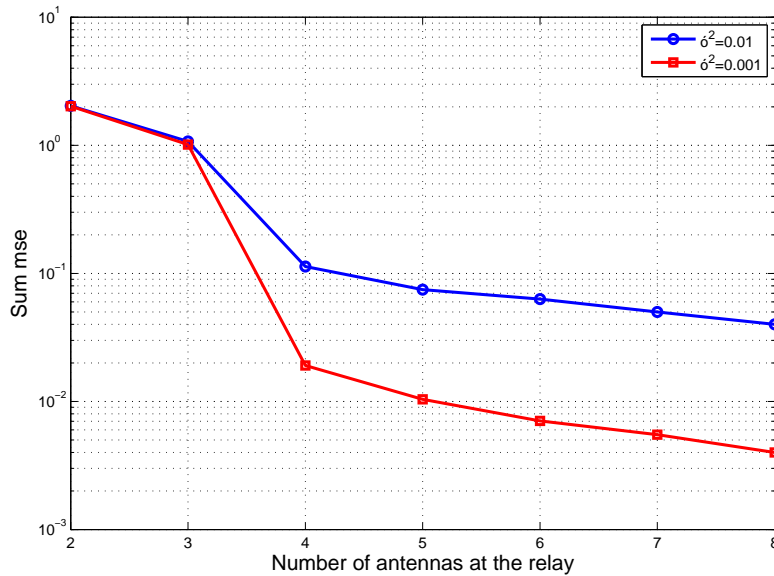
Figure 4.4. The sum MMSE against number of antennas at the relay for a network with 2 users on both sides equipped with 2 antennas each. P_T=1W, P_{Rmax}=10W

tennas required. The stopping criterion for the iteration was $\xi_1, \xi_2 = 0.005$ and approximately four to eight iterations have been observed. As seen in Figure 4.4, the MMSE value decreases as the number of antennas at the relay increases. As there are two users transmitting two data streams each, the relay needs at least four antennas to perform satisfactory spatial multiplexing, as such the MMSE value drops significantly beyond the use of four antennas. When there is adequate number of antennas at the relays, the achieved MMSE is of the order of the variance of the noise, confirming satisfactory performance of the proposed iterative method.

According to the proposed iterative method, the relay matrix had to be initialized when designing the transmitter and receiver filters in the first iteration. For the above simulation, the relay matrix $\mathbf{W}$ has been initialized with randomly generated circularly symmetric complex Gaussian variables. However, as the overall problem is non-convex, the algorithm is not expected to converge to the same final MMSE value for different initializations. In
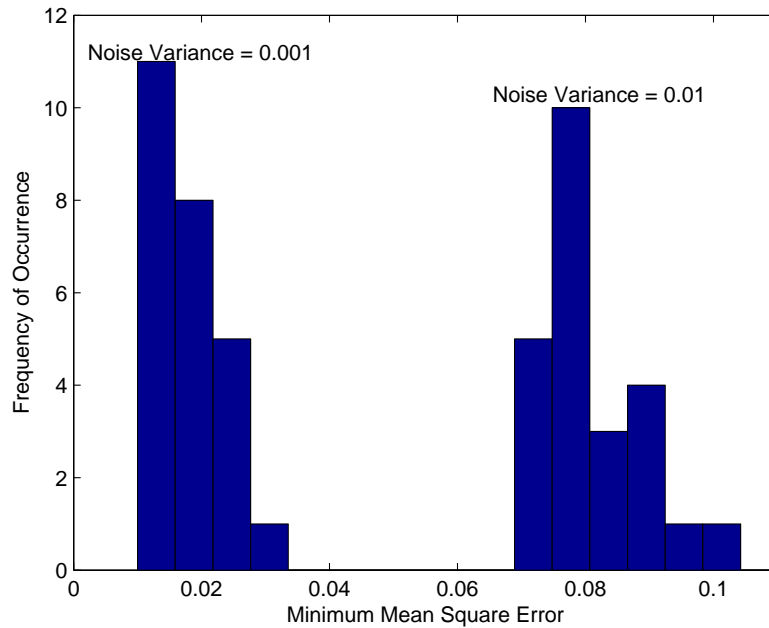
Figure 4.5. The histogram of the final sum MMSE due to different random initializations of the relay matrix. The results are shown for two different noise variances 0.01W and 0.001W.

order to investigate the susceptibility of the algorithm for different initializations, the above simulations have been performed for 25 different random initializations and the histogram of the final sum MMSE values obtained has been plotted. The number of antennas at the relay is five. The histograms are shown separately for two different noise variances 0.01W and 0.001W. As seen in Figure 4.5, the final MMSE values differ only slightly for different random initializations of the relay matrices. Moreover, the higher MMSE values occur only with small probabilities.

Finally, the sum MMSE for various values of SNR has been computed. The simulation scenario was set the same as before, i.e. two users, two antennas at the transmitter and receiver and five antennas at the relay. The SNR (SNR1) of the link from the transmitters to the relay has been set at 30dB and the SNR (SNR2) of the link between the relay and the users has been varied. SNR1 has been defined as the ratio of the total transmitter
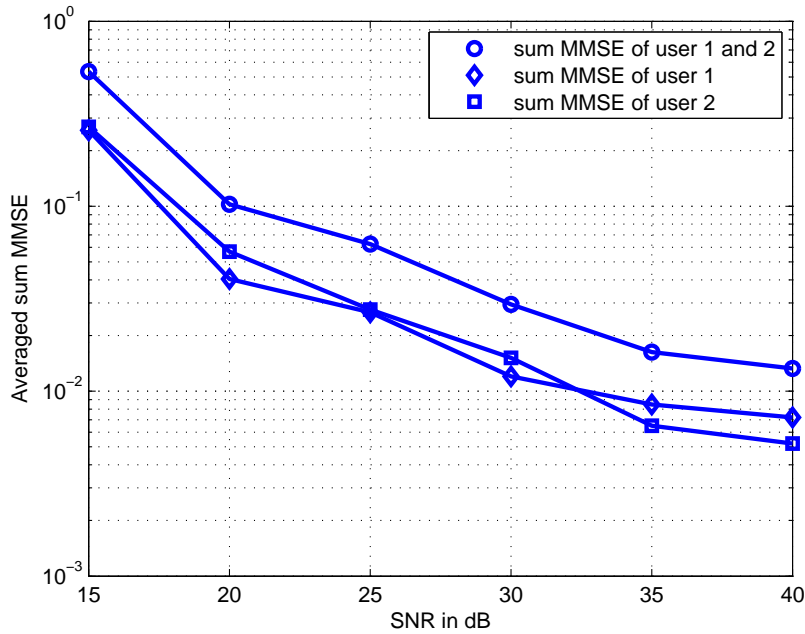
Figure 4.6. The sum MMSE against SNR of the relay-user terminal link, averaged over 10 set of random channels. The results are shown for sum MMSE of both users and sum MMSE of user 1 and user 2. SNR of transmitter-relay link was set to 30dB.

power and noise variance at the relay and SNR2 as the ratio between the relay power and the noise variance at the user terminals. To obtain 30dB for the first link (SNR1), the transmitter power was set to unity (W) and the noise variance at the relay to 0.001 W. In order to change SNR2, the relay power has been varied but the noise variance has been set at both user terminals to 0.001W. Ten monte-carlo simulations have been performed for various random channels and averaged the sum MMSE values. The average sum MMSE of all users as well as averaged sum MMSE of each users (i.e. summed over multiple data streams for each user) is depicted in Figure 4.6. As seen in Figure 4.6, the sum MMSE decreases as the SNR increases. More-over, even though sum MMSE has been considered as the design criterion, both users attain more or less the same averaged sum MMSE values over a set of random channels ensuring a degree of fairness.

4.8 Conclusion

Based on the derivation of the mean square error uplink-downlink duality for a MIMO peer-to-peer relay network, a transmitter, receiver and relay design technique has been proposed. The algorithm was based on an iterative method and second order cone programming. The simulation results demonstrate satisfactory performance in term of achievable mean square error and susceptibility to local minimum.

Chapter 5

A COORDINATED MULTIUSER RELAYING TECHNIQUE THROUGH INTERFERENCE PRECODING AT THE BASE STATION

A new precoder design technique for a cooperative BS and relay station network is proposed. The system model considered is a BS serving a number of users directly and another set of users through a MIMO wireless relay. A set of beamformers are designed at the relay and the BS to achieve a set of target rates for users. Most of the works on relay design have not focused on the backhaul overhead required in terms of transmission of the data from the BS to the relay. The work in [150], however, proposed a clever optimization approach to include this overhead in the design of BS and relay transceivers. A set of beamformers for BS and relay was designed while ensuring the users served by the BS and the relay achieve a set of target data rates and the data throughput from BS to relay is no less than the sum data rate of users served by the relay. Inspired by this work in [150], a technique is proposed

121
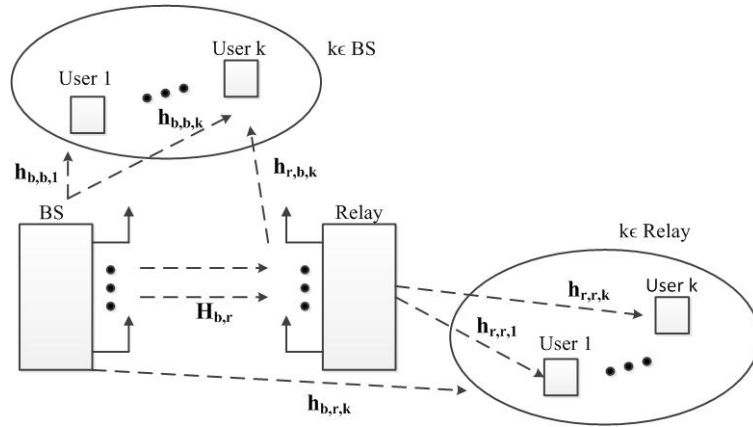
Figure 5.1. Block diagram of a cellular relay system consisting of one BS, one MIMO Relay and a number of users served by either BS or relay.

in this chapter to improve the network power usage of this coordinated BS and relay scheme by exploiting the knowledge available at the BS in terms of the interference caused by the relay transmission. The benefit of the proposed scheme is demonstrated in terms of transmission power usage by the simulation results.

5.1   System Model

A downlink multiuser network in a cellular relay as depicted in Figure 5.1 is considered with a single BS, a DF based relay in half duplex mode and users attached to BS and relay denoted by the set B and R respectively.The transmitters are assumed to have complete CSI, and there are $|B \cup R| = K$ number of users with single-antenna, where $|.|$ is the cardinality of a set. The BS and relay are equipped with N_B and N_R antennas, respectively. The communications between the BS, the relay and the users are established in two time-slots. In the first time slot (TS1), the BS is concurrently sending independent information to its $|B|$ single-antenna users and to the relay. The channel $\mathbf{H}_{b,r}$ between the BS and the relay is assumed to be of rank

$N_L = \min\{N_B, N_R\}$, thus allowing up to N_L data streams to be transmitted to the relay. The relay decodes the received signal relevant to its assigned users and during the second time slot, it transmits the decoded messages to $|R|$ users. At the same time, in the second time slot, the BS also transmits new independent messages to its $|B|$ users. Consequently, both the users belonging to BS and relay experience interference in the second time slot.

The work in [150] presented joint beamformer design and power allocation techniques for BS and relay so that users served by both the BS and the relay achieve a set of target data rates. To avoid an iterative design and to maintain low complexity, the work in [150] suggested performing singular value decomposition of $\mathbf{H}_{b,r}$ and set the receiver filter of the relay at the first time slot as the left singular vectors of $\mathbf{H}_{b,r}$. In this case, the output of the relay can be considered as N_L single antenna users. So, in first time slot, the problem is reduced to designing $N_L + |B|$ number of beamformers to serve $|B|$ users and the relay. In the second time slot, the BS transmits data to its $|B|$ users and the relay transmits data to its $|R|$ users. Hence, beamformers are required at the BS and the relay to ensure the users belong to the BS and relay attain certain data rate targets. However, the design of the beamformers should incorporate a constraint that the data rate achieved between the BS and the relay in the first time slot should be equal to the sum rate of users served by the relay in the second time slot.

5.2   Beamformer design with interference precoding

When designing the beamformers for BS in the second time slot, the work in [150] considered the signal transmitted by the relay as interference. However, the information transmitted by the relay in the second time slot is the very same information that was conveyed to the relay by the BS in the first time slot. Hence, the interference induced at the receiver of users served

by the BS is already known at the transmitter. Therefore, this knowledge of the interference structure is exploited at the transmitter (BS) to design beamformers which is the main focus of the work presented in this chapter. The optimization required in the first time slot will remain the same as in [150], however, optimization of the proposed algorithm will be different in the second time slot due to interference precoding. For completeness, the optimization for the first time slot is summarized before proceeding to the second time slot. In the first time slot, the BS transmits signal to both the users attached to it and to the relay by processing the signal through a set of beamformers, $\mathbf{w}_{b,k}^{(1)} \in \mathbb{C}^{N_B \times 1}$, $k = 1, \cdots, |B|$ and $\mathbf{w}_{b,r,l}^{(1)} \in \mathbb{C}^{N_B \times 1}$, $l = 1, \cdots, N_L$, designed through the following optimization:

$$
\min_{\mathbf{w}_{b,k}^{(1)}, \mathbf{w}_{b,r,l}^{(1)}, \bar{R}_l} \quad \sum_{k=1}^{|B|} ||\mathbf{w}_{b,k}^{(1)}||^2 + \sum_{l=1}^{N_L} ||\mathbf{w}_{b,r,l}^{(1)}||^2
$$
$$
\text{s.t.} \quad \Gamma_k^{(1)} \geq 2^{R_k} - 1, \forall\, k \in B
$$
$$
\bar{\Gamma}_l^{(1)} \geq 2^{2\bar{R}_l} - 1, \ l = 1, \cdots, N_l
$$
$$
\sum_{l=1}^{N_l} \bar{R}_l = \sum_{j=1}^{|R|} R_j \qquad\qquad (5.2.1)
$$

where the parenthesis (t) denotes time slot t, $\Gamma_k^{(1)}$ is the signal-to-interference plus noise ratio (SINR) achieved by user $k \in B$ and $\bar{\Gamma}_l^{(1)}$ is the SINR achieved through the lth spatial sub channel between the BS and relay. Both the $\Gamma_k^{(1)}$ and $\bar{\Gamma}_l^{(1)}$ are functions of $\mathbf{w}_{b,k}^{(1)}$ and $\mathbf{w}_{b,r,l}^{(1)}$ [150]. The last constraint ensures that the data rate of the BS-relay link is equal to the sum-rate of the users attached to relay. For a given set of $\bar{R}_l$, the overall optimization problem is SOCP, however, it is necessary to determine the optimum split of data throughput (i.e. $\bar{R}_l$) through the BS-relay channel, which can be achieved iteratively through GP together with signomial approximation [150]. In the second time slot, the BS transmits signal to only its $|B|$ users while the relay transmits signal to its $|R|$ users. The users attached to BS will receive signals

transmitted by both the BS and the relay. The work in [150] considered the relay signal at the BS users' terminals as interference in the optimization. However, as the BS knows the messages transmitted by the relay to its users (which were transmitted by the BS to the relay in the first time slot), the BS could aim to cancel the interference seen by its own users in the second time slot by performing the following linear preprocessing at the transmitter:

$$\mathbf{x}_B^{(2)} = \sum_{k=1}^{|B|} \mathbf{w}_{b,k}^{(2)} s_{b,k}^{(2)} + \sum_{k=1}^{|R|} \boldsymbol{\theta}_{b,k}^{(2)} s_{r,k}^{(2)} \tag{5.2.2}$$

where $\boldsymbol{\theta}_{b,k}^{(2)} \in \mathbb{C}^{N_B \times 1}$ is an interference precoding vector aimed at mitigating interference caused by the relay to the users $k \in B$ served by the BS and $\mathbf{w}_{b,k}^{(2)} \in \mathbb{C}^{N_B \times 1}$ are the beamformers for the users attached to BS. $\mathbf{x}_B$ is the signal transmitted by BS in the second time slot. In this case, the received signal at the kth BS user can be written as

$$\mathbf{y}_k^{(2)} = \sum_{i=1}^{|B|} \mathbf{h}_{b,b,k} \mathbf{w}_{b,i}^{(2)} s_{b,i}^{(2)} + \sum_{i=1}^{|R|} (\mathbf{h}_{b,b,k} \boldsymbol{\theta}_{b,i}^{(2)} s_{r,i}^{(2)} + \mathbf{h}_{r,b,k} \mathbf{w}_{r,i}^{(2)} s_{r,i}^{(2)}) + \mathbf{n}_k^{(2)}$$
$$\tag{5.2.3}$$

where $\mathbf{w}_{r,i}^{(2)} \in \mathbb{C}^{N_R \times 1}$ are the beamformers at the relay for its users. The subscripts b and r indicate the channel considered corresponds to BS and relay respectively. For example, $\mathbf{h}_{r,b,k}$ denotes the channel vector of the kth stream (considered as the kth user) from the BS to the relay. $s_{b,k}^{(2)}$ and $s_{r,k}^{(2)}$ are the normalized complex data symbols transmitted from the BS and relay, while $\mathbf{n}_k^{(2)}$ is circularly symmetric additive white Gaussian noise with zero mean and variance σ_k^2. The optimization is expected to determine an appropriate set of interference precoders $\boldsymbol{\theta}_{b,k}^{(2)}$, in addition to the beamformers $\mathbf{w}_{b,k}^{(s)}$ and $\mathbf{w}_{r,k}^{(2)}$ to ensure that the users achieve the target data rates with minimum possible transmission power. Since the same signal $s_{r,k}^{(2)}$ is transmitted from both the BS and the relay, we define augmented vectors

