# Multi-User MIMO Relay Transceiver Optimization: Robust, Energy-Efficient and Secure Design Approaches

*Jiaxin Yang*

Department of Electrical & Computer Engineering
McGill University
Montreal, Canada

July 2018

A thesis submitted to McGill University in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

# Abstract

Multiple-input multiple-output (MIMO)-aided wireless relaying can improve the quality of the communication links between the source and destination nodes, hence significantly increasing system throughput, especially in multi-user networks. Relaying strategies can mainly be classified as amplify-and-forward (AF) and decode-and-forward (DF). The AF relaying technique imposes lower signal processing complexity and latency; therefore, it is preferred in many operational applications. In this regard, transceiver design becomes crucial to fully leverage the benefits of multi-user MIMO relay systems. The primary objective of this thesis is to develop new transceiver design approaches for multi-user MIMO relay networks from the perspectives of robustness, energy efficiency and secrecy.

First, we propose new transceiver design approaches for a multi-user MIMO AF relay network. It is well known that the performance of wireless relaying is significantly deteriorated under realistic conditions such as imperfect channel state information (CSI) for radio links involved in the transmission. To address this issue, two popular CSI error models, namely, the statistical and norm-bounded models, are considered. Based on these models, a robust joint transceiver design framework relying on modern convex optimization theory is proposed. The resulting design algorithms lead to a relaying performance that is notably less sensitive to different types of CSI errors, as demonstrated by the simulation results.

Then, we address the energy efficient design of a multi-user cooperative relay network. Assuming a flexible centralized network structure where relays can be adaptively activated/deactivated, we formulate the problem as a quality-of-service (QoS)-based network energy minimization problem that facilitates joint relay selection and transceiver optimization. An iterative solution based on re-weighted $l_1$-norm minimization along with a block-coordinate descent (BCD)-type algorithm is proposed and its convergence properties investigated. The new algorithm is shown to provide a significantly lower energy consumption of the relay network than that required by a conventional relaying scheme.

Finally, we propose a secure transceiver design approach for an MIMO relay network in the presence of multiple eavesdroppers. Under a realistic assumption of imperfect knowledge of the eavesdropper channels, we formulate the relay transceiver design as a signal-to-interference-plus-noise ratio (SINR) maximization subject to robust secrecy constraints. To solve the resulting non-convex problem, a penalized difference-of-convex (DC) algorithm is developed and its properties analyzed. Results show that the proposed algorithm can improve the secrecy of the relay-aided transmission at the physical layer.

## Sommaire

Le relayage sans fil à entrées multiples et sorties multiples (MIMO) peut améliorer la qualité des liaisons de communication entre les nœuds source et destination, augmentant ainsi considérablement le débit de transmission, en particulier dans les réseaux multi-utilisateurs. Les stratégies de relayage peuvent principalement être classifiées comme amplification-et-transfert (AF) et décodage-et-transfert (DF). La technique de relais AF impose une complexité de traitement du signal et une latence plus faibles; par conséquent, elle est préférée dans de nombreuses applications pratiques. À cet égard, la conception de l'émetteur-récepteur devient cruciale pour tirer pleinement parti des avantages des systèmes de relais MIMO multi-utilisateurs. L'objectif principal de cette thèse est de développer de nouvelles approches de conception d'émetteurs-récepteurs pour les réseaux de relais MIMO multi-utilisateurs du point de vue de la robustesse, de l'efficacité énergétique et de la confidentialité.

Nous proposons tout d'abord nouvelles approches de conception d'émetteurs-récepteurs pour un réseau de relais MIMO AF multi-utilisateurs. Il est bien connu que les performances des relais sans fil sont considérablement détériorées dans des conditions réalistes d'utilisation, par exemple lorsque les informations d'état de canal (CSI) pour les liaisons radio impliquées dans la transmission sont impreécises. Pour résoudre ce problème, deux modèles d'erreur CSI, sont pris en compte, à savoir les modèles statistique et normés. Sur la base de ces modèles, un cadre de conception d'émetteur-récepteur robuste reposant sur la théorie moderne de l'optimisation convexe est proposé. Les algorithmes de conception qui en résultent conduisent à une performance de relais notablement moins sensible aux différents types d'erreurs CSI, comme le démontrent nos résultats de simulation.