$\tilde{\mathbf{w}}_k^{(2)} = [\boldsymbol{\theta}_{b,k}^{(2)T}; \mathbf{w}_{r,k}^{(2)T}]^T$ and $\tilde{\mathbf{h}}_k = [\mathbf{h}_{b,b,k}\ \mathbf{h}_{r,b,k}]$ in order to write (5.2.3) as:

$$\mathbf{y}_k^{(2)} = \sum_{i=1}^{|B|} \mathbf{h}_{b,b,k} \mathbf{w}_{b,i}^{(2)} s_{b,i}^{(2)} + \sum_{i=1}^{|R|} \tilde{\mathbf{h}}_k \tilde{\mathbf{w}}_i^{(2)} s_{r,i}^{(2)} + \mathbf{n}_k^{(2)} \qquad (5.2.4)$$

Hence, the SINR of the above kth user can be written as

$$\Gamma_k^{(2)} = \frac{|\mathbf{h}_{b,b,k}\mathbf{w}_{b,k}^{(2)}|^2}{\sum\limits_{\substack{i=1 \\ i \neq k}}^{|B|} |\mathbf{h}_{b,b,k}\mathbf{w}_{b,i}^{(2)}|^2 + \sum\limits_{i=1}^{|R|} |\tilde{\mathbf{h}}_k\tilde{\mathbf{w}}_i^{(2)}|^2 + \sigma_k^2} \qquad (5.2.5)$$

The signal received by users served by relay can be written as

$$\mathbf{y}_k^{(2)} = \sum_{i=1}^{|R|} (\mathbf{h}_{r,r,k}\mathbf{w}_{r,i}^{(2)} s_{r,i}^{(2)} + \mathbf{h}_{b,r,k}\boldsymbol{\theta}_{b,i}^{(2)} s_{r,i}^{(2)}) + \sum_{i=1}^{|B|} \mathbf{h}_{b,r,k}\mathbf{w}_{b,i}^{(2)} s_{b,i}^{(2)} + \mathbf{n}_k^{(2)}$$

$$(5.2.6)$$

In the second time slot, the BS could also help the users served by the relay by coherent transmission of messages $s_{r,i}^{(2)}$. However, since the channel gains from BS to relay users are likely to be very small (otherwise these users are not needed to be served by the relay) and since the relay users will experience considerable delay on signals from BS as compared to signals coming from relay, it is not viable to coherently add the signals transmitted by BS and relay at the receiver of the relay users. Hence, the signals transmitted by the BS, including $\boldsymbol{\theta}_{b,i}^{(2)} s_{r,i}^{(2)}$ that was intended to cancel BS user interference is considered as interference to relay users. Therefore, the SINR of the relay users $k \in R$ is written as:

$$\Gamma_k^{(2)} = \frac{|\mathbf{h}_{r,r,k}\mathbf{w}_{r,k}^{(2)}|^2}{\sum\limits_{\substack{i=1 \\ i \neq k}}^{|R|} |\mathbf{h}_{r,r,k}\mathbf{w}_{r,i}^{(2)}|^2 + \sum\limits_{i=1}^{|R|} |\mathbf{h}_{b,r,k}\boldsymbol{\theta}_{b,i}^{(2)}|^2 + \sum\limits_{i=1}^{|B|} |\mathbf{h}_{b,r,k}\mathbf{w}_{b,i}^{(2)}|^2 + \sigma_k^2} \qquad (5.2.7)$$

Since (5.2.7) is written in terms of $\mathbf{w}_{b,k}^{(2)}$, $\mathbf{w}_{r,k}^{(2)}$ and $\boldsymbol{\theta}_{b,k}^{(2)}$, we write $\boldsymbol{\theta}_{b,k}^{(2)}$ and $\mathbf{w}_{r,k}^{(2)}$ in terms of the optimization variable $\tilde{\mathbf{w}}$ as:

$$\boldsymbol{\theta}_{b,k}^{(2)} = [\mathbf{I}_{N_B \times N_B} \ \mathbf{0}_{N_B \times N_R}]\tilde{\mathbf{w}}_k^{(2)} = \mathbf{M}_1 \tilde{\mathbf{w}}_k^{(2)} \text{ and}$$

$$\mathbf{w}_{r,k}^{(2)} = [\mathbf{0}_{N_R \times N_B} \ \mathbf{I}_{N_R \times N_R}]\tilde{\mathbf{w}}_k^{(2)} = \mathbf{M}_2 \tilde{\mathbf{w}}_k^{(2)}.$$

Hence, the user-rate constrained power minimization problem for the second time slot is written as:

$$\min_{\mathbf{w}_{b,k}^{(2)}, \tilde{\mathbf{w}}_k^{(2)}} \sum_{k=1}^{|B|} ||\mathbf{w}_{b,k}^{(2)}||^2 + \sum_{k=1}^{|R|} ||\tilde{\mathbf{w}}_k^{(2)}||^2 \text{ , subject to}$$

$$\frac{|\mathbf{h}_{b,b,k}\mathbf{w}_{b,k}^{(2)}|^2}{\sum\limits_{\substack{i=1 \\ i \neq k}}^{|B|} |\mathbf{h}_{b,b,k}\mathbf{w}_{b,i}^{(2)}|^2 + \sum\limits_{i=1}^{|R|} |\tilde{\mathbf{h}}_k\tilde{\mathbf{w}}_i^{(2)}|^2 + \sigma_k^2} \geq 2^{R_k} - 1, \forall \ k \in B$$

$$\frac{|\mathbf{h}_{r,r,k}\mathbf{m}_2\tilde{\mathbf{w}}_k^{(2)}|^2}{\sum\limits_{\substack{i=1 \\ i \neq k}}^{|R|} |\mathbf{h}_{r,r,k}\mathbf{M}_2\tilde{\mathbf{w}}_i^{(2)}|^2 + \sum\limits_{i=1}^{|R|} |\mathbf{h}_{b,r,k}\mathbf{M}_1\tilde{\mathbf{w}}_i^{(2)}|^2 + \sum\limits_{i=1}^{|B|} |\mathbf{h}_{b,r,k}\mathbf{w}_{b,i}^{(2)}|^2 + \sigma_k^2}$$

$$\geq 2^{2R_k} - 1, \ \forall \ k \in R \tag{5.2.8}$$

The above optimization can be solved using SOCP and $\boldsymbol{\theta}_{b,k}$, $\mathbf{w}_{r,k}$ can be obtained from $\tilde{\mathbf{w}}_{\mathbf{k}}$. This proposed scheme requires symbol level synchronization between BS and relay transmission, however this is not impossible as both the BS and relay as well as all users belong to the same network, and the BS will have direct control channels to relay and the users attached to it. Technical requirement and benefits of coherent multi-point transmission have been investigated in [151] and references therein.

As the overall problem has been approximated to convex, the complexity is polynomial time. The overall complexity is limited by the complexity of SOCP. Compared to the work in [150], the extra complexity of the proposed method is due to the increased dimension of the optimization variable (i.e. $\tilde{\mathbf{w}}$ instead of $\mathbf{w}_r$). Assuming $N_B \geq N_R$, the worst case complexity of the algorithm in [150] is $O(K^{1.5}(2N_B)^3)$ whereas it is $O(K^{1.5}(2(N_B + N_R))^3)$ for the proposed algorithm, [152].

For a given set of users, optimum allocation of users to BS and relay is a combinatorial optimization problem which is NP hard. However, a suboptimum solution was proposed in [150] which assumed zero interference between users, however, for both methods proposed here and in [150] that were based on SOCP allow interference between users in the SINR formulation. For both these cases, the user allocation algorithm in [150] can still be used, but they will be suboptimum. If the complexity is tolerable, it is possible to develop user allocation based on mixed integer programming and SOCP based on branch and bound method as similar to the work in [153]. The complexity can be reduced further by allocating users that have strong channel gains to either BS or relay to the corresponding transmitters and performing optimization only for those users whose channel gains are approximately the same to either BS or relay.

5.3   Simulation Results

In order to assess the performance of the proposed design using the interference precoding and compare this with the design proposed in [150] a scenario where BS and relay are equipped with 4 and 2 antennas respectively and both BS and relay serve two single-antenna users each is considered. A set of random channels using zero mean complex Gaussian random variables has been generated. The average channel power gain from BS to relay, the BS to its users and the relay to its users was set to one. The channel gain from BS to users assigned to relay was set to 0.01, however for the first simulation an additional value of 0.1 was also considered for this channel as well as the channel gain from relay to users served by the BS was set to 0.1 . In Figure 5.2, the required total transmission power (in dB) against the users' target data rates (in bits/sec/Hz) is depicted with and without interference precoding. As target rate increases, a substantial reduction in
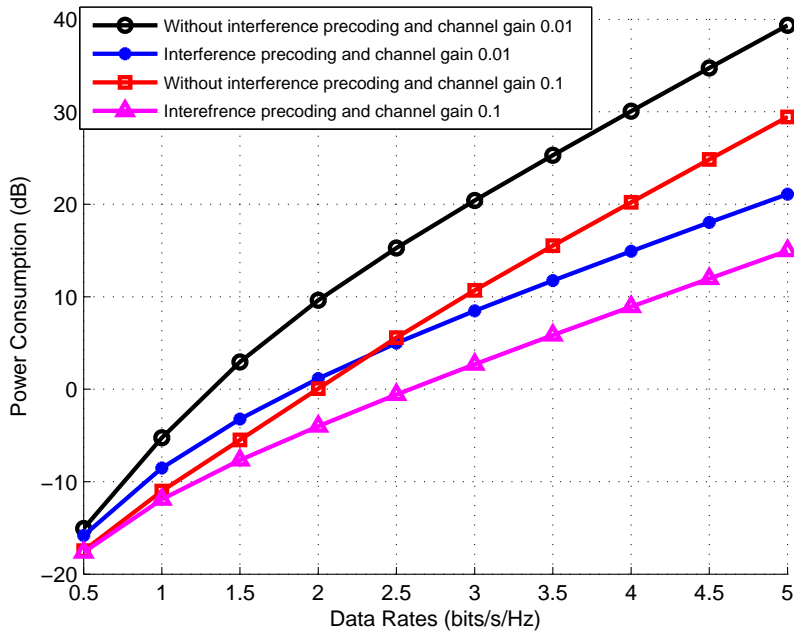
Figure 5.2. The total transmission power against various target data rates.

transmission power using our proposed scheme is observed. For example, at 4 bits/sec/Hz, a power saving of up to factor 10 is achieved.

For the second simulation, the channels have been set the same as before except that the average channel gain between the relay and the users served by the BS has been varied from -20dB to 10dB. The target data rates for all 4 users were set to 9 and 5. The required transmission power for the proposed scheme and the scheme that does not consider interference precoding is depicted in Figure 5.3. As seen, the proposed design provides substantial reduction in transmission power for a range of channel gains between the relay and BS users. For example, at 0dB channel gain, for target rate of 5 bits/s/Hz, we observe 15dB reduction in transmission power using our scheme.

The BS to relay channel may introduce bit errors that may affect the SINR of BS users in the second time slot due to mismatch between the precoded interference and the true interference from relay. However, for
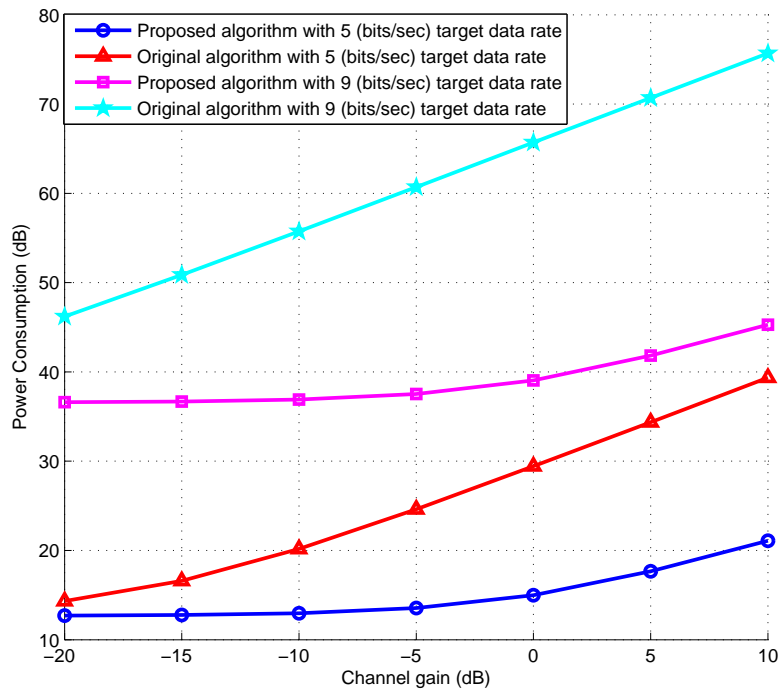
Figure 5.3. The total transmission power against various interference channel gains. The unit of target data rates is bits/sec/Hz

the DF relay, the relay should normally perform code redundancy check on each packet and forwards the signals to the users only when packets are received correctly, otherwise it should seek retransmission of packets from BS. Therefore, by setting appropriate delay between BS transmission and relay transmission of signals, the BS and relay could coordinate to avoid transmission of precoded interference whenever relay does not transmit signals. If for any reason, the relay forwards the signal without performing code redundancy check check (e.g. for the case of voice packets or real time services where delay is unacceptable), certain bits transmitted by relay may be in error. Consequently, the interference precoding performed at the BS may not match exactly the interference seen from relay at the BS user terminals. In this case, the SINR achieved at the BS users will be lower than the target SINR. Normally, the BER will be very small due to forward error control

coding. However, in order to consider a worst case scenario, a simulation to include possible errors caused at relay is performed when the signal from the BS to relay is transmitted without forward error control coding. The target rate for the 2 users attached to relay is set as 1bit/s/Hz and 0.5 bit/s/Hz respectively. Since, only one out of two time slots is used for relay users, the capacity of BS to relay channel should be above 3bits/s/Hz. The SINR required to achieve this is 7 (i.e. 8.4510 dB). Hence Eb/No is 3.6797 dB. Considering 8 phase-shift-keying modulation for the BS-RS transmission and using the BER formula ($BER = \frac{2}{log_2 M} Q(\sqrt{2 \frac{E_b}{N_0} log_2 M} sin(\frac{\pi}{M})))$, $M = 8$, the BER is computed as 0.15. A packet of 300 bits has been considered; 200 bits for the relay user 1 and 100 bits for the relay user 2. From relay to the first user, a quadrature phase shift keying modulation has been considered and for the other user a binary phase shift keying modulation. With 0.15 BER and assuming grey coding, 15 quadrature phase shift keying and 15 binary phase shift keying symbols are in error. These errors have been randomly generated and the true SINR achieved at the BS user terminals using the beamformer designed has been computed. As the achieved SINR is lower than the target SINR, the target SINR of the users attached to the BS has been increased in the optimization until the data rate achieved meets the target rate in the presence of bit errors. The average transmission power required for 300 different random realizations of channels has been computed and the result has been compared with our scheme (assuming no error propagation) and with the scheme of [150]. As seen in Figure 5.4, even in the presence of errors, the proposed scheme achieves the target rates with substantially lower transmission power. Having presented the worst case scenario, it should be highlighted that the signals should be transmitted with appropriate forward error control coding, and the SINR operating points are normally set slightly above the theoretical SINR value obtained from the Shannon capacity. This together with advanced forward error control

coding such as Turbo codes will result into arbitrarily small BER, e.g., in the range of 10^{-3} to 10^{-6} [154]. To account for this possible scenario, our simulation has been repeated with a lower BER than the uncoded BER of 0.15. Even though, it is possible to get very low BER with forward error control coding such as Turbo codes, once again a worst case scenario has been considered by setting the BER as 0.01, and obtained the result as in Figure 5.4. As seen, the required transmission power is very close to as if there is no error in relay transmission.
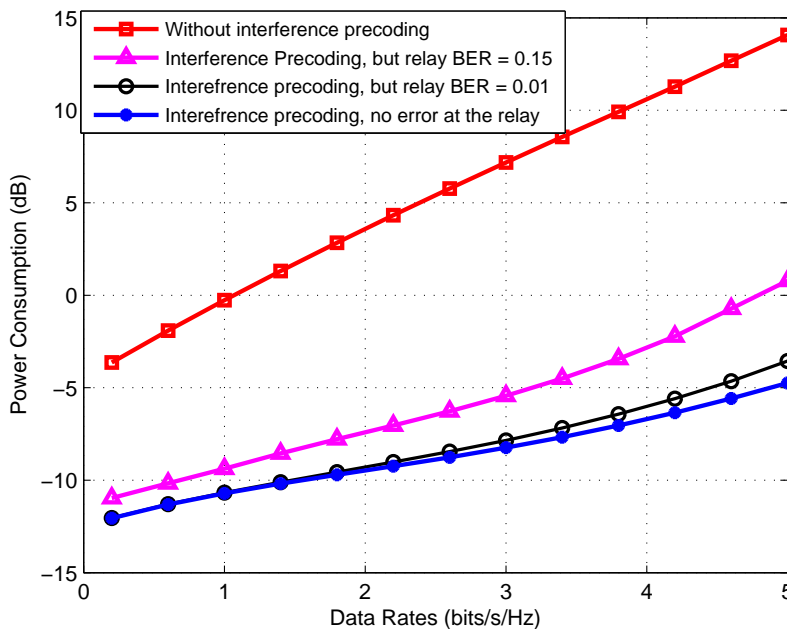
Figure 5.4. The total transmission power against various target data rates for BS users, averaged over 300 random channel realizations.

5.4 Conclusion

An interference pre-subtraction based design technique for coordinated BS and relay beamforming has been proposed. Since the interference introduced by the relay to the users served by the BS is known to the BS, the BS is able to mitigate the interference seen by its users through interference pre-

cancelation.   The proposed method outperforms the coordinated BS and relay beamforming technique that does not consider interference precoding at the transmitter, in terms of transmission power for a wide range of target data rates.

Chapter 6

A BASESTATION BEAMFORMING TECHNIQUE USING MULTIPLE SINR BALANCING CRITERIA

A coordinated multi-cell beamforming technique for SINR balancing under multiple BS power constraints is proposed in this chapter. Instead of balancing SINR of all users in all cells to the same level, a new approach to balance SINR of users in various cells to different maximum possible values is proposed. This has the ability to allow users in cells with relatively more transmit power or better channel condition to achieve a higher balanced SINR than that achieved by users in the worst case cells. According to the proposed approach, first, minimum SINR of all users in all cells is maximized. After this optimization, certain BSs may not use all of their transmission power. Hence, a subsequent optimization to design beamformers is proposed to enable users in the BSs that have excess power to maximize their balanced SINR values while keeping the SINRs of users in other BSs unchanged. This problem is solved using SINR constraint based SINR balancing techniques with multiple linear power constraints.

6.1   Signal Model and Problem Formulation

A multi-cell multiuser downlink beamforming design problem with N cells that operate in the same frequency band is considered. At each cell, a BS consisting of M antennas performs spatial multiplexing through beamforming and transmits signals simultaneously to K single-antenna terminals. Let $s_{i,j}$ denote the information symbol for the jth user in the ith cell, with $E\{|s_{i,j}|^2\} = 1$, $\mathbf{u}_{i,j} \in \mathbb{C}^{M \times 1}$ is the associated unity-norm beamformer vector (i.e., $\|\mathbf{u}_{i,j}\|_2 = 1$, $\forall\ i,\ j$) and $p_{i,j}$ is the power allocated to the jth user in the ith cell. Moreover, $\mathbf{p}_i = [p_{i1} \cdots p_{iK}]^T$ and $\mathbf{p} = [\mathbf{p}_1^T \cdots \mathbf{p}_N^T]^T$ is introduced. Let $\mathbf{U}_i = [\mathbf{u}_{i,1} \cdots \mathbf{u}_{i,K}] \in \mathbb{C}^{M \times K}$ denote the beamforming matrix for the users in the ith cell and $\mathbf{U} \in \{\mathbf{U}_1, \mathbf{U}_2, \cdots, \mathbf{U}_N\}$ denotes the set of all beamforming matrices. The received signal at the jth user in the ith cell consists of the desired signal, intra-cell and inter-cell interference, as expressed by

$$\mathbf{y}_{i,j} = \sum_{k=1}^{K} \mathbf{h}_{i,i,j}^H \sqrt{p_{i,k}} \mathbf{u}_{i,k} s_{i,k} + \sum_{\substack{m=1 \\ m \neq i}}^{N} \sum_{l=1}^{K} \mathbf{h}_{m,i,j}^H \sqrt{p_{m,l}} \mathbf{u}_{m,l} s_{m,l} + n_{i,j},$$

$$(6.1.1)$$

where $\mathbf{h}_{m,i,j}$ denotes the channel vector from the BS of the mth cell to the jth user in the ith cell and $n_{i,j}$ is the complex additive white Gaussian noise (AWGN) with zero mean and variance $\sigma_{i,j}^2$, i.e., $n_{i,j} \sim \mathcal{CN}(0, \sigma_{i,j}^2)$. Fig. 6.1 illustrates the system model.

Let P_i^{max} be the maximum possible transmit power of the ith BS. Then it should hold that $\mathbf{1}_K^T \mathbf{p}_i \leq P_i^{max}$. By defining $\mathbf{R}_{i,i,j} \triangleq \mathbf{h}_{i,i,j} \mathbf{h}_{i,i,j}^H$, the downlink SINR for the jth user in the ith cell is expressed as

$$\Gamma_{i,j}^{DL} = \frac{p_{i,j} \mathbf{u}_{i,j}^H \mathbf{R}_{i,i,j} \mathbf{u}_{i,j}}{\sum_{\substack{l=1 \\ l \neq j}}^{K} p_{i,l} \mathbf{u}_{i,l}^H \mathbf{R}_{i,i,j} \mathbf{u}_{i,l} + \sum_{\substack{m=1 \\ m \neq i}}^{N} \sum_{l=1}^{K} p_{m,l} \mathbf{u}_{m,l}^H \mathbf{R}_{m,i,j} \mathbf{u}_{m,l} + \sigma_{i,j}^2}.$$

$$(6.1.2)$$
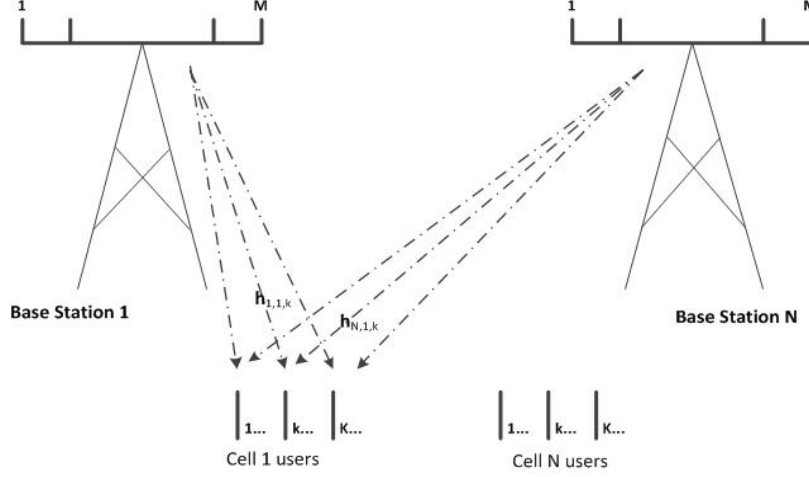
Figure 6.1. Multicell wireless network

The goal is to jointly optimize the beamformers $\mathbf{U}$ and the power allocation vector $\mathbf{p}$ of all cells, so that the SINRs of users in each cell are balanced and maximized. Let t_i be the balanced SINR of the users in the ith cell. The problem of interest can be stated as

$$\max_{\mathbf{U},\mathbf{p}\succeq 0} \quad t_i \,, \forall i \in \{1, 2, \cdots, N\} \tag{6.1.3a}$$

$$\text{s.t.} \quad \min_{1 \leq j \leq K} \frac{\Gamma_{i,j}^{DL}(\mathbf{U}, \mathbf{p})}{\rho_{i,j}} \geq t_i, \forall i \tag{6.1.3b}$$

$$\mathbf{1}_K^T \mathbf{p}_i \leq P_i^{max}, \ \forall i, \tag{6.1.3c}$$

where $\rho_{i,j}$ is a priority factor for user j in the ith cell. It should be highlighted that all available works in the literature so far considered $t_1 = t_2 = \cdots = t_N = t$, i.e., SINRs of all users in all cells are balanced to an identical t value, however by considering different $t_i = 1, 2, \cdots, N$, we aim to balance SINRs of all users in various cells to different possible maximum values.

6.2   SINR-Balancing in the Coordinated Multi-cell

The first stage of the proposed algorithm aims to balance the SINR of all users in all cells. This is achieved using the SINR balancing techniques proposed, for example, [17], [155] and [156]. For the purpose of completeness and to facilitate development of the proposed algorithm in Section IV, the SINR balancing technique for multi-cell beamforming is summarized in this section. The minimum SINR of users in any cell is maximized, subject to an upper bound on the transmit power as follows:

$$\max_{\mathbf{U},\mathbf{p}\succeq 0} \min_{i,j} \quad \frac{\Gamma_{i,j}^{DL}(\mathbf{U},\mathbf{p})}{\rho_{i,j}} \tag{6.2.1a}$$

$$\text{s.t.} \quad \mathbf{1}_K^T \mathbf{p}_i \leq P_i^{max}, \forall i. \tag{6.2.1b}$$

The multiple per-BS power constraints in (6.2.1b) can be put into a single linear constraint by introducing auxiliary variables $a_i \in \mathbb{R}^+$ as $\sum_{i=1}^{N} a_i(\mathbf{1}_K^T \mathbf{p}_i - P_i^{max}) \leq 0$. The auxiliary variables need to be adapted using a subgradient method as described in [156]. Defining a vector $\mathbf{a} = [a_1 \mathbf{1}_K^T \cdots a_N \mathbf{1}_K^T]^T$ and a scalar $P^{max} \triangleq \sum_{i=1}^{K} a_i P_i^{max}$, the problem in (6.2.1) can be restated as

$$\max_{\mathbf{U},\mathbf{p}\succeq 0} \min_{i,j} \quad \frac{\Gamma_{i,j}^{DL}(\mathbf{U},\mathbf{p})}{\rho_{i,j}} \tag{6.2.2a}$$

$$\text{s.t.} \quad \mathbf{a}^T \mathbf{p} \leq P^{max}. \tag{6.2.2b}$$

This problem can be solved iteratively using the following three sub-problems: downlink power allocation, uplink power allocation and beamforming design.

6.2.1   Downlink Power Assignment for a given set of Beamformers

Let $\widetilde{\mathbf{p}}$ be the optimal power allocation for the problem in (6.2.2). At the optimum, the inequality constraint should be satisfied with equality [17]. Hence for a given set of beamformers $\widetilde{\mathbf{U}}$, the optimization in (6.2.2) should

satisfy

$$\frac{\Gamma_{i,j}^{DL}(\widetilde{\mathbf{U}}, \widetilde{\mathbf{p}})}{\rho_{i,j}} = \gamma^{DL}(\widetilde{\mathbf{U}}, P^{max}), \quad (6.2.3a)$$

$$\mathbf{a}^T \widetilde{\mathbf{p}} = P^{max}, \quad (6.2.3b)$$

where $\gamma^{DL}(\widetilde{\mathbf{U}}, P^{max})$ is the balanced SINR. Substituting the definition of SINR given in (6.1.2) and using the method proposed in [17], the following eigensystem can be constructed to determine the optimum downlink power allocation $\widetilde{\mathbf{p}}$:

$$\begin{bmatrix} \mathbf{D}\boldsymbol{\Psi} & \mathbf{D}\boldsymbol{\sigma} \\ \frac{1}{P^{max}}\mathbf{a}^T\mathbf{D}\boldsymbol{\Psi} & \frac{1}{P^{max}}\mathbf{a}^T\mathbf{D}\boldsymbol{\sigma} \end{bmatrix} \begin{pmatrix} \widetilde{\mathbf{p}} \\ 1 \end{pmatrix} = \frac{1}{\gamma^{DL}(\widetilde{\mathbf{U}}, P^{max})} \begin{pmatrix} \widetilde{\mathbf{p}} \\ 1 \end{pmatrix}, \quad (6.2.4)$$

where $\mathbf{D} = \text{blkdiag}\{\mathbf{D}_1, \cdots, \mathbf{D}_N\}$,

$\mathbf{D}_n = \text{diag}[(\frac{\rho_{n,1}}{\widetilde{\mathbf{u}}_{n1}^H \mathbf{R}_{n,n,1}\widetilde{\mathbf{u}}_{n1}}) \cdots (\frac{\rho_{n,K}}{\widetilde{\mathbf{u}}_{nK}^H \mathbf{R}_{n,n,K}\widetilde{\mathbf{u}}_{nK}})]$,

$\boldsymbol{\sigma} = [\boldsymbol{\sigma}_1 \cdots \boldsymbol{\sigma}_N]^T$, $\boldsymbol{\sigma}_n = [\sigma_{n,1}^2 \cdots \sigma_{n,K}^2]^T$,

$$\boldsymbol{\Psi} = \begin{bmatrix} \boldsymbol{\Psi}_1 & \boldsymbol{\Psi}_{2,1} & \dots & \boldsymbol{\Psi}_{N,1} \\ \boldsymbol{\Psi}_{1,2} & \boldsymbol{\Psi}_2 & \dots & \boldsymbol{\Psi}_{N,2} \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{\Psi}_{1,N} & \boldsymbol{\Psi}_{2,N} & \dots & \boldsymbol{\Psi}_N \end{bmatrix}, \quad (6.2.5)$$

$$[\boldsymbol{\Psi}_n]_{ki} = \begin{cases} \widetilde{\mathbf{u}}_{ni}^H \mathbf{R}_{nnk}\widetilde{\mathbf{u}}_{ni} & \text{if } i \neq k, i = 1, \cdots, K, \ j = 1, \cdots, K \\ 0 & \text{if } i = k \end{cases} \quad (6.2.6)$$

and

$$[\boldsymbol{\Psi}_{l,n}]_{ki} = \begin{cases} \widetilde{\mathbf{u}}_{li}^H \mathbf{R}_{lnk}\widetilde{\mathbf{u}}_{li} & k = 1, \cdots, K, \ i = 1, \cdots, K, \ l \neq n \cdot \end{cases} \quad (6.2.7)$$

where $\text{blkdiag}\{\mathbf{S}_1, \cdots, \mathbf{S}_M\}$ denotes a block-diagonal square matrix with $\mathbf{S}_1, \cdots, \mathbf{S}_M$ as the diagonal square matrices and $\text{diag}\{a_1, \cdots, a_M\}$ is a diag-

onal matrix whose diagonal elements are $a_1, \cdots, a_M$.

6.2.2   Uplink Power Assignment for a given set of Beamformers

Designing beamformers directly in the downlink is difficult due to coupling of all beamformer vectors and power allocation in the SINR equation of every user. Hence, SINR uplink-downlink duality has been proposed in [17, 156] which states that the set of SINRs achievable in the downlink can also be achieved in a virtual uplink using the same set of beamformers under identical total power constraint. Using the results in [156], for the multiple linear constraints as well, it can be shown that for a given set of beamformers $\widetilde{\mathbf{u}}_{i,j}$, the downlink optimization problem in (6.2.1) is equivalent to the following virtual uplink optimization problem:

$$\max_{\widetilde{\mathbf{q}} \succeq 0} \min_{i,j} \quad \frac{\Gamma_{i,j}^{UL}(\widetilde{\mathbf{u}}_{i,j}, \widetilde{\mathbf{q}})}{\rho_{i,j}} \tag{6.2.8a}$$

$$\text{s.t.} \quad \boldsymbol{\sigma}^T \widetilde{\mathbf{q}} \leq P^{max}, \tag{6.2.8b}$$

where $\Gamma_{i,j}^{UL}$ is the uplink SINR of user j in the ith cell given by

$$\Gamma_{i,j}^{UL} = \frac{\widetilde{q}_{i,j}\widetilde{\mathbf{u}}_{i,j}^{H}\mathbf{R}_{i,i,j}\widetilde{\mathbf{u}}_{i,j}}{\widetilde{\mathbf{u}}_{i,j}^{H}(\sum\limits_{\substack{l=1 \\ l \neq j}}^{K} \widetilde{q}_{i,l}\mathbf{R}_{i,i,l} + \sum\limits_{\substack{m=1 \\ m \neq i}}^{N} \sum\limits_{l=1}^{K} \widetilde{q}_{m,l}\mathbf{R}_{i,m,l} + \boldsymbol{\Omega}_i)\widetilde{\mathbf{u}}_{i,j}}, \tag{6.2.9}$$

and $\boldsymbol{\Omega}_i = a_i\mathbf{I}$ is the equivalent noise covariance matrix in the uplink, $q_{i,j}$ is the virtual uplink transmit power allocated to the jth user in the ith cell, $\widetilde{\mathbf{q}}_i = [\widetilde{q}_{i1}, \cdots, \widetilde{q}_{iK}]^T$ and $\widetilde{\mathbf{q}} = [\widetilde{\mathbf{q}}_i^T, \cdots, \widetilde{\mathbf{q}}_N^T]^T$. Similar to the downlink solution in (6.2.4), the uplink power allocation can be obtained by solving the following eigensystem:

$$\begin{bmatrix} \mathbf{D}\boldsymbol{\Psi}^T & \mathbf{Db} \\ \frac{1}{P^{max}}\boldsymbol{\sigma}^T\mathbf{D}\boldsymbol{\Psi}^T & \frac{1}{P^{max}}\boldsymbol{\sigma}^T\mathbf{Db} \end{bmatrix} \begin{pmatrix} \widetilde{\mathbf{q}} \\ 1 \end{pmatrix} = \frac{1}{\gamma^{UL}(\widetilde{\mathbf{U}}, P^{max})} \begin{pmatrix} \widetilde{\mathbf{q}} \\ 1 \end{pmatrix}, \tag{6.2.10}$$

where $\mathbf{b} = [\mathbf{b}_1, \cdots, \mathbf{b}_N]^T$ and

$$[\mathbf{b}_n]_i = \begin{cases} \widetilde{\mathbf{u}}_{ni}^H \mathbf{\Omega}_i \widetilde{\mathbf{u}}_{ni} & n = 1, \cdots, N, \ i = 1, \cdots, K. \end{cases}$$

6.2.3   Beamformer Design for a Given Power Allocation

For a given uplink power allocation $\widetilde{\mathbf{q}}$, beamformers in the uplink are designed by solving the following generalized eigenvector problem

$$\widetilde{\mathbf{u}}_{i,j} = \arg \max_{\mathbf{u}_{i,j}} \frac{\mathbf{u}_{i,j}^H \mathbf{R}_{i,i,j} \mathbf{u}_{i,j}}{\mathbf{u}_{i,j}^H \mathbf{Q}_{ij} \mathbf{u}_{i,j}}, \ \text{s.t. } ||\mathbf{u}_{i,j}||_2 = 1.$$

$$\text{where } \mathbf{Q}_{ij} = \sum_{\substack{l=1 \\ l \neq j}}^{K} \widetilde{q}_{i,l} \mathbf{R}_{i,i,l} + \sum_{\substack{m=1 \\ m \neq i}}^{N} \sum_{l=1}^{K} \widetilde{q}_{m,l} \mathbf{R}_{i,m,l} + \mathbf{\Omega}_i, \ \ \forall i, \forall j$$

$$(6.2.11)$$

6.2.4   Iterative Solution

The optimization in (6.2.1) is solved using an iterative algorithm. For a given set of auxiliary variables $\mathbf{a}$, the beamformers from (6.2.11) and the power allocation from (6.2.10) are iteratively obtained until convergence. From the uplink-downlink duality, the same SINR values can be achieved in both the uplink and the downlink with the same set of beamformers and total transmit power but with different power allocation. Hence, the beamformers $\mathbf{U}$ that have been computed in the virtual uplink, are used in the downlink to obtain the downlink power allocation using (6.2.4). In order to update the auxiliary variables, a subgradient method is used as follows [156], [157]:

$$a_i^{(m+1)} = a_i^{(m)} + t(\mathbf{1}_K^T \mathbf{p}_i^{(m)} - P_i^{max}), \ i = 1, 2, \cdots N, \qquad (6.2.12)$$

Algorithm 1

---

1) **Initialize** $m \Leftarrow 0$, $n \Leftarrow 0$, $\mathbf{q}^{(0)}$, $\mathbf{a}^{(0)}$, t, ϵ
2) **Repeat**
3) $m \Leftarrow m + 1$
4)   **Repeat**
5) $n \Leftarrow n + 1$
6)   Solve (6.2.11) using $\mathbf{q}^{(n-1)}$ to obtain $\widetilde{\mathbf{U}}^{(n-1)}$
7)   Compute $\mathbf{D}^{(n-1)}$, $\mathbf{\Psi}^{(n-1)}$, $\mathbf{b}^{(n-1)}$, using $\widetilde{\mathbf{U}}^{(n-1)}$
8)   Solve (6.2.10) and obtain $\gamma^{(n)}$ and $\widetilde{\mathbf{q}}^{(n)}$
9)   **Until** $\gamma^{(n-1)} - \gamma^{(n)} \leq \epsilon$
10) $\gamma^* = \gamma^{(n)}$ and $\widetilde{\mathbf{U}}^* = \widetilde{\mathbf{U}}^{(n-1)}$
11)   Obtain $\widetilde{\mathbf{p}}^{(m)}$ using (6.2.4)
12)   Update auxiliary variables using (6.2.12)
13) **Until** (6.2.13) is satisfied.