Nous abordons ensuite la conception d'un réseau de relais coopératif multi-utilisateurs utilisateurs du point de vue de l'économie de l'énergie. En supposant une structure de réseau centralisée souple où les relais peuvent être activés/désactivés de manière adaptative, nous formulons le problème de conception comme celui de la comme un problème de minimisation d'énergie de réseau avec contrainte sur la qualité de service (QoS) permettant ainsi la sélection conjointe des relais et l'optimisation de l'émetteur-récepteur. Une solution itérative basée sur la minimisation de la norme $l_1$ repondérée et l' algorithme de type descente de coordonnées par blocs (BCD) est proposée et ses propriétés de convergence étudiées. Les résultats de simulations montrent que le nouvel algorithm peut fournir une

consommation d'énergie significativement plus faible du réseau de relais que celle requise par un schéma de relayage conventionnel.

Enfin, nous proposons une approche de conception d'émetteur-récepteur sécurisé pour un réseau de relais MIMO en présence de multiples écouteurs clandestins. Sous l'hypothèse réaliste d'une connaissance imparfaite des canaux de ces écouteurs, nous formulons la conception de l'émetteur-récepteur relais comme une maximisation du rapport signal-à-interférence-plus-bruit (SINR) soumise à des contraintes de confidentialité. Pour résoudre le problème non convexe qui en résulte, un algorithme de différence de convexité (DC) pénalisé est développé et ses propriétés analysées. Les résultats montrent que l'algorithme proposé peut améliorer la confidentialité de la transmission assistée par relais au niveau de la couche physique.

## Acknowledgments

My deepest gratitude goes to my supervisor, Prof. Benoit Champagne for his continuous support, inspiration and motivation for high quality research throughout my Ph.D. studies. Without his guidance and encouragement, this thesis would have never been possible. I would also like to express my sincere appreciation to Prof. Lajos Hanzo from University of Southampton, U.K., for hosting my research visit to his laboratory as well as his generous support and invaluable technical advice that significantly improved the quality of my work during the past three years. I am also very thankful to Prof. Yunlong Cai from Zhejiang University, Prof. Yulong Zou from Nanjing University of Posts and Telecommunications, and Prof. Qiang Li from University of Electronic Science and Technology of China for their constructive comments during our research collaborations. I would also like to thank Dr. Benoit Pelletier for his technical guidance during my internship at InterDigital.

I would like to thank the other members of my Ph.D. supervisory committee, Profs. Ioannis Psaromiligkos and Bruce Shepherd, who have regularly reviewed the progress of my work and provided valuable feedbacks throughout my Ph.D. program. I also would like to thank Prof. Yue Rong from Curtin University of Technology, Australia, for serving as the external examiner and for spending time reviewing my thesis.

I have thoroughly enjoyed spending time with my fellow labmates (present and past) at the Telecommunication and Signal Processing Lab. Here, I wish to give special thanks to: Ali, Alireza, Chao, Djamel, Fang, Hanwook, Hao, Kangli, Lu, Reza, Ryan, Siamak, Tianze for their generous help and discussions.

I gratefully acknowledge the financial support provided by McGill University through the McGill Engineering Doctoral Award (MEDA), by the Natural Sciences and Engineering Research Council (NSERC) of Canada and InterDigital Canada through the NSERC Collaborative Research and Developments (CRD) grant, by the NSERC of Canada through the Collaborative Research and Training Experience (CREATE) Program, and by the Fonds Québécois de la Recherche sur la Nature et les Technologies (FQRNT) through the International Internship Program.

I am forever grateful to my parents for their unconditional love and support, which helped me overcome various difficulties despite the distance. Last but not least, my heartfelt love is devoted to my wife, Erica Yanwen Kuang, for her patience, continuous support and encouragement over the course of my Ph.D. studies. I hope that this accomplishment makes all of you proud.

# Contents

# List of Figures

# List of Tables

# List of Acronyms

| | |
|---|---|
| 3G | third generation |
| 4G | fourth generation |
| 5G | fifth generation |
| AF | amplify-and-forward |
| AN | artificial noise |
| BBU | baseband unit |
| BCD | block-coordinate descent |
| BER | bit error rate |
| BS | base station |
| CCCP | constrained concave-convex procedure |
| CDF | cumulative distribution function |
| CDI | channel direction information |
| CMI | channel magnitude information |
| C-RAN | cloud radio access network |
| CSI | channel state information |
| D2D | device-to-device |
| DC | difference-of-convex |
| DF | decode-and-forward |
| DFT | discrete Fourier transform |
| DL | downlink |
| ECSI | eavesdropper's channel state information |
| GLA-VQ | generalized Lloyd algorithm of vector quantization |
| HetNet | heterogeneous network |
| HSPA | high speed packet access |

| | |
|---|---|
| ICT | information and communication technology |
| i.i.d. | independently and identically distributed |
| KKT | Karush-Kuhn-Tucker |
| LMI | linear matrix inequality |
| LTE | long-term evolution |
| MCS | modulation and coding scheme |
| MFCQ | Mangasarian Fromovitz constraint qualification |
| MIMO | multiple-input multiple-output |
| MM | majorization minimization |
| MMSE | minimum mean square error |
| MRC | maximum ratio combining |
| MSE | mean square error |
| NR | new radio |
| PSD | positive semi-definite |
| QCLP | quadratically-constrained linear program |
| QCQP | quadratically-constrained quadratic program |
| QoS | quality of service |
| QPSK | quadrature phase-shift keying |
| RRH | remote radio head |
| RVQ | random vector quantization |
| SC | selection combining |
| SDP | semidefinite program |
| SDR | semidefinite relaxation |
| SINR | signal-to-interference-plus-noise ratio |
| SNR | signal-to-noise ratio |
| SOC | second-order cone |
| SOCP | second-order cone programming |
| SVD | singular-value decomposition |
| TPC | transmit precoder |
| UE | user equipment |
| UL | uplink |
| WCDMA | wideband code division multiple access |
| ZF | zero-forcing |

# Chapter 1

# Introduction

A conventional wireless cellular network is characterized by the use of multiple non-overlapping cells covering a given geographical area such as a city. In each cell, a base station (BS) is deployed by the network operator to provide cellular coverage and services to all mobile subscribers within the cell. The use of multiple cells makes it possible to reduce the transmission power and interference and therefore increase the network capacity (i.e., supporting larger number of users). Examples of such cellular networks include third generation (3G) systems such as Wideband Code Division Multiple Access (WCDMA) and High Speed Packet Access (HSPA) and fourth generation (4G) Long-Term Evolution (LTE).

Over the past few decades, wireless communication technologies and services have witnessed an explosive growth. Due to the proliferation of mobile devices, and especially smart phones, there has been an increased demand for ubiquitous data access with enhanced data rate and quality-of-service (QoS) requirements. In particular, the popularity of advanced mobile multimedia applications (e.g., video streaming) has brought significant challenges to the design of next-generation wireless networks. In a typical point-to-point communication scenario where one wireless transmitter sends data to its receiver, the link performance is limited by and already quite close to the fundamental capacity limit. To achieve higher data rates, a relatively high signal-to-noise ratio (SNR) is therefore required from a pure link budget perspective. In current wireless cellular networks, a natural approach to achieve higher SNR would be to deploy a denser network infrastructure, i.e., to use smaller cells where user-to-BS distance is reduced. However, the ultra-dense deployment of BSs leads to significantly higher deployment and maintenance costs, which may also make it more