Table 6.1.   Algorithmic solution of the SINR Balancing Problem (6.2.1)

where t denotes the step size of the subgradient method. The algorithm stops when the following criterion is satisfied

$$|a_i^{(m+1)}(\mathbf{1}_K^T \mathbf{p}_i^{(m)} - P_i^{max})| \leq \epsilon, \ i = 1, 2, \cdots N, \qquad (6.2.13)$$

where ϵ is a very small positive constant. The algorithm is summarized in Table I. The quantities that are associated with the nth iteration are denoted by the superscript (n). We also use the shorthand notation $\gamma^{(n)}$, which is the inverse of the largest eigenvalue obtained in (6.2.10).

6.3   SINR-Balancing with per Basestation SINR-Target-Constraint

Once the SINR balancing is achieved as described in section III, at least one of the BS power constraints must be active, i.e., at least one of the BSs should be drawing the maximum possible transmit power. The rationale behind this argument is that suppose at the optimum, if all BSs do not use the maximum transmit power, the transmit power of all BSs can be scaled up by a factor $\alpha > 1$ until one of the BSs attains its full transmit power. Since

SINR increases monotonically with increasing transmit power, the balanced SINR will be increased. As the aim was to maximize the balanced SINR, at least one of the BS power constraints must be active. Let us denote this BS as l. The corresponding balanced SINR γ_l obtained from Algorithm 1, is the maximum SINR that this BS can achieve. At this stage, other BSs may not have used all of their transmit power. Hence, a second stage optimization is performed to enable other BSs to increase their balanced SINRs using their remaining transmit powers while ensuring SINRs of all users' in the lth BS are kept at γ_l. Hence, the aim of the optimization problem is to maximize minimal SINR of all users in all BSs except the lth BS, while keeping all transmit power constraints and introducing an additional constraint that the users in the lth BS should achieve a target SINR γ_l as follows:

$$\max_{\mathbf{U},\mathbf{p}\succeq 0} \min_{i,j} \quad \frac{\Gamma_{i,j}^{DL}(\mathbf{U},\mathbf{p})}{\rho_{i,j}}, \ \forall i, i \neq l, \forall j \qquad (6.3.1a)$$

$$\text{s.t.} \quad \frac{\Gamma_{l,j}^{DL}(\mathbf{U},\mathbf{p})}{\rho_{i,j}} \geq \gamma_l, \forall j \qquad (6.3.1b)$$

$$\mathbf{1}_K^T \mathbf{p}_i \leq P_i^{max}, \forall i. \qquad (6.3.1c)$$

The above optimization aims to maximize SINRs of users in the remaining *(N-1)* BSs, while ensuring the balanced SINR of users in the lth BS is γ_l. At this stage, the BS l and possibly one of the remaining *(N-1)* BSs will use the available transmit power fully. Let us denote the BS that uses its full transmit power from the remaining *(N-1)* BS as l' and the corresponding balanced SINR as $\gamma_{l'}$. It should be noted that $\gamma_{l'} \geq \gamma_l$. This is because if $\gamma_{l'} < \gamma_l$, then BS l' should have used all of its transmit power at the first stage of the optimization. This target SINR based SINR balancing beamforming design is repeated until SINR of users in the last BS is maximized. At the end of this repeated optimization, users in various BSs should have achieved different levels of balanced SINR. Suppose we are at the N_1th stage of this repeated optimization. Also, without loss of generality, assume the first N_1

BSs have achieved full use of their transmit power. Hence, the optimization is formulated as maximizing the SINR of all users in all BSs except the first N_1 BSs subject to the maximum transmit power constraints and SINR target constraint for the users in the first N_1 BSs as follows:

$$\max_{\mathbf{U},\mathbf{p}\succeq 0} \min_{i,j} \frac{\Gamma_{i,j}^{DL}(\mathbf{U},\mathbf{p})}{\rho_{i,j}}, \ i = N_1 + 1, \cdots N, \forall j \qquad (6.3.2a)$$

$$\text{s.t.} \quad \frac{\Gamma_{i,j}^{DL}(\mathbf{U},\mathbf{p})}{\rho_{i,j}} \geq \gamma_i, \ i = 1, \cdots N_1, \forall j, \qquad (6.3.2b)$$

$$\mathbf{1}_K^T \mathbf{p}_i \leq P_i^{max}, \forall i, \qquad (6.3.2c)$$

where $\Gamma_{i,j}^{DL}(\mathbf{U},\mathbf{p})$ is the SINR of the jth user in the ith cell in the downlink, which is given by (6.1.2) and γ_i is the balanced SINR achieved by the ith BS. First, the multiple linear constraints are put in (6.3.2c) into a single linear constraint by introducing the auxiliary variables $\mathbf{b} = [b_1 \mathbf{1}_K^T \cdots b_N \mathbf{1}_K^T]^T$. These auxiliary variables are updated based on the subgradient method described in [158]. Defining $P^{max} \triangleq \sum_{i=1}^K b_i P_i^{max}$, the problem in (6.3.2) can be written as

$$\max_{\mathbf{U},\mathbf{p}\succeq 0} \min_{i,j} \frac{\Gamma_{i,j}^{DL}(\mathbf{U},\mathbf{p})}{\rho_{i,j}}, \ i = N_1 + 1, \cdots N, \forall j \qquad (6.3.3a)$$

$$\text{s.t.} \quad \frac{\Gamma_{i,j}^{DL}(\mathbf{U},\mathbf{p})}{\rho_{i,j}} \geq \gamma_i, \ i = 1, \cdots N_1, \forall j \qquad (6.3.3b)$$

$$\mathbf{b}^T \mathbf{p} \leq P^{max}. \qquad (6.3.3c)$$

Based on the uplink-downlink duality [17, 156], for a given set of auxiliary variables, the dual uplink problem can be formulated as

$$\max_{\mathbf{U},\mathbf{q}\succeq 0} \min_{i,j} \frac{\Gamma_{i,j}^{UL}(\mathbf{U},\mathbf{q})}{\rho_{i,j}}, \ i = N_1 + 1, \cdots N, \forall j \qquad (6.3.4a)$$

$$\text{s.t.} \quad \frac{\Gamma_{i,j}^{UL}(\mathbf{U},\mathbf{q})}{\rho_{i,j}} \geq \gamma_i, \ i = 1, \cdots N_1, \forall j \qquad (6.3.4b)$$

$$\boldsymbol{\sigma}^T \mathbf{q} \leq P^{max}, \qquad (6.3.4c)$$

where $\Gamma_{i,j}^{UL}(\mathbf{U}, \mathbf{q})$ is the uplink SINR, given by (6.2.9). Now, this problem can be solved iteratively using the following three sub-problems: uplink power allocation, downlink power allocation and beamforming design.

6.3.1   Uplink Power Allocation for a given set of Beamformers

As the first step, for a given set of beamformers $\widehat{\mathbf{u}}_{i,j}\ \forall\ i,\ j$, in the uplink, we determine the uplink power allocation $\widehat{\mathbf{q}}$. The optimal power allocation should satisfy the following set of equations:

$$\frac{\Gamma_{i,j}^{UL}(\widehat{\mathbf{u}}_{i,j}, \widehat{\mathbf{q}})}{\rho_{i,j}} = \lambda(\widehat{\mathbf{U}}, P^{max}),\ i = N_1 + 1, \cdots, N, \forall j, \qquad (6.3.5a)$$

$$\frac{\Gamma_{i,j}^{UL}(\widehat{\mathbf{u}}_{i,j}, \widehat{\mathbf{q}})}{\rho_{i,j}} = \gamma_i(\widehat{\mathbf{U}}, P^{max}),\ i = 1, \cdots, N_1, \forall j, \qquad (6.3.5b)$$

$$\boldsymbol{\sigma}^T \widehat{\mathbf{q}} = P^{max}, \qquad (6.3.5c)$$

where $\lambda(\widehat{\mathbf{U}}, P^{max})$ is the balanced SINR of users in the cells that have not used their full transmit power. $\widehat{\mathbf{q}}_A = [\widehat{\mathbf{q}}_1 \cdots \widehat{\mathbf{q}}_{N_1}]^T$ is defined as the power allocation of users in the first N_1 cells and $\widehat{\mathbf{q}}_B = [\widehat{\mathbf{q}}_{N_1+1} \cdots \widehat{\mathbf{q}}_N]^T$ as the power allocation of users in the remaining cells. By evolving an earlier approach [159] to write power allocation of a set of users in terms of power allocation of the remaining users and solve the uplink power allocation in (6.3.4), it can be shown from (6.2.9) and (6.3.5) that $\widehat{\mathbf{q}}_{\mathbf{A}}$ and $\widehat{\mathbf{q}}_{\mathbf{B}}$ satisfy

$$\widehat{\mathbf{q}}_A = \mathbf{D}_A \boldsymbol{\Psi}_A \widehat{\mathbf{q}}_A + \mathbf{D}_A \boldsymbol{\Psi}_B \widehat{\mathbf{q}}_B + \mathbf{D}_A \mathbf{b}_A, \qquad (6.3.6a)$$

$$\frac{1}{\lambda}\widehat{\mathbf{q}}_B = \mathbf{D}_B \boldsymbol{\Psi}_C \widehat{\mathbf{q}}_A + \mathbf{D}_B \boldsymbol{\Psi}_D \widehat{\mathbf{q}}_B + \mathbf{D}_B \mathbf{b}_B, \qquad (6.3.6b)$$

$$P^{max} = \boldsymbol{\sigma}_A^T \widehat{\mathbf{q}}_A + \boldsymbol{\sigma}_B^T \widehat{\mathbf{q}}_B, \qquad (6.3.6c)$$

where $\boldsymbol{\sigma}_A = [\sigma_1 \cdots \sigma_{N_1}]^T$, $\boldsymbol{\sigma}_B = [\sigma_{N_1+1} \cdots \sigma_N]^T$,

$\mathbf{D}_A = \text{blkdiag}\{\gamma_1 \mathbf{D}^1, \cdots, \gamma_{N_1} \mathbf{D}^{N_1}\}$, $\mathbf{D}_B = \text{blkdiag}\{\mathbf{D}^{N_1+1}, \cdots, \mathbf{D}^N\}$,

$$\boldsymbol{\Psi}_A = \begin{bmatrix} \boldsymbol{\Psi}_1 & \dots & \boldsymbol{\Psi}_{1,N_1} \\ \vdots & \ddots & \vdots \\ \boldsymbol{\Psi}_{N_1,1} & \dots & \boldsymbol{\Psi}_{N_1} \end{bmatrix}, \tag{6.3.7}$$

$$\boldsymbol{\Psi}_B = \begin{bmatrix} \boldsymbol{\Psi}_{1,N_1+1} & \dots & \boldsymbol{\Psi}_{1,N} \\ \vdots & \ddots & \vdots \\ \boldsymbol{\Psi}_{N_1,N_1+1} & \dots & \boldsymbol{\Psi}_{N_1,N} \end{bmatrix}, \tag{6.3.8}$$

$$\boldsymbol{\Psi}_C = \begin{bmatrix} \boldsymbol{\Psi}_{N_1+1,1} & \dots & \boldsymbol{\Psi}_{N_1+1,N_1} \\ \vdots & \ddots & \vdots \\ \boldsymbol{\Psi}_{N,1} & \dots & \boldsymbol{\Psi}_{N,N_1} \end{bmatrix}, \tag{6.3.9}$$

$$\boldsymbol{\Psi}_D = \begin{bmatrix} \boldsymbol{\Psi}_{N_1+1} & \dots & \boldsymbol{\Psi}_{N_1+1,N} \\ \vdots & \ddots & \vdots \\ \boldsymbol{\Psi}_{N,N_1+1} & \dots & \boldsymbol{\Psi}_N \end{bmatrix}, \tag{6.3.10}$$

$\mathbf{b}_A = [\mathbf{b}_1, \cdots, \mathbf{b}_{N_1}]$ and $\mathbf{b}_B = [\mathbf{b}_{N_1+1}, \cdots, \mathbf{b}_N]$. According to [159], $(\mathbf{I} - \mathbf{D}_A \boldsymbol{\Psi}_A)$ is invertible and $(\mathbf{I} - \mathbf{D}_A \boldsymbol{\Psi}_A)^{-1}$ is nonnegative. Therefore, $\widehat{\mathbf{q}}_A$ can be written in terms of $\widehat{\mathbf{q}}_B$ as

$$\widehat{\mathbf{q}}_A = (\mathbf{I} - \mathbf{D}_A \boldsymbol{\Psi}_A)^{-1} \mathbf{D}_A \boldsymbol{\Psi}_B \widehat{\mathbf{q}}_B + (\mathbf{I} - \mathbf{D}_A \boldsymbol{\Psi}_A)^{-1} \mathbf{D}_A \mathbf{b}_A. \tag{6.3.11}$$

Substituting (6.3.11) in (6.3.6b), we can get

$$\frac{1}{\lambda} \widehat{\mathbf{q}}_B = \mathbf{B} \widehat{\mathbf{q}}_B + \mathbf{t}, \tag{6.3.12}$$

where

$$\mathbf{B} = \mathbf{D}_B\mathbf{\Psi}_C(\mathbf{I} - \mathbf{D}_A\mathbf{\Psi}_A)^{-1}\mathbf{D}_A\mathbf{\Psi_B} + \mathbf{D}_B\mathbf{\Psi}_D \tag{6.3.13a}$$

$$\mathbf{t} = \mathbf{D}_B\mathbf{\Psi}_C(\mathbf{I} - \mathbf{D}_A\mathbf{\Psi}_A)^{-1}\mathbf{D}_A\mathbf{b}_A + \mathbf{D}_B\mathbf{b}_B \tag{6.3.13b}$$

By substituting (6.3.11) into (6.3.6c), we obtain the following

$$\mathbf{c}^T\widehat{\mathbf{q}}_B = P^{max} - c, \tag{6.3.14}$$

where

$$\mathbf{c}^T = \boldsymbol{\sigma}_A^T(\mathbf{I} - \mathbf{D}_A\mathbf{\Psi}_A)^{-1}\mathbf{D}_A\mathbf{\Psi}_A + \boldsymbol{\sigma}_B^T \tag{6.3.15a}$$

$$c = \boldsymbol{\sigma}_A^T(\mathbf{I} - \mathbf{D}_A\mathbf{\Psi}_A)^{-1}\mathbf{D}_A\mathbf{b}_A. \tag{6.3.15b}$$

Multiplying both sides of (6.3.12) by $\mathbf{c}^T$ and using (6.3.14), we obtain

$$\frac{1}{\lambda} = \frac{1}{P^{max} - c}\mathbf{c}^T\mathbf{B}\widehat{\mathbf{q}}_B + \frac{1}{P^{max} - c}\mathbf{c}^T\mathbf{t}. \tag{6.3.16}$$

From (6.3.12) and (6.3.16), $\widehat{\mathbf{q}}_{ext} = [\widehat{\mathbf{q}}_B^T\ 1]^T$ satisfies

$$\frac{1}{\lambda}\widehat{\mathbf{q}}_{ext} = \begin{bmatrix} \mathbf{B} & \mathbf{t} \\ \frac{1}{P^{max}-c}\mathbf{c}^T\mathbf{B} & \frac{1}{P^{max}-c}\mathbf{c}^T\mathbf{t} \end{bmatrix} \widehat{\mathbf{q}}_{ext}. \tag{6.3.17}$$

Let

$$\mathbf{W}(\widehat{\mathbf{U}}) = \begin{bmatrix} \mathbf{B} & \mathbf{t} \\ \frac{1}{P^{max}-c}\mathbf{c}^T\mathbf{B} & \frac{1}{P^{max}-c}\mathbf{c}^T\mathbf{t} \end{bmatrix} \tag{6.3.18}$$

In [159], it was also shown that the conditions

$$\rho(\mathbf{D}_A\mathbf{\Psi}_A) \leq 1 \tag{6.3.19a}$$

$$c \leq P^{max} \tag{6.3.19b}$$

will imply that $\mathbf{W}(\widehat{\mathbf{U}})$ is a nonnegative matrix.

According to the Perron-Frobenious theory [160], among all eigenvalues of $\mathbf{W}(\widehat{\mathbf{U}})$, $\lambda_{max}(\mathbf{W}(\widehat{\mathbf{U}}))$ is the unique eigenvalue of maximum modulus. It is real and positive and associated with a positive eigenvector $\widehat{\mathbf{q}}_{ext}$. Hence, $\widehat{\mathbf{q}}_B$ can be obtained from $\widehat{\mathbf{q}}_{ext}$ by scaling the elements such that the last element is equal to one. Then $\widehat{\mathbf{q}}_A$ can be obtained using (6.3.11).

6.3.2   Beamformer Design for a Given Power Allocation

For a given power allocation, the optimal beamformers for all the users in the virtual uplink are obtained by maximizing independently the SINR of each user in the uplink. Hence, the same optimization problem as in (6.2.11) has to be solved.

6.3.3   Downlink Power Allocation and Iterative Solution

From the uplink-downlink duality, the same SINR values can be achieved in both the uplink and the downlink with the same set of beamformers and total power constraints. Thus, the uplink beamformers $\widehat{\mathbf{U}}$ obtained in the virtual uplink can be used to achieve the same SINR values in the downlink. Let us denote the power allocation in the downlink by $\mathbf{p} = [\mathbf{p}_A^T \ \mathbf{p}_B^T]^T$, where $\mathbf{p}_A$ and $\mathbf{p}_B$ are the downlink power allocation vectors for the users belonging to the first N_1 BSs that have already used their powers fully and the remaining $(N - N_1)$ BSs, respectively.

Similar to (6.3.12), (6.3.13a) and (6.3.13b), the following equations for the power allocation of the users in the downlink can be written:

$$\frac{1}{\lambda}\widehat{\mathbf{p}}_B = \mathbf{B}^{DL}\widehat{\mathbf{p}}_B + \mathbf{t}^{DL} \qquad (6.3.20)$$

where

$$\mathbf{B}^{DL} = \mathbf{D}_B \mathbf{\Psi}_B^T (\mathbf{I} - \mathbf{D}_A \mathbf{\Psi}_A^T)^{-1} \mathbf{D}_A \mathbf{\Psi_C}^T + \mathbf{D}_B \mathbf{\Psi}_D^T \qquad (6.3.21a)$$

$$\mathbf{t}^{DL} = \mathbf{D}_B \mathbf{\Psi}_B^T (\mathbf{I} - \mathbf{D}_A \mathbf{\Psi}_A^T)^{-1} \mathbf{D}_A \boldsymbol{\sigma}_A + \mathbf{D}_B \boldsymbol{\sigma}_B \qquad (6.3.21b)$$

It can be shown that $(\frac{1}{\lambda}\mathbf{I} - \mathbf{B}^{DL})$ is nonsingular and $(\frac{1}{\lambda}\mathbf{I} - \mathbf{B}^{DL})^{-1}$ is a nonnegative matrix. Using this assumption, the power allocation in the downlink is given by

$$\widehat{\mathbf{p}}_B = (\frac{1}{\lambda}\mathbf{I} - \mathbf{B}^{DL})^{-1} \mathbf{t}^{DL} \qquad (6.3.22a)$$

$$\widehat{\mathbf{p}}_A = (\mathbf{I} - \mathbf{D}_A \mathbf{\Psi}_A^T)^{-1} \mathbf{D}_A \mathbf{\Psi}_C^T \widehat{\mathbf{p}}_B + (\mathbf{I} - \mathbf{D}_A \mathbf{\Psi}_A^T)^{-1} \mathbf{D}_A \boldsymbol{\sigma}_A \qquad (6.3.22b)$$

Having obtained the beamformers and the power allocations for a given set of auxiliary variables $\mathbf{b}$, the auxiliary variables are updated according to the following subgradient method [156], [157].

$$b_i^{(m+1)} = b_i^{(m)} + t(\mathbf{1}_K^T \mathbf{p}_i^{(m)} - P^{max}), \ i = 1, 2, \cdots N, \qquad (6.3.23)$$

where t denotes the step size of the subgradient method. The algorithm stops when the following criterion is met

$$|b_i^{(m+1)}(\mathbf{1}_K^T \mathbf{p}_i^{(m)} - P^{max})| \leq \epsilon, \ i = 1, 2, \cdots N. \qquad (6.3.24)$$

The proposed algorithm is summarized in Table II.

Algorithm 2
---
1) **Initialize** $m \Leftarrow 0$, $n \Leftarrow 0$, $\widehat{\mathbf{q}}_A^{(0)}$, $\widehat{\mathbf{q}}_B^{(0)}$ $\mathbf{b}^{(0)}$, t, ϵ

2) **Repeat**

3) $m \Leftarrow m + 1$

4)   **Repeat**

5) $n \Leftarrow n + 1$

6)   Solve (6.2.11) using $\widehat{\mathbf{q}}_A^{(n-1)}$, $\widehat{\mathbf{q}}_B^{(n-1)}$ to obtain $\widehat{\mathbf{U}}^{(n-1)}$

7)   Compute $\mathbf{D}_A^{(n-1)}$, $\mathbf{D}_B^{(n-1)}$, $\mathbf{\Psi}_A^{(n-1)}$, $\mathbf{\Psi}_B^{(n-1)}$,
          $\mathbf{\Psi}_C^{(n-1)}$, $\mathbf{\Psi}_D^{(n-1)}$, $\mathbf{b}_A^{(n-1)}$, $\mathbf{b}_B^{(n-1)}$ using $\widehat{\mathbf{U}}^{(n-1)}$

8)   Solve (6.3.17) and obtain $\lambda^{(n)}$ and $\widehat{\mathbf{q}}_B^{(n)}$

9)   Obtain $\widehat{\mathbf{q}}_A^{(n)}$ from (6.3.11)

10)   **Until** $\lambda^{(n-1)} - \lambda^{(n)} \leq \epsilon$

11) $\lambda^* = \lambda^{(n)}$ and $\widehat{\mathbf{U}}^* = \widehat{\mathbf{U}}^{(n-1)}$

12)   Obtain $\widehat{\mathbf{p}}_B^{(m)}$, $\widehat{\mathbf{p}}_A^{(m)}$ using (6.3.22a) and (6.3.22b), respectively.

13)   Update auxiliary variables using (6.3.23)

14) **Until** (6.3.24) is satisfied.

Table 6.2. Algorithmic solution of the mixed SINR balancing based beamforming design (6.3.2)

6.4 Complexity and Convergence Analysis

6.4.1 Complexity Analysis

For a given set of auxiliary variables, the complexity of the proposed algorithms in Table I and Table II mainly depends on the complexities of the matrix inversion and the eigenvalue decomposition. For a given $l \times l$ matrix, the required arithmetic operations to determine its inverse and the eigenvectors are given by $\mathcal{O}(l^3)$ and $\mathcal{O}(l^3 + (l\log^2 l)\log b))$, respectively, where b is the relative error bound [161]. Based on this, the number of arithmetic operations required for every iteration of the algorithm in Table I and Table II are shown in Table III and Table IV, respectively. Only steps 6, 8 and 11 of Algorithm 1 and steps 6, 8, 9 and 12 of Algorithm 2 require matrix inversion or eigenvalue decomposition. Hence the total arithmetic operations required for each iteration is the summation of arithmetic operations needed for matrix inversion and the eigenvalue decomposition. In both tables, n denotes the number of iterations required to satisfy the stopping criterion, i.e., step 9 in Table I and step 10 in Table II. It has been observed through simulations that the average number of iterations n is equal to two.

Table 6.3. Complexity of the 1st stage of the algorithm

	Matrix Inversion
Step 6	$nKN\mathcal{O}(M^3)$

	Eigenvalue Decomposition
Step 6	$nKN\mathcal{O}\left[M^3 + (M\log^2 M)\log b\right]$
Step 8	$n\mathcal{O}\{(NK+1)^3 + \left[(NK+1)\log^2(NK+1)\right]\log b)\}$
Step 11	$\mathcal{O}\{(NK+1)^3 + \left[(NK+1)\log^2(NK+1)\right]\log b)\}$

6.4.2   Convergence Analysis

Here, we discuss the convergence of the auxiliary variables. Assume that the ith constraint (i.e., $\mathbf{1}_K^T \mathbf{p}_i \leq P_i^{max}$) is not satisfied at the first iteration. Since $\mathbf{1}_K^T \mathbf{p}_i^{(1)} > P_i^{max}$, the corresponding auxiliary variable a_i will be increased based on the subgradient method as follows:

$$a_i^{(m+1)} = a_i^{(m)} + t(\mathbf{1}_K^T \mathbf{p}_i^{(m)} - P_i^{max}). \tag{6.4.1}$$

At the next iteration, the corresponding uplink noise covariance will be increased as $a_i^{(1)} > a_i^{(0)}$ (i.e., $\Omega_i = a_i \mathbf{I}$). This reduces the achievable SINR at the next iteration and yields into less power allocation (i.e., $\mathbf{1}_K^T \mathbf{p}_i^{(1)} \geq \mathbf{1}_K^T \mathbf{p}_i^{(2)}$). Hence, the corresponding auxiliary variable a_i increases with the iteration number n until the stopping criterion $|a_i^{(m+1)}(\mathbf{1}_K^T \mathbf{p}_i^{(m)} - P_i^{max})| \leq \epsilon$ is satisfied as follows:

$$a_i^{(0)} \leq a_i^{(1)} \leq \cdots \leq a_i^{(n)}. \tag{6.4.2}$$

The stopping criterion is satisfied when either a_i reaches a small value or the associated constraint is satisfied. Since, one of these conditions is achieved, the auxiliary variable adaptation converges. The similar argument holds for the scenario where the constraint is satisfied at the first iteration. In this case, the corresponding auxiliary variable will be decreasing with the iteration. This monotonic convergence of the auxiliary variable can be observed

Table 6.4. Complexity of the 2nd stage of the algorithm

	Matrix Inversion
Step 6	$nKN\mathcal{O}(M^3)$
Step 8	$-$
Step 9	$n\mathcal{O}\{[N(K-1)+1]^3\}$
Step 12	$2\mathcal{O}\{[N(K-1)+1]^3\}$

	Eigenvalue Decomposition
Step 6	$nKN\mathcal{O}\left[M^3 + (M\log^2 M)\log b\right]$
Step 8	$n\mathcal{O}\{[(N-1)K+1]^3 + \left[(N-1)K+1)\log^2[(N-1)K+1]\right]\log b)\}$
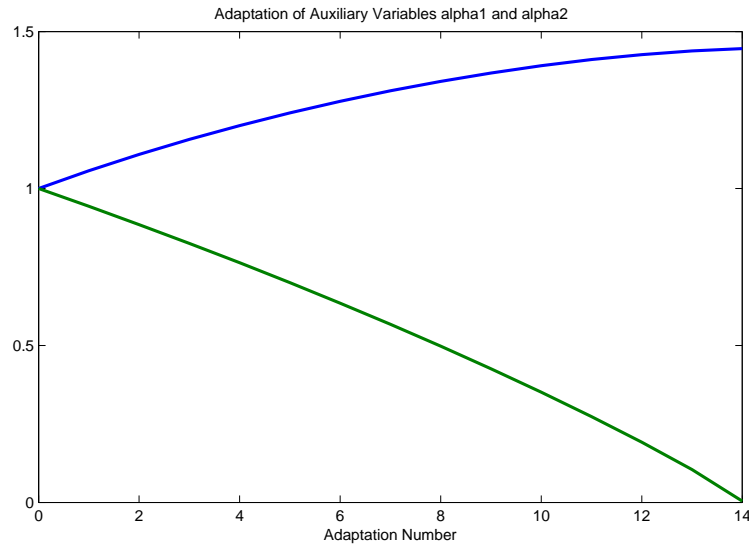
Figure 6.2. Evolution of auxiliary variables a_1 and a_2 against adaptation number.

in Figure 6.2.

Having established the monotonic convergence of the auxiliary variables, SINR convergence of the inner-loop and the outer-loop of the proposed algorithm is investigated. In the inner-loop, the beamformers and the power allocations are determined for a given set of auxiliary variables whereas the outer-loop updates the auxiliary variables through a subgradient method. In the inner-loop, the optimal beamformers and the power allocations are iteratively determined by maximizing the virtual uplink SINR of each user [17, 156]. Hence, the virtual uplink SINR increases at each iteration. Since the available transmit power is limited, the achievable uplink SINR is upper bounded. Hence, the inner-loop converges to a fixed solution. In addition, the inner-loop provides the optimal solution for a given set of auxiliary variables [17, 156]. If the auxiliary variables are chosen to satisfy the constraints through the outer-loop, then the beamformers and power allocation obtained from the inner-loop is optimal solution for the original problem. As discussed earlier, the auxiliary variables are updated to satisfy the constraints and they either monotonically increase or decrease.

Hence, the proposed iterative algorithms yield optimal solution which has been verified in the simulation section using the SDP formulation of the power minimization problem.

6.4.3 Beamforming Design based on SDP

For a given set of target SINRs, beamformer weight vectors can be designed using the following optimization [162]. footnotesize

$$\min_{\mathbf{w}_{i,j}} \sum_{i=1}^{N} \sum_{j=1}^{K} ||\mathbf{w}_{i,j}||_2^2$$
s.t.
$$\frac{\mathbf{w}_{i,j}^H \mathbf{R}_{i,i,j} \mathbf{w}_{i,j}}{\sum_{l=1,l\neq j}^{K} \mathbf{w}_{i,l}^H \mathbf{R}_{i,i,j} \mathbf{w}_{i,l} + \sum_{m=1,m\neq i}^{N} \sum_{l=1}^{K} \mathbf{w}_{m,l}^H \mathbf{R}_{m,i,j} \mathbf{w}_{m,l} + \sigma_{i,j}^2} \geq \gamma_{i,j},$$
$$\forall i, j, \tag{6.4.3}$$

where $\mathbf{w}_{i,j} = p_{i,j}\mathbf{u}_{i,j}$ includes both the power allocation and beamformer vector. The above optimization can be converted into convex form and solved using SDP [162], hence it provides optimum solution. This formulation is used to verify the optimality of the results in Section VI. However, it should be highlighted that the SINR balancing problem cannot be solved directly using the SDP approaches because the optimal balanced SINR values are not known a priori.

6.5 Simulation Results

First, the performance of the SINR balancing algorithm without consideration to balancing SINR of each BS at different levels is assessed i.e., balancing SINR of all users in all BSs together but subject to individual BS power constraints. A scenario with two BSs has been considered, each serving two users. The BSs and the users consist of four antennas and single antenna, respectively. The noise power at each user terminal is set to 0.01. Random

channels for each user from both BSs using zero mean complex Gaussian random variables have been generated. The average channel gain from every BS to the corresponding users was set to one while the gain of channels from each BS to users in other BS (co-channel gain) was normalized to 0.1 so that the average power gain for the co-channels is 0.01. The maximum power limitations at the first and the second BSs were set to 0.1W and 10W, respectively. In all the simulations, the priority factor is set to $\rho_{i,j} = 1$ for all users. The SINR balancing algorithm described in Section III has been employed. The stopping criterion threshold ϵ in (6.2.13) and (6.3.24) is set to 0.1 and the adaptation step size t in (6.2.12) and (6.3.23) to 0.015. The adaptation of auxiliary variables for a given set of random channels is shown in Figure 6.2. The adaptation is stopped when one of the auxiliary variables goes to zero or when there is insignificant change in subsequent values of both the auxiliary variables. In this example, the second auxiliary variable converges to zero. At the convergence, it was observed that the first BS uses all available power of 0.1W while the second BS uses only 0.5425W. i.e., the second BS is not able to use all its available power, hence the constraint associated with the second BS power is inactive confirming as to why the second auxiliary variable converges to zero.

After the convergence, it is observed that for the random channels used, the SINRs of users belonging to both the BS were balanced at 15.35. In order to confirm the optimality of this result, the SDP based beamforming design [162] as described in Section V has been used and the SINR constraints for all four users are set to 15.35 and the power constraints for the two BSs at 0.1W and 10W, respectively as in the previous simulation. It is observed that the SDP based design consumed the same power for BS1 and BS2, as obtained by the proposed algorithm i.e., 0.1W and 0.5425W. Also, it is observed that the beamformers obtained using the proposed method and the SDP method are identical. When the SINR target of all four users has

Table 6.5. Power allocations and the achieved SINRs using the proposed method.

	Power Allocation				Total Power	
	BS1		BS2		BS1	BS2
Channels	User 1	User 2	User 1	User 2		
Channel 1	0.0634	0.0366	7.0120	2.9868	0.1	10
Channel 2	0.0466	0.0534	3.2116	6.7878	0.1	10
Channel 3	0.0581	0.0419	3.9419	6.0579	0.1	10
Channel 4	0.0295	0.0705	3.2601	6.7399	0.1	10
Channel 5	0.0479	0.0521	1.4086	8.5912	0.1	10
	Achieved SINR (in linear scale)				Total Power	
	BS1		BS2		BS1	BS2
Channels	User 1	User 2	User 1	User 2		
Channel 1	15.35	15.35	34.54	34.54	0.1	10
Channel 2	18.52	18.52	341.46	341.46	0.1	10
Channel 3	22.37	22.37	806.82	806.82	0.1	10
Channel 4	18.95	18.95	659.56	659.56	0.1	10
Channel 5	15.08	15.08	280.01	280.01	0.1	10

been increased to slightly above 15.35, the SDP confirmed that the problem is infeasible. This confirms that the balanced SINR obtained using our method is optimum. It should be noticed that for a given set of target SINRs, the beamforming design based on SDP is optimum. However the SDP method cannot be used directly to design beamformers for SINR balancing, because the balanced SINR value is not known a priori. SDP method is used only to verify the optimum resource allocation for a given set of target SINRs which in this case is the balanced SINR obtained using the method proposed.

Having confirmed the optimality of the first algorithm, the second algorithm described in Section IV has been used to optimize the SINR of users in the second BS to a higher level. For the set of random channels considered, as can be seen for the first Channel of Table 6.5 the BS2 achieved a higher SINR of 34.54. The power allocations and the balanced SINR values for each user at each cell for five different set of random channels have been obtained, as depicted in Table 6.5. The results with the solutions obtained using the SDP approach in Table 6.6 have been compared. The random channels used

Table 6.6. Target SINRs and the user power consumptions using the SDP-Based Method.