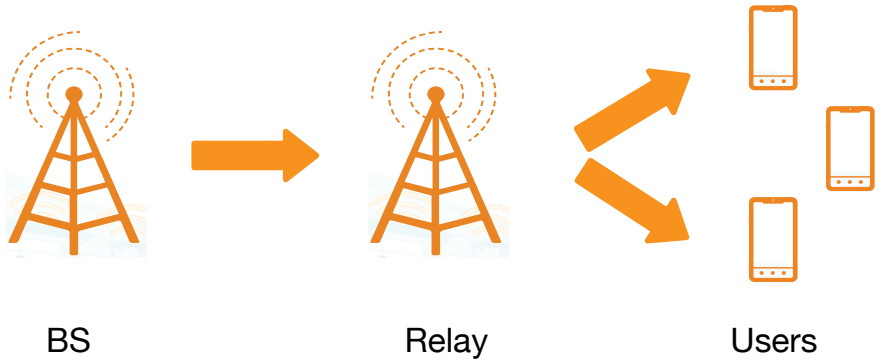



**Fig. 1.1** A typical wireless relaying scenario.

difficult to transit from existing to next-generation standards.

To address this issue, current LTE-Advanced and emerging standards such as fifth generation (5G) New Radio (NR) are embracing relaying technologies to extend the network coverage and improve the throughput in hot spots. Different from traditional or macro BS, the relay stations may be remote radio heads (RRHs) transmitting at low power levels. Other benefits of deploying relay stations include low implementation cost and flexible site acquisition. Below we briefly review the various strategies currently available for wireless relaying and then expose the main technical challenges brought about by their utilization. This sets the stage for the presentation of the thesis goals and contributions.

## 1.1 The Wireless Relaying Concept

Wireless relaying, where one or more intermediate nodes, called *relays*, assist in forwarding data from a source node to one or more destination nodes (as depicted in Fig. 1.1), has been proposed as a cost-effective solution to meet some of the more stringent demands of LTE-Adavned and 5G NR networks [1]. Indeed, wireless relaying offers many benefits, either by boosting the performance of traditional cellular networks, or extending their functionality. First, relay stations can be deployed to improve the QoS of users near the cell boundary (which suffer from higher path loss and interference), or to extend the network coverage by providing services to out-of-coverage users (as in rural areas). Compared to the use of additional macro BSs, using relays can save the high costs associated with site acquisition and backhaul deployment. In this way, uniform broadband experience to users almost everywhere within the network is provided in a cost-efficient manner. Second, simi-

lar to the concept of frequency reuses in cellular systems, relay stations can provide better spatial resolution. Hence, more users can be served simultaneously within the same frequency band, due to the combined effect of reduced transmit-receive distance (hence higher received signal strength) and smaller co-channel interference from neighboring small cells served by other relay stations, thereby contributing to a higher system throughput. Third, in Release 12 of LTE, device-to-device (D2D) communication has been proposed, which allows two mobile devices in proximity of each other to establish a direct local link as an underlay mode of operation and thus to bypass the BS [2]. The inherent property of wireless relaying makes it perfectly suitable for D2D scenarios. For instance, for D2D-based public safety applications, user equipments (UEs) with better link quality and longer battery life may serve as relays for destination UEs that are out of coverage of the network. Finally, in a multi-user scenario, when combined with multiple-input multiple-output (MIMO) techniques, wireless relaying is capable of simultaneously serving multiple users concurrently over the same physical channel by exploiting the so-called distributed spatial multiplexing, which also contributes to higher network throughput.

In general, wireless relaying can be classified based on the chosen relaying strategy, duplexing mode, transmission mode (i.e., one-way or two-way), number of equipped antennas and so on. The relaying strategy refers to how the source information is processed and forwarded to the intended destination by the relay. The two main categories of relaying strategies are amplify-and-forward (AF) (non-regenerative) and decode-and-forward (DF) (regenerative). In AF relaying, which is the simplest strategy, the relay amplifies the received signal from the source and then transmits the resultant to the destination without doing any decoding. Ideally, the signal amplification can be represented by a linear transformation: for single-antenna relays, the linear transformation amounts to multiplying the received signal by a scaling factor while for multi-antenna relays, it amounts to multiplying the received signal vector by a transformation matrix. In contrast, a relay adopting the DF strategy first needs to decode the binary information bits received from the source, and then re-encodes, modulates and finally transmits the resultant signal to the destination. Compared with AF, the DF strategy suffers higher signal processing complexity and longer delay. For these reasons, our focus in this thesis will be on the AF relaying strategy[1]. It is also worth noting that in the literature, other relaying strategies have been proposed, such

[1]In this thesis, unless otherwise noted, we use the term "relay transceiver design" to denote transceiver design for relays adopting AF strategy.

as compress-and-forward [3] and selective relaying [4]. However, such strategies fall outside the scope of the present thesis.

Wireless relays can operate in either full-duplex or half-duplex mode. In the half-duplex mode, orthogonal duplexing is employed to avoid any possible interference between the receive and transmit functions. In this regard, time duplexing is commonly used, where the relay receives the signal from the source during a first time slot (hop), and then retransmits an AF or DF processed signal during the second time slot, etc. In the full-duplex mode, the relay is allowed to receive and transmit simultaneously, which in general is difficult to implement since the transmitted signal from the relay can severely interfere with the co-located receiver [5]. Although the half-duplex constraint impacts negatively on the theoretical spectral efficiency, the latter can be compensated by the additional capacity gain. For these reasons, we focus on half-duplex relaying mode in this thesis.

Relays can also be classified based on their transmission modes, i.e., one-way or two-way communications. In one-way communications, the relay simply forwards the information from one end-user (the source) to another end-user (the destination) while in two-way communications, the two end-users simultaneously transmit to the relay in the first time slot, and the relay then broadcasts the processed signals back to the two end-users in the second slot [6]. Two-way relaying may improve the network throughput since only two transmission phases are needed to exchange the information between two end-users while for one-way relaying, four transmission phases are needed instead. However, technical challenges such as interference management (mutual interference and self-residual interference) and synchronization need to be carefully considered in the design and implementation of the network. In this thesis, we shall limit our attention to the one-way relaying, although some of the proposed solutions may be extended to the two-way scenario.

Finally, the introduction of MIMO techniques into the relaying framework, through the use of multiple antennas at the relays and possibly also at the sources and destinations, brings further performance benefits as compared to the single-antenna configuration. However, the design and implementation of the transceiver sub-systems for MIMO relaying becomes significantly more challenging. For DF relaying, transceiver design at the relay is relatively straightforward since it can be realized in two separate parts: use of an MIMO receiver in the first hop, and an MIMO transmitter in the second hop. Still, implementation of these MIMO receiver and transmitter increases complexity, cost and processing delay at the relays. In contrast, for AF relaying, the transceiver structures can now be repre-

sented by linear matrix transformation, i.e., both the receive and transmit functions are now incorporated into a single linear matrix transformation. Compared with single-antenna configurations, where the transceiver structure can be represented by a single multiplicative factor (or weight), the shift from scalar to matrix optimization significantly complicates the transceiver design problem. Note also that in AF relaying, not only the useful signals are amplified and forwarded, but also are the additive noise components at the relays. Therefore, a good transceiver design for AF relaying must prevent over-amplification of noise components while preserving the desired signal components. The transceiver design problem for MIMO AF relaying is exacerbated in more general multi-user scenarios involving multiple source-destination pairs or multiple cooperating relays. In these configurations, in order to achieve optimum performance, the transceivers of all sources, relays and destinations need to be jointly optimized. This type of problems, where the cross interference between multiple users needs to be carefully handled, is quite mathematically challenging, especially when considering additional power constraints and other system limitations, as further discussed in the next section.

Relaying is one of the proposed features for the current LTE standard. In Release 8 [7], AF relays, also referred to as repeaters, were adopted as a tool for handling coverage holes. Later, Release 10 introduced support for a DF relaying scheme [8]. Main use cases include extending the coverage to mountainous and sparsely populated regions, providing temporary coverage when there is an emergency, enhancing throughput in urban/indoor hot spots and providing services to mobile users in public vehicles to reduce excessive handover. Motivated by these use cases, the so-called *self-backhauling