	Target SINR (in linear scale)				Total Power	
	BS1		BS2		BS1	BS2
Channels	User 1	User 2	User 1	User 2		
Channel 1	15.35	15.35	34.54	34.54	0.1	10
Channel 2	18.52	18.52	341.46	341.46	0.1	10
Channel 3	22.37	22.37	806.82	806.82	0.1	10
Channel 4	18.95	18.95	659.56	659.56	0.1	10
Channel 5	15.08	15.08	280.01	280.01	0.1	10
	Power Allocation				Total Power	
	BS1		BS1		BS2	
Channels	User 1	User 2	User 1	User 2		
Channel 1	0.0634	0.0366	7.0120	2.9868	0.1	10
Channel 2	0.0466	0.0534	3.2116	6.7878	0.1	10
Channel 3	0.0581	0.0419	3.9419	6.0579	0.1	10
Channel 4	0.0295	0.0705	3.2601	6.7399	0.1	10
Channel 5	0.0479	0.0521	1.4086	8.5912	0.1	10

in both Tables are the same and the SINR targets in Table 6.6 are the same SINR values achieved in Table 6.5. For example, for Channel 2 in Table 6.6, the SINR targets for the two users in the first BS is set to [18.52 18.52] and for the two users in the BS2 to [341.46 341.46] as obtained with our method in Table 6.5. Comparing Table 6.5 and 6.6, for the same SINR targets, the power allocation obtained using the SDP approach and the proposed method is the same. In addition, it is observed that both methods provided the same set of beamformers. For all five random channels, for the verification of results using SDP, the SINR target of users has been increased slightly above the balanced SINR values obtained by the proposed method and observed that SDP optimization turns out to be infeasible. Since the SDP method for the beamforming design for a given set of SINR targets is optimum [162], and since both the proposed method and the SDP method provide the same result, as in Tables 6.5 and 6.6, the proposed method of balancing SINRs of users at different BSs to different levels is optimum.

Figure 6.3 depicts the balanced SINRs (in dB) that the users achieve
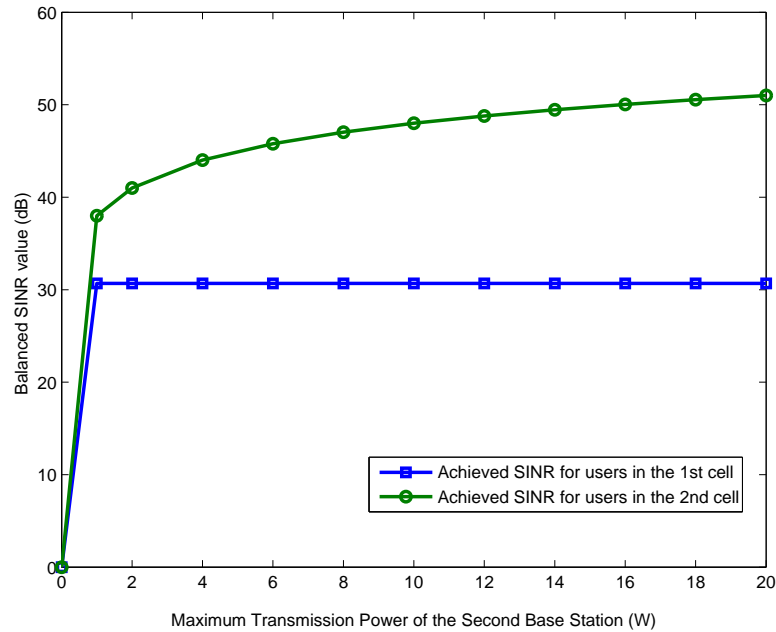
Figure 6.3. Achieved SINR (dB) of the users in the 1st and 2nd cell versus the transmit power constraint of the 2nd cell.

in the first and the second cells against the maximum transmit power of the second BS for a given set of random channels. The transmit power of the first BS has been set to 0.1W and the maximum power of the BS2 has been varied. As observed in Figure 6.3, the balanced SINR achieved from the users belonging to the second cell grows substantially as the maximum transmit power of BS2 increases, while the balanced SINR of the first cell remains at the same value.

For the results presented so far, four antennas for each BS have been used. Since there are only four users altogether in both BSs, each BS has adequate degrees of freedom to mitigate interference satisfactorily, hence they are able to fully utilize the available transmit power to maximize the balanced SINR as high as possible. However, if the number of antennas are limited, while one of the BS uses its full transmit power (usually the one with lower maximum transmit power), the other BSs may not fully use its transmit power as otherwise significant interference may be introduced to the

users in the first BS. For example, a scenario has been simulated as before for a set of random channels, but with three antennas for each BS. The first BS power was set to 1W, while the second BS power was set to 10W. The first BS used all its transmit power of 1W and achieved a balanced SINR of 57.81 for its users, however, the second BS used only 1.7279W and achieved a balanced SINR of 57.81 for its users. The second BS could not increase its transmit power beyond 1.7279W, because it may introduce interference to the first BS and it will result into lower balanced SINR for the first BS. However, a slight decrease in the SINR target for the first BS will increase the second BS transmit power utilization and will enhance the second BS users balanced SINR significantly. This tradeoff has been depicted in Figure 6.4. Again, the proposed algorithm has the ability to first determine the maximum possible balanced SINR that the first BS could achieve and secondly to balance and maximize the SINR of the users in the second cell, while satisfying a wide range of target SINRs for the users in the first cell.

Finally, the performance of the proposed algorithm has been compared with that of an extension of the work proposed in [2]. The work in [2] proposed to perform SINR balancing of all users in all the cells with multiple BS power constraints. Observing that after SINR balancing, certain BSs may not have used all the available transmit power, as also described in the proposed algorithm, the authors in [2] proposed to use any excess power available at these BSs to increase the SINR of the best performing users in those cells without affecting the SINR of other users. However, the work in [2] did not aim to balance the SINR of users in other BSs. Therefore, for the purpose of comparison, the work in [2] is extended to perform a second level of SINR balancing for the users in the BS that has not used all available power, by distributing the remaining power among all users in that BS such that the SINR of users in that BS is balanced and maximized. Assume, that the ith BS used all its transmit power. Suppose if we wish to use all the
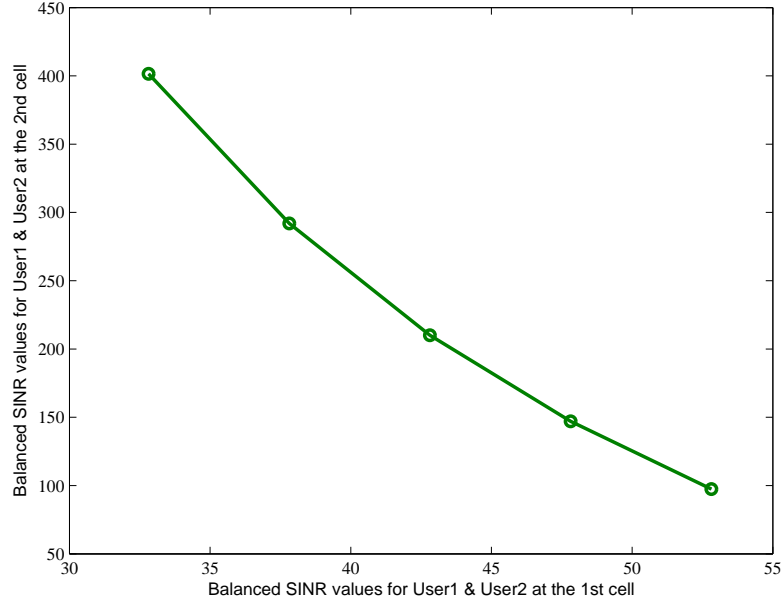
Figure 6.4.   Achieved SINR (dB) of the users in the 2nd cell for decreasing values of target SINRs (dB) for the users in the 1st cell with M=3.

excess transmit power of the jth BS to its kth user, according to the work in [2], a new beamformer should be designed for the kth user as follows:

$$\mathbf{u}_{j,k}^{new} = \mathbf{u}_{j,k} + a_{j,k}e^{j\theta_k}\mathbf{v}_{j,k}^{ZF} \tag{6.5.1}$$

where $\theta_{j,k} = \angle(\mathbf{u}_{j,k}^H \mathbf{h}_{j,j,k})$ and $a_{j,k}$ is determined such that $||\mathbf{u}_{j,k}^{new}|| = P_j^{max}$ and $\mathbf{v}_{j,k}^{ZF}$ is the projection of $\mathbf{h}_{j,j,k}$ onto the complement of the column space of $\tilde{\mathbf{H}}_{\mathbf{j,k}} = [\mathbf{h}_{j,i,1}\cdots\mathbf{h}_{j,i,K}\mathbf{h}_{j,j,1}\cdots\mathbf{h}_{j,j,k-1},\mathbf{h}_{j,j,k+1},\cdots\mathbf{h}_{j,j,K}]$, i.e., $\mathbf{v}_{j,k}^{ZF} = (\mathbf{I} - \tilde{\mathbf{H}}_{\mathbf{j,k}}(\tilde{\mathbf{H}}_{\mathbf{j,k}}^H\tilde{\mathbf{H}}_{\mathbf{j,k}})^{-1}\tilde{\mathbf{H}}_{\mathbf{j,k}}^H)\mathbf{h}_{j,j,k}$. However, since the aim is to balance SINR of all users in the jth cell, rather than allocating all the excess power to only one user, the work in [2] is modified and extra beamformers for all the users are designed and the total excess power is distributed among all users in the jth cell in such a way the SINRs of all users in the jth cell are balanced. Bisection method has been used to determine the optimum excess power allocation. Then the performance of our proposed algorithm has been

compared with the above said extension of [2]. In order to compare the proposed method with the method described in [2], a scenario with 2 BSs and 3 users for each BS has been considered. The transmit power at the first and second BSs were set to 0.1W and 10W respectively. Five sets of different random channels were generated as shown in Table 6.7. As can be seen, for example in the first row of Table 6.7, the second BS achieved a higher SINR of 34.2 using the proposed method as compared to the SINR of 25.89 achieved using the extended work of [2]. For all the set of random channels depicted in Table 6.7, the SINR values achieved using the proposed method are higher than that achieved using the extended work of [2]. This is because, the work in [2] proposed to use the remaining power by using extra beamforming vector components that were obtained by projecting each user channel onto the complement of the column space of all other user channels. This can be viewed as zero forcing multiplexing. Due to zero forcing there is loss of flexibility, however, the proposed work does not impose this strict zero forcing condition, but aims to maximize minimal SINR directly using an optimization framework. Hence, it outperforms the extended work of [2].

  Simulation Remarks: Some additional key points that have been used to assist the simulation are summarized. If any of the BSs does not use all its transmit power, the associated auxiliary variable during the adaptation as in (6.2.12) will converge to zero. In this case, $\mathbf{\Omega}_i$ in the SINR of the equivalent uplink in (6.2.9) is zero. Suppose, the number of antennas at the BS is greater than the number of interference components shown in the denominator of the SINR in (6.2.9), the matrix $\mathbf{M} = (\sum\limits_{\substack{l=1 \\ l \neq j}}^{K} q_{i,l}\mathbf{R}_{i,i,l} + \sum\limits_{\substack{m=1 \\ m \neq i}}^{N} \sum\limits_{l=1}^{K} q_{m,l}\mathbf{R}_{i,m,l} + \mathbf{\Omega}_i)$ will become rank deficient. Therefore, the beamformer for this case cannot be designed using the generalized singular value solution. For this case, the beamformer is designed as the vector that is in the null space of the above matrix $\mathbf{M}$. Also, when the number of antennas is higher than that of the users, the BS has the ability to manage interference if possible through

Table 6.7. The achieved balanced SINRs using the extended work of [2] and our proposed method

	Extended work of [2]	
Channels	SINR of U1-U3 i.e., BS1	SINR of U4-U6 i.e., BS2
Channel 1	14.4	25.9
Channel 2	9.5	186.5
Channel 3	8.7	337.1
Channel 4	13.7	302.7
Channel 5	8.8	151.5
	Our method	
Channels	SINR of U1-U3 i.e., BS1	SINR of U4-U6 BS2
Channel 1	14.4	34.2
Channel 2	9.5	199.9
Channel 3	8.7	351.4
Channel 4	13.7	317.3
Channel 5	8.8	154.9

steering appropriate nulls. In this case, the multiple BSs have the ability to perform SINR balancing efficiently and the BSs are likely to use the maximum available transmit power. However, with fewer antennas, if the whole network is severely interference limited, certain BSs may not have the flexibility to trade off the power between users and use all available transmit power. The auxiliary variables associated with these BS will be non-zero in this case.

6.6 Conclusion

A coordinated multi-cell beamforming for multiple cells with various transmit power constraints has been considered. The aim was to balance and maximize the SINR of users in various cells to different maximum possible levels. This facilitates BSs with better channel conditions and more power to balance SINR of their serving users to a higher level than the balanced SINR of users attached to BSs with either low transmit power or bad channel conditions. The algorithm was solved sequentially by balancing SINR of users in various cells with constraints on transmit power and already achieved balanced SINRs. The results have been compared with SDP based beamforming design for optimality.

Chapter 7

SUMMARY, CONCLUSION

AND FUTURE WORK

In this chapter, the novel results of this thesis and the conclusions that can be drawn from them are summarized, followed by a discussion on future work.

7.1 Summary and Conclusions

The focus of this thesis has been on proposing and analyzing signal processing algorithms, in particular mathematical optimization techniques, for the enhancement of coverage and capacity of wireless networks. There are various techniques that have the potential to enhance capacity and coverage, for example spatial diversity techniques, cognitive radio techniques, relay networks, heterogenous networks and advanced error control coding techniques. Among these available techniques, the spatial diversity and cognitive radio techniques have the ability to enhance the usage of the spectrum and to increase the capacity substantially. Relay technology has the ability to enhance the coverage and also to minimize the power consumption for transmission of signals in the network. The emphasize on the work presented in this thesis has been on spatial diversity techniques and relay network as well as cognitive radio technology. In particular, various mathematical optimization techniques have been proposed for optimal resource allocation and

spatial diversity for wireless relay networks. The work also considered designing such systems under interference constraints, so that is applicable to cognitive radio networks as well. Having designed various relay networking technologies, the final part of the thesis focused on coordinated multi-cell processing, in particular coordinated beamformer design for enhancing capacity.

The first contributing chapter "Wireless Peer-to-Peer Relay Networks" had four contributions on peer-to-peer wireless networks. There have been various works available in the literature on designing wireless relay networks to achieve a certain QoS. However the work considered in this thesis focused on designing wireless relay networks while ensuring users' fairness in the form of maximizing the worst case user SINR in the network. This is termed as SINR balancing. The work considered both one way and two way relay networks with and without consideration to interference leakage to primary users as in a cognitive radio network. Accordingly, the work considered multiple peer-to-peer users and a number of spatially distributed relays. The aim of the relay is to receive the signal from multiple users, amplify and rotate the phase and forward them to the destinations. In that process, the relay is expected to perform spatial diversity, as the aim is to maximize the worst case SINR of the users at the destination. This problem was solved using a combination of SDP and GP and the simulation results shown multiple spatially distributed relays have the ability to perform spatial multiplexing and forward the signal to the destination, even in the presence of multiple users and even if the relays have only one antenna. This work has been extended to two-way relay networks as well. Finally the work has been extended to a CR network, where the aim of the relay is not only to perform spatial multiplexing and to forward the signal to the destination, however the relay should also need to ensure the interference leakage to a number of primary users is below a specific threshold.

In the second contributing chapter, we extended the peer-to-peer wireless relay network to a relay that consists of multiple antennas and the peer to peer users have multiple antennas at both the transmitting and the receiving end. Hence, we used MMSE as the criterion for optimization. The aim of the MIMO relay as well as the transceivers of the peer-to-peer users is to minimize the weighted sum MMSE of the network for a given total network power. This problem has been solved using uplink-downlink duality and SOCP. According to the uplink-downlink duality, the receiver is fixed and the transmitter is designed by converting the problem into a virtual uplink problem. Once the transmitter has been designed, it is fixed and the receivers are designed directly using MMSE technique. The optimum relay transceiver matrix has been designed using SOCP. Hence an iterative method was proposed. It is understood that the overall problem is not convex, however this iterative method has provided reasonably good performance for a variety of random channels.

The third contributing chapter focused again on relay network problem. Specifically a BS with multiple antennas serving a number of users directly and another set of users through a MIMO relay has been considered. The aim is to jointly design the beamformers required at the BS and at the relay to ensure that the users served by both the BS and the relay achieve a set of target SINRs (for a given transmission power at the BS and at the relay). According to this setup, for each complete transmission, there are two time slots. In the first time slot, the BS sends information to the users that serves directly and to the relay. In the second time slot, both the BS and the relay transmit signal to their corresponding users. The signal that needs to be transmitted to the corresponding users of the relay, has been transmitted by the BS during the first time slot. The work available in the literature, considered the signal received by the users in the second time slot that are directly served by the BS from the relay as interference. However,

as this is the very same signal that has been transmitted by the BS in the first time slot, it should not be considered as interference, as the interference structure is already known at the BS. This concept has been exploited and a novel technique has been proposed. The BS in the second time slot not only designs beamformer to send information to the users it serves, however also to mitigate interference caused by the relay as the BS knows the exact interference structure. This technique has been shown to outperform other techniques available in the literature.

Final contributing chapter has been on multi-cell beamforming. Traditionally, spatial diversity or beamforming techniques, have been designed for each BS considering the interference from other BSs as noise. However, coordinated multi-cell processing aims to design beamformers jointly for a set of BSs. There are various beamforming techniques known in the literature, including designing beamformers to achieve a set of SINR targets for users served by various BSs and also to maximize the worst case user SINR, namely SINR balancing. However, all the techniques available in the literature on SINR balancing based beamformer design aimed to balance the SINR of all users in all cells together. This is however, not desirable, because users that are in a BS with good channel conditions or higher transmission power may be disadvantaged as the balanced overall SINR values are limited by the worst case user in the worst case BS. Therefore, a new approach has been proposed that aims to balance SINR of users in various cells to various levels. This has been solved using an SINR target based SINR balancing technique. The results have been proved to be optimal using a set of simulation results that compares the performance with that obtained using SDP techniques.

In summary this thesis has investigated various spatial diversity techniques and resource allocation techniques for enhancing the capacity and coverage of wireless networks. The solutions were based on exploitation of

uplink-downlink duality and convex optimization techniques, in particular SDP, SOCP and GP.

7.2 Future Work

The work presented in this thesis can be extended in a number of ways. First of all, all the techniques presented in the thesis considered perfect channel state information. It is possible to design spatial diversity techniques and resource allocation techniques considering certain level of uncertainty in the channel state information. Various forms of robust optimization techniques can be applied to solve this problem. In Chapter 3, we considered the relays to have single antennas. It is possible to employ multiple antennas at the relay and can be solved within the context of one and two way relay networks. Employing multiple antennas has the potential to enhance the capacity and coverage further. The work in Chapter 4 can extended to MIMO peer to peer network with multiple relays with multiple antennas. This has also the potential to enhance the capacity and coverage further. The work presented in Chapter 5 considered single antennas for the user terminals. Again it is possible to extend this work to multiple antennas based users. In this case MIMO based transmission from the BS to users as well as from the relay to users will be established. In this setup, instead of SINR target either data rate target or MMSE target should be considered. The work in Chapter 6 could also be extended in a number of ways. First we considered multiple BSs with coordinated multi-cell processing to balance the SINR of users in various cells to different levels. This technique can be extended to achieve distributed design, either using Lagrangian duality or using game theoretic methods. This work can be extended further to consider multiple antennas at the users terminals. Here instead of SINR balancing, either data rate balancing or MMSE balancing criteria should be considered.

References

[1] Cisco Visual Networking Index: Global mobile data traffic forecast update *National Telecommunications and Information Administration, U.S. Department of Commerce*, February 2013, Available: `http://www.cisco.com/en/US/solutions/collateral/ns341/ns525/ns537/ns705/ns827/white_paper_c11-520862.html`.

[2] Y. Huang, G. Zheng, M. Bengtsson, K.-K. Wong, L. Yang, and B. Ottersten, "Distributed multicell beamforming design approaching pareto boundary with max-min fairness," *IEEE Transactions on Wireless Communications*, vol. 11, no. 8, pp. 2921–2933, August 2012.

[3] U. S. frequency allocation chart *National Telecommunications and Information Administration, U.S. Department of Commerce*, October 2003, Available: `www.ntia.doc.gov/osmhome/allochrt.pdf`.

[4] M. Cave, C. Doyle, and W. Webb, *Essentials of Modern Spectrum Management.* Cambridge University Press, 2007.

[5] T. S. Rappaport, *Wireless Communications: Principals and Practice.* Upper Saddle River, NJ: Prentice Hall, 2 ed., 1996.

[6] M.-S. Alouini and A. J. Goldsmith, "Area spectral efficiency of cellular mobile radio systems," *IEEE Transactions on Vehicular Technology*, vol. 48, pp. 1047–1066, Jul. 1999.

[7] The Nobel Prize in Physics 1968, Available: `www.nobelprize.org/nobelprizes/physics/laureates/1968/alvarezbio.html`.

[8] B. D. V. Veen and K. M. Buckley, "Beamforming: a versatile approach to spatial filtering," *IEEE ASSP Magazine*, vol. 5, pp. 4–24, 1988.

[9] O. L. F. III, "An algorithm for linearly constrained adaptive array processing," vol. 60, pp. 926–935, Aug. 1972.

[10] B. Widrow, P. E. Mantey, L. Griffiths, and B. B. Goode, "Adaptive antenna systems," *Proc. IEEE*, vol. 55, pp. 2143–2159, 1967.

[11] R. Monzingo and T. Miller, "Introduction to adaptive arrays," *Wiley and Sons, New York*, 1980.

[12] D. Gerlach and A. Paulraj, "Adaptive transmitting antenna array with feedback," *IEEE Signal Processing Letters*, vol. 1, no. 10, October 1994.

[13] D. J. Love, R. W. Heath, and T. Strohmer, "Grassmannian beamforming for multiple-input multiple-output wireless systems," *IEEE Transactions on Information Theory*, vol. 49, no. 10, pp. 2735–2747, October 2003.

[14] Z. Xiong, R. Krishna, K. Cumanan, and S. Lambotharan, "Grassmannian beamforming and null space broadcasting protocols for cognitive radio networks," *IET Signal Processing*, vol. 5, no. 5, pp. 451–460, August 2011.

[15] R. T. Derryberry, S. D. Gray, D. M. Ionescu, G. Mandyam, and B. Raghothaman, "Transmit diversity in 3G CDMA systems," *IEEE Communications Magazine*, April 2002,.

[16] M. Bengtsson and B. Ottersten, "Optimal downlink beamforming using semidefinite optimization," *In Proceedings of 37th Annual Allerton Conference on Communication, Control, and Computing*, p. 987996, September 1999.

[17] M. Schubert and H. Boche, "Solution of the multiuser downlink beam-forming problem with individual SINR constraints," *IEEE Transactions on Vechicular Technology*, vol. 53, no. 1, pp. 18–28, January 2004.

[18] A. Naguib, A. Paulraj, and T. Kailath, "Capacity improvement with base-station antenna arrays in cellular CDMA," *IEEE Transactions on Vechicular Technology*, vol. 43, no. 3, pp. 691–698, August 1994.

[19] S. M. Alamouti, "A simple transmit diversity technique for wireless communications," *IEEE Journal on Select Areas in Communications*, vol. 16, pp. 1451–1468, October 1998.

[20] V. Tarokh, N. Seshadri, and A. R. Calderbank, "Space-time codes for high data rate wireless communication: Performance criterion and code construction," *IEEE Transactions on Information Theory*, vol. 44, no. 2, pp. 744–765, March 1998.

[21] A. Goldsmith, S. A. Jafar, N. Jindal, and S. Vishwanath, "Capacity limits of MIMO channels," *IEEE Journal on Selected Areas in Communications*, vol. 21, no. 5, pp. 691–698, June 2003.

[22] M. Costa, "Writing on dirty paper," *IEEE Transactions on Information Theory*, vol. 29, pp. 439 – 441, May 1983.

[23] G. T. . v7.1.0, "Physical layer aspects for evolved univeral terrestrial radio access (UTRA)," 2006.

[24] I. . W. S. on Mobile Multihop Relay, "IEEE 802 tutorial: 802.16 mobile multihop relay," `http://www.ieee802.org/16/sg/mmr/docs/80216mmr06006.zip`.

[25] S. Haykin, "Cognitive radio: brain-empowered wireless communications," *IEEE Journal on Selected Areas in Communications*, vol. 23, pp. 201–220, 2005.

[26] J. Proakis and M. Salehi, *Digital Communications*. McGraw-Hill Science, 5 ed., 2007.

[27] Hayes and H. Monson, "Statistical digital signal processing and modeling," *John Wiley and Sons*, 1996.

[28] I. H. Kim and D. J. Love, "On the capacity and design of limited feedback multiuser MIMO uplink," *Linear Algebra Its Applications*, vol. 396, pp. 373–384, February 2005.

[29] D. J. Love and R. W. Heath, "Limited feedback unitary precoding for spatial multiplexing systems," *IEEE Transactions on Information Theory*, vol. 51, no. 8, pp. 2967–2976, August 2005.

[30] T. Pande, D. J. Love, and J. V. Krogmeier, "Reduced feedback MIMO-OFDM precoding and antenna selection," *IEEE Transactions on Signal Processing*, vol. 55, no. 5, pp. 2284–2293, May 2007.

[31] D. J. Love, R. W. Heath, V. K. N. Lau, D. Gesbert, B. D. Rao, and M. Andrews, "An overview of limited feedback in wireless communication systems," *IEEE Transactions on Information Theory*, vol. 26, no. 8, pp. 1341–1365, October 2008.

[32] Z. Pan, K. K. Wong, and T. S. Ng, "Generalized multiuser orthogonal space-division multiplexing," *IEEE Transactions on Wireless Communications*, vol. 3, no. 6, pp. 1969–1973, November 2004.

[33] Q. H. Spencer, A. L. Swindlehurst, and M. Haardt, "Zero-forcing methods for downlink spatial multiplexing in multiuser MIMO channels," *IEEE Transactions on Signal Processing*, vol. 52, no. 2, pp. 461–471, February 2004.

[34] N. Ravindran and N. Jinal, "MIMO broadcast channels with block diagonalization and finite rate feedback," *IEEE International Conference*

on Acoustics, Speech and Signal Processing, ICASSP, vol. 3, pp. 13–16, April 2007, Honolulu, HI.

[35] N. Ravindran and N. Jinal, "Limited feedback-based block diagonalization for the MIMO broadcast channel," *IEEE Journal on Selected Areas in Communications*, vol. 26, no. 8, pp. 1473–1482, Oct. 2008.

[36] K. K. Wong, R. D. Murch, and K. B. Letaief, "A joint-channel diagonalization for multiuser MIMO antenna systems," *IEEE Transactions on Wireless Communications*, vol. 2, no. 4, pp. 773–786, July 2003.

[37] A. Wiesel, Y. Eldar, and S. Shamai, "Optimal generalized inverses for zero forcing precoding," *41st IEEE Annual Conference on Information Sciences and Systems*, no. Baltimore, MD, pp. 130–134, March 2007.

[38] M. Bengtsson and B. Ottersten, "Handbook of antennas in wireless communications. optimal and suboptimal transmit beamforming," *Boca Raton FL:CRC*, August 2001.

[39] J. Lofberg, "Yalmip: A toolbox for modelling and optimization in MATLAB," *Proceedings of IEEE International Symposium on Computer Aided Control Systems Design*, pp. 284–289, Tapei, September 2004.

[40] S. Boyd and L. Vandenberghe, "Convex optimisation," *Cambridge M.A.: Cambridge University Press*, 2004.

[41] J. Sturn, "Using seDuMi 1.02, a MATLAB toolbox for optimization over symmetric cones," *Optimization Methods and Software. Special issue on Interior Point Methods (CD supplement with software)*, vol. 11-12, pp. 625–653, 1999.

[42] H. Boche and M. Schubert, "Solution of the SINR downlink beamforming problem," *Proceedings of 36th Conference on Information Sciences and Systems (CISS) Princeton USA*, March 2002.

[43] H. Boche and M. Schubert, "A general duality theory for uplink and downlink beamforming," *Proceedings of 56th IEEE Semiannual Vehicular Technology Conference, Fall Vancouver Canada*, vol. 1, pp. 87–91, September 2002.

[44] H. Boche and M. Schubert, "A unifying theory for uplink and downlink multiuser beamforming," *Proceedings of IEEE International Zurich Seminar (IZS)*, pp. 271–276, February 2002.

[45] J. Palomar and M. Lagunas, "Joint transmit-receive space-time equalisation in spatially correlated MIMO channels: a beamforming approach," *IEEE Journal on Selected Areas on Communications*, vol. 21, no. 5, pp. 730–743, June 2003.

[46] J. Palomar, M. Lagunas, and J. Cioffi, "Optimum linear joint transmit-receive processing for MIMO channels with QoS constraints," *IEEE Transactions on Signal Processing*, vol. 52, no. 5, pp. 1179–1197, May 2004.

[47] J. Palomar, J. Cioffi, and M. Lagunas, "Joint Tx-Rx beamforming design for multicarrier MIMO channels: a unified framework for convex optimisation," *IEEE Transactions on Signal Processing*, vol. 51, no. 9, pp. 2381–2401, September 2003.

[48] A. Goldsmith, *Wireless Communications*. Gambridge University Press, 2005.

[49] G. Strang, *Linear Algebra and its Applications*. Fourth Edition: Thomson Brooks/Cole, 2005.

[50] G. Kramer, M. Gastpar, and M. Gupta, "Cooperative strategies and capacity theorems for relay networks," *IEEE Transactions on Information Theory*, vol. 51, no. 9, pp. 3037–3063, September 2005.

[51] J. N. Laneman, G. W. Wornell, and D. N. C. Tse, "An efficient protocol for realizing cooperative diversity in wireless networks," *Proceedings IEEE International Symposium on Information Theory, Waschington, DC.*, p. 294, June 2001.

[52] 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," *IEEE Transactions on Signal Processing*, vol. 56, no. 9, pp. 4306–4316, September 2008.

[53] A. Abdelkader, S. Shahbazpanahi, and A. Gershman, "Joint subcarrier power loading and distributed beamforming in OFDM-based asynchronous relay networks," *3rd IEEE International Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP)*, pp. 105–108, 13-16 December 2009.

[54] S. Fazeli-Dehkordy, S. Shahbazpanahi, and S. Gazor, "Multiple peer-to-peer communications using a network of relays," *IEEE Transactions on Signal Processing*, vol. 57, pp. 3053–3062, August 2009.

[55] H. Chen, A. B. Gershman, and S. Shahbazpanahi, "Distributed peer-to-peer beamforming for multiuser relay networks," *Proceedings of International Conference on Acoustics, Speech and Signal Processing (ICASSP'09), Tapei, Taiwan*, pp. 2265–2268, April 2009.

[56] Y. Rong, X. Tang, and Y. Hua, "A unified framework for optimizing linear nonregenerative multicarrier MIMO relay communication systems," *IEEE Transactions on Signal Processing*, vol. 57, no. 12, pp. 4837–4851, December 2009.

[57] X. Tang and Y. Hua, "Optimal design of non-regenerative mimo wireless relays," *IEEE Transactions on Wireless Communications*, vol. 6, no. 4, pp. 1398–1407, April 2007.

[58] O. Munoz-Medina, J. Vidal, and A. Augustin, "Linear transceiver design in non-regenerative relays with channel state information," *IEEE Transactions on Signal Processing*, vol. 55, no. 6, pp. 2593–2604, June 2007.

[59] D. P. Palomar and Y. Jiang, "MIMO transceiver design via majorization theory," *Now Publishers, Foundations and Trends in Communications and Information Theory*, vol. 3, no. 4-5, pp. 331–551, 2006.

[60] W. Zhang, U. Mitra, and M. Chiang, "Optimization of amplify-and-forward multicarrier two-hop transmission," *IEEE Transactions on Communications*, vol. 59, no. 5, pp. 1434–1445, May 2011.

[61] A. Behbahani, R. Merched, and A. Eltawil, "Optimizations of a MIMO relay network," *IEEE Transactions on Signal Processing*, vol. 56, no. 10, pp. 5602–5073, October 2008.

[62] W. Guan, H. Luo, and W. Chen, "Joint MMSE transceiver design in non-regenerative MIMO relay systems," *IEEE Communications Letters*, vol. 12, no. 7, pp. 517–519, July 2008.

[63] A. Toding, M. Khandaker, and Y. Rong, "Optimal joint source and relay beamforming for parallel MIMO relay networks," *Proceedings of 6th International Conference on Wireless Communications, Networking and Mobile Computing (WiCOM)*, pp. 1–4, September 2010.

[64] N. Laneman, "Cooperative diversity in wireless networks: Algorithms and architectures," *Ph.D dissertation, MIT, Cambridge, MA*, August 2002.

[65] A. Sendonaris, E. Erkip, and B. Aazhang, "User cooperation diversity-part I: System description," *IEEE Transactions on Communications*, vol. 51, pp. 1939–1948, November 2003.

[66] A. Stefanov and E. Erkip, "Cooperative coding for wireless networks,"

IEEE Transactions on Communications, vol. 52, no. 9, pp. 1470–1476, September 2004.

[67] B. Zhao and M. C. Valenti, "Distributed turbo coded diversity for relay channel," *Electronics Letters*, vol. 39, no. 10, pp. 786–787, May 2003.

[68] Z. Zhang and T. Duman, "Capacity approaching turbo coding for half duplex relaying," *IEEE Transactions on Communications*, vol. 55, no. 10, pp. 1895–1906, October 2007.

[69] A. Chakrabarti, A. de Baynast, A. Sabharwal, and B. Aazhang, "Low density parity check codes for the relay channel," *IEEE Journal on Selected Areas in Communications*, vol. 25, no. 2, pp. 280–291, February 2007.

[70] P. Razaghi and W. Yu, "Bilayer ldpc codes for the relay channel," *Proceedings of IEEE International Conference on Communications (ICC)*, pp. 1574–1579, June 2006.

[71] J. Hu and T. Duman, "Low density parity check codes over wireless relay channels," *IEEE Transactions on Wireless Communications*, vol. 6, no. 9, pp. 3384–3394, September 2007.

[72] Y. R. Tsai and L. C. Lin, "Optimal power allocation for decode-and-forward cooperative diversity under an outage performance constraint," *IEEE Communications Letters*, vol. 14, no. 10, pp. 945–947, October 2010.

[73] M. Chen, S. Serbetli, and A. Yener, "Distributed power allocation strategies for parallel relay networks," *IEEE Transactions on Wireless Communications*, vol. 7, no. 2, pp. 552–561, October 2010.

[74] K. Vardhe, D. Reynolds, and B. Woerner, "Joint power allocation and relay selection for multiuser cooperative communication," *IEEE Transactions on Wireless Communications*, vol. 9, no. 4, pp. 1255–1260, April 2010.

[75] A. Lau and S. Cui, "Joint power minimization in wireless relay channels," *IEEE Transactions on Wireless Communications*, vol. 6, no. 8, pp. 2820–2824, August 2007.

[76] M. Kaneko and P. P. K. Hayashi, "Throughput-guaranteed resource-allocation algorithms for relay-aided cellular ofdma system," *IEEE Trans. Commun.*, vol. 58, no. 4, pp. 1951–1964, May 2009.

[77] D. Ng and R. Schober, "Resource allocation and scheduling in multicell ofdma systems with decode-and-forward relaying," *IEEE Transactions on Wireless Communications*, vol. 10, no. 7, pp. 2246–2258, July 2011.

[78] Y. Fan, A. Adinoyi, J. S. Thompson, H. Yanikomeroglu, and H. V. Poor, "A simple distributed antenna processing scheme for cooperative diversity," *IEEE Transactions on Communications*, vol. 57, no. 3, pp. 626–629, 2009.

[79] C. E. Shannon, "Two-way communication channels," *Proceedings of the 4th Berkeley Symposium on Mathematical Statistics and Probability*, vol. 1, pp. 611–644, 1961.

[80] R. Vaze and R. W. Heath, "Capacity scaling for mimo two-way relaying," *IEEE International Symposium on Information Theory*, pp. 1451–1455, 24-29 June 2007.

[81] Y. Han, S. H. Ting, C. K. Ho, and W. H. Chin, "High rate two-way amplify-and-forward half-duplex relaying with ostbc," *IEEE Vehicular Technology Conference*, pp. 2426–2430, 11-14 May 2008.

[82] Y. Wu, P. A. Chou, and S. Y. Kung, "Information exchange in wireless networks with network coding and physical layer broadcast," *Proceedings of the 39th Annual Conference on Information Sciences and Systems (CISS'05)*, March 2005.

[83] P. Larsson, N. Johansson, and K. E. Sunell, "Coded bi-directional relaying," *IEEE 63rd Vehicular Technology Conference VTC 2006-Spring*, vol. 2, pp. 851–855, 7-10 May 2006.

[84] I. Hammerstrom, M. Kuhn, C. Esli, J. Zhao, A. Wittneben, and G. Bauch, "Mimo two-way relaying with transmit csi at the relay," *IEEE 8th Workshop on Signal Processing Advances in Wireless Communications SPAW 2007*, pp. 1–5, 17-20 June 2007.

[85] R. Zhang, Y.-C. Liang, C. C. Chai, and S. Cui, "Optimal beamforming for two-way multi-antenna relay channel with analogue network coding," *IEEE International Journal on Selected Areas in Communications*, vol. 27, no. 5, pp. 699–712, June 2009.

[86] C. Yuen, W. H. Chin, Y. L. Guan, W. Chen, and T. Tee, "Bi-directional multi-antenna relay communications with wireless network coding," *IEEE Vehicular Technology Conference, VTC Spring 2008.*, pp. 1385–1388, May 2008.

[87] S. J. Kim, P. Mitran, C. John, R. Ghanadan, and V. Tarokh, "Coded bi-directional relaying in combat scenarios," *IEEE Military Communications Conference, MILCOM 2007*, pp. 1–7, October 2007.

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

[89] C. K. Ho, R. Zhang, and Y.-C. Liang, "Two-way relaying over ofdm: Optimized tone permutation and power allocation," *IEEE International Conference on Communications*, pp. 3908–3912, May 2008.

[90] K. Jitvanichphaibool, R. Zhang, and Y.-C. Liang, "Optimal resource allocation for two-way relay-assisted ofdma," *IEEE Global Telecommunications Conference, GLOBECOM 2008*, pp. 1–5, November-December 2008.

[91] I. Hammerstrom, M. Kuhn, C. Esli, J. Zhao, A. Wittneben, and G. Bauch, "Mimo two-way relaying with transmit csi at the relay," *IEEE 8th Workshop on Signal Processing Advances in Wireless Communications SPAW 2007*, pp. 1–5, 17-20 June 2007.

[92] T. Cui, F. Gao, T. Ho, and A. Nallanathan, "Distributed space-time coding for two-way wireless relay networks," *IEEE International Conference on Communications ICC '08*, pp. 3888–3892, 19-23 May 2008.

[93] R. F. Wyrembelski, T. J. Oechtering, I. Bjelakovic, C. Schnurr, and H. Boche, "Capacity of gaussian mimo bidirectional broadcast channels," *IEEE International Symposium on Information Theory, ISIT 2008*, pp. 584–588, 6-11 July 2008.

[94] T. Unger and A. Klein, "Linear transceive filters for relay stations with multiple antennas in the two-way relay channel," *Mobile and Wireless Communications Summit IST*, pp. 1–5, 1-5 July 2007.

[95] E. A. Jorswieck and H. Boche, "Strategies and impact of correlation in multiantenna systems with different types of channel state information," *IEEE Transactions on Sigal Processing*, vol. 52, no. 12, December 2004.

[96] M. T. Ivrlac, W. Utschick, and J. A. Nossek, "Fading correlations in wireless mimo communication systems," *IEEE Journal on Selected Areas in Communications*, vol. 21, no. 5, pp. 819–828, June 2003.

[97] D. shan Shiu, G. J. Foschini, M. J. Gans, and J. M. Kahn, "Fading correlation and its effect on the capacity of multielement antenna systems," *IEEE Transactions on Communications*, vol. 48, no. 3, pp. 502–513, March 2000.

[98] S. Catreux, P. F. Driessen, and L. J. Greenstein, "Simulation results for an interference-limited multiple-input multiple-output cellular system," *IEEE Communications Letters*, vol. 4, no. 11, pp. 334–336, November 2000.

[99] D. Gesbert, M. Shafi, D. shan Shiu, P. J. Smith, and A. Naguib, "From theory to practice: an overview of mimo space-time coded wireless systems," *IEEE Journal on Selected Areas in Communications*, vol. 21, no. 3, pp. 281–302, April 2003.

[100] S. V. Hanly, "Information capacity of radio networks," *Ph.D dissertation, Gambridge University*, August 1993.

[101] S. V. Hanly, "Capacity and power control in spread spectrum macro-diversity radio networks," *IEEE Transactions on Communications*, vol. 44, no. 2, pp. 247–256, February 1996.

[102] S. V. Hanly and D. N. Tse, "Resource pooling and effective bandwidths in cdma networks with multiuser receivers and spatial diversity," *IEEE Transactions on Information Theory*, vol. 47, no. 4, pp. 1328–1351, March 2001.

[103] S. V. Hanly and P. A. Whiting, "Information-theoretic capacity of multi-receiver networks," *Telecommunication Systems*, vol. 1, no. 1, pp. 1–42, March 1993.

[104] A. D. Wyner, "Shannon-theoretic approach to a gaussian cellular multiple-access channel," *IEEE Transactions on Information Theory*, vol. 40, no. 6, pp. 1713–1727, November 1994.

[105] D. Gesbert, M. Kountouris, R. W. Heath, C.-B. Chae, and T. Salzer, "Shifting the mimo paradigm," *IEEE Signal Processing Magazine*, vol. 24, no. 5, pp. 36–46, 2007.

[106] H. Zhang and H. Dai, "Cochannel interference mitigation and cooperative processing in downlink multicell multiuser MIMO networks," *EURASIP Journal on Wireless Communications and Networking*, December 2004.

[107] S. Jing, D. N. C. Tse, J. B. Soriaga, J. Hou, J. Smee, and R. Padovani, "Multicell downlink capacity with coordinated processing," *Eurasip Journal on Wireless Communications and Networking*, January 2008.

[108] G. J. Foschini, K. Karakayali, and R. Valenzuela, "Coordinating multiple antenna cellular networks to achieve enormous spectral efficiency," *IEE Proceedings on Communications*, August 2006.

[109] W. Choi and J. G. Andrews, "Downlink performance and capacity of distributed antenna systems in a multicell environment," *IEEE Transactions on Wireless Communications*, January 2007.

[110] H. Zhang, H. Dai, and Q. Zou, "Base station cooperation for multiuser MIMO: joint transmission and bs selection," *Proceedings of the Conference on Information Sciences and Systems (CISS), Princeton, NJ*, March 2004.

[111] F. Boccardi and H. Huang, "Limited downlink network coordination in cellular networks," *IEEE 18th International Symposium on Personal, Indoor and Mobile Radio Communications, Athens*, pp. 1–5, September 2007.

[112] S. Venkatesan, "Coordinating base stations for greater uplink spectral efficiency in a cellular network," *IEEE 18th International Symposium on Personal, Indoor and Mobile Radio Communications*, pp. 3–7, September 2007.

[113] J. Zhang, R. Chen, J. G. Andrews, and R. W. Heath, "Coordinated multi-cell MIMO systems with cellular block diagonalization," *Conference Record of the Forty-First Asilomar Conference on Signals, Systems and Computers, 2007*, pp. 1669–1673, November 2007.

[114] C. Botella, G. Pinero, A. Gonzalez, and M. D. Diego, "Coordination in a multi-cell multi-antenna multi-user w-cdma system: A beamforming approach," *IEEE Transactions on Wireless Communications*, vol. 7, no. 11, pp. 4479–448, November 2008.

[115] H. Huang and M. Trivellato, "Performance of multiuser mimo and network coordination in downlink cellular networks," *6th International Symposium on odeling and Optimization in Mobile, Ad Hoc, and Wireless Networks and Workshops, 2008*, pp. 85–90, April 2008.

[116] R. Veronesi, V. Tralli, J. Zander, and M. Zorzi, "Distributed dynamic resource allocation for multicell sdma packet access net," *IEEE Transactions on Wireless Communications*, vol. 5, no. 10, pp. 2772–2783, October 2006.

[117] M. Vemula, D. Avidor, J. Ling, and C. Papadias, "Inter-cell coordination, opportunistic beamforming and scheduling," *IEEE International Conference on Communications, 2006*, vol. 12, pp. 5319–5324, June 2006.

[118] H. Dahrouj and W. Yu, "Coordinated beamforming for the multi-cell multi-antenna wireless system," *IEEE Transactions on Wireless Communications*, vol. 9, no. 5, pp. 1748–1759, May 2010.

[119] R. Zakhour, Z. K. M. Ho, and D. Gesbert, "Distributed beamforming coordination in multicell mimo channels," *IEEE 69th Vehicular Technology Conference, 2009*, pp. 1–5, April 2009.

[120] H. Huh, H. C. Papadopoulos, and G. Caire, "Multiuser MIMO transmitter optimization for inter-cell interference mitigation," *arXiv:0909.1344v1*.

[121] R. Zakhour and S. V. Hanly, "Base station cooperation on the downlink: Large system analysis," *IEEE Transactions on Information Theory*, vol. 58, no. 4, pp. 2079–2106, April 2012.

[122] F. Rashid-Farrokhi, K. Liu, and L. Tassiulas, "Transmit beamforming and power control for cellular wireless systems," *IEEE Journal on Selected Areas in Communications*, vol. 16, no. 8, pp. 1437–1450, October 1998.

[123] M. Bengtsson and B. Ottersten, "Optimal downlink beamforming using semidefinite optimization," *In Proceedings of 37th Annual Allerton Confer-*

ence on Communication, Control, and Computing, pp. 987–996, September 1999.

[124] A. Wiesel, Y. C. Eldar, and S. Shamai, "Linear precoding via conic optimization for fixed MIMO receivers," *IEEE Transactions on Signal Processing*, vol. 54, pp. 161–176, 2006.

[125] M. Schubert and H. Boche, "Iterative multiuser uplink and downlink beamforming under sinr constraints," *IEEE Transactions on Signal Processing*, vol. 53, no. 7, pp. 2324–2334, July 2005.

[126] E. Visotsky and U. Madhow, "Optimum beamforming using transmit antenna arrays," *IEEE 49th Vehicular Technology Conference*, vol. 1, pp. 851–856, July 1999.

[127] W. Yu and T. Lan, "Transmitter optimization for the multi-antenna downlink with per-antenna power constraints," *IEEE Transactions on Signal Processing*, vol. 55, no. 6, pp. 2646–2660, June 2007.

[128] W. Yu and T. Lan, "Transmitter optimization for the multi-antenna downlink with per-antenna power constraints," *IEEE Transactions on Signal Processing*, June 2007.

[129] S. Bongyong, R. L. Cruz, and B. D. Rao, "Network duality for multiuser mimo beamforming networks and applications," *IEEE Transactions on Communications*, vol. 55, no. 3, pp. 618–630, March 2007.

[130] R. Krishna, K. Cumanan, Z. Xiong, and S. Lambotharan, "A novel cooperative relaying strategy for wireless networks with signal quantization," *IEEE Transactions on Vechicular Technology*, vol. 59, no. 1, pp. 458–489, January 2010.

[131] S. Fazeli-Dehkordy, S. Gazor, and S. Shahbazpanahi, "Distributed peer-

to-peer multiplexing using ad hoc relay networks," pp. 2373–2376, April 2008.

[132] Y. Ying and H. Jafarkhani, "Network beamforming using relays with perfect channel information," *Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing (ICASSP '07), Honoloulou, HI*, vol. 3, pp. 473–476, Apr. 2007.

[133] A. El-Keyi and B. Champagne, "Cooperative MIMO beamforming for multiuser relay networks," pp. 2749–2752, April 2008.

[134] J. Joung and A. H. Sayed, "Multiuser two-way amplify-and-forward relay processing and power control methods for beamforming systems," *IEEE Transactions on Signal Processing*, vol. 58, pp. 1833 – 1846, March 2010.

[135] R. Krishna, Z. Xiong, and S. Lambotharan, "Cooperative relays for an underlay cognitive radio network," *International Conference on Wireless Communications and Signal Processing*, pp. 1–4, November 2009.

[136] N. Sidiropoulos, T. Davidson, and Z. Q. Luo, "Transmit beamforming for physical-layer multicasting," *IEEE Transactions on Signal Processing*, vol. 54, no. 6, pp. 2239 – 2251, June 2006.

[137] M. Chiang, "Geometric programming for communication systems," *now Publishers Inc*, 2005.

[138] S. Boyd and L. Vandenberghe, "Convex optimization," *Gambridge UK: Gambridge University Press*, 2004.

[139] P. Tseng, "Further results on approximating nonconvex quadratic optimization by semidefinite programming relaxation," *SIAM Journal on Optimization*, vol. 14, no. 1, pp. 268– 283, July 2003.

[140] S. Zhang, "Quadratic maximization and semidefinite relaxation," *Mathematical Programming*, vol. 87, no. 3, pp. 453–465, 2000.

[141] S. Srinivasa and S. A. Jafar, "Cognitive radios for dynamic spectrum access - The throughput potential of cognitive radio: A theoretical perspective," *IEEE Communications Magazine*, vol. 45, no. 5, pp. 73–79, May 2007.

[142] L. Zhang, Y. C. Liang, and Y. Xin, "Joint beamforming and power allocation for multiple access channels in cognitive radio networks," *IEEE Journal on Selected Areas in Communications*, vol. 26, no. 1, pp. 38– 51, January 2008.

[143] A. Jovicic and P. Viswanath, "Cognitive radio: An information-theoretic perspective," *IEEE Transactions on Information Theory*, vol. 55, no. 9, pp. 3345– 3958, September 2009.

[144] K. Phan, S. Vorobyov, N. Sidiropoulos, and C. Tellambura, "Spectrum sharing in wireless networks via QoS-aware secondary multicast beamforming," *IEEE Transactions on Signal Processing*, vol. 57, pp. 2323–2335, June 2009.

[145] Z.-Q. Luo and W. Yu, "An introduction to convex optimization for communications and signal processing," *IEEE Journal on Selected Areas in Communications*, vol. 24, no. 8, pp. 1426 – 1438, August 2006.

[146] R. A. M. Khandaker and Y. Rong, "Interference mimo relay channel: Joint power control and transceiver-relay beamforming," *IEEE Transactions on Signal Processing*, vol. 60, pp. 6509 – 6518, 2012.

[147] S. Shuying, M. Schubert, and H. Boche, "Downlink MMSE transceiver optimization for multiuser MIMO systems: Duality and sum-MSE minimization," *IEEE Transactions on Signal Processing*, vol. 55, pp. 5436 – 5446, November 2007.

[148] R. Hunger, M. Joham, and W. Utschick, "On the MSE-duality of the broadcast channel and the multiple access channel," *IEEE Transactions on Signal Processing*, vol. 57, pp. 698 – 713, February 2009.

[149] A. Berman and R. Plemmons, "Nonnegative matrices in the mathematical sciences," *New York: Academic*, 1979.

[150] J. Kaleva, A. Tolli, G. Venkatraman, and M. Juntti, "Downlink precoder design for coordinated regenerative multi-user relaying," *IEEE Transactions on Signal Processing*, vol. 61, no. 5, pp. 1215–1229, 2013.

[151] G. F. M. Karakayali and R. Valenzuela, "Network coordination for spectrally efficient communications in cellular systems," *IEEE Wireless Communications Magazine*, vol. 13, no. 4, pp. 56–61, August 2006.

[152] M. Lobo, L. Vandenberghe, S. Boyd, and H. Lebret, "Applications of second-order cone programming," *Linear Algebra and Its Applications*, vol. 284, pp. 193–228, 1998.

[153] K. Cumanan, R. Krishna, L. Musavian, and S. Lambotharan, "Joint beamforming and user maximization techniques for cognitive radio networks based on branch and bound method," *IEEE Transactions on Wireless Communications*, vol. 9, pp. 3082–3092, 2010.

[154] S. Benedetto, D. Divsalar, G. Montorsi, and F. Pollara, "Parallel concatenated trellis coded modulation," *IEEE International Conference on Communications (ICC)*, vol. 2, pp. 974–978, 1996.

[155] H. Shiwen, H. Yongming, Y. Luxi, A. Nallanathan, and L. Pingxiang, "A multi-cell beamforming design by uplink-downlink max-min SINR duality," *IEEE Transactions on Wireless Communications*, vol. 11, pp. 2858 – 2867, 2012.

[156] L. Zhang, R. Zhang, Y. Liang, Y. Xin, and H. Poor, "On Gaussian MIMO BC-MAC duality with multiple transmit covariance constraints," *IEEE Transactions on Information Theory*, vol. 58, no. 4, pp. 2064–2078, April 2012.

[157] S. Boyd, L. Xiao, and A. Mutapcic, "Subgradient methods," October 2003.

[158] W. Yu and R. Lui, "Dual methods for nonconvex spectrum optimization of multicarrier systems," *IEEE Transactions on Communications*, vol. 54, no. 7, pp. 1310–1322, July 2006.

[159] Y. Rahulamathavan, K. Cumanan, and S. Lambotharan, "A mixed SINR-balancing and SINR-target-constraints-based beamformer design technique for spectrum-sharing networks," *IEEE Transactions on Vehicular Tecnhology*, vol. 60, no. 9, pp. 4403 – 4414, November 2011.

[160] C. D. Meyer, "Matrix analysis and applied linear algebra," *Philadelphia, PA: SIAM*, 2000.

[161] V. Y. Pan and Z. Q. Chen, "The complexity of the matrix eigenproblem," *in 31st Annual Symposium on Theory of Computing, ACM Press*, vol. 1999, New York.

[162] M. Bengtsson and B. Ottersten, "Optimal and suboptimal transmit beamforming," *Handbook of Antennas in Wireless Communications, L.C. Godara, Ed. Boca Raton, FL:CRC Press*, 2001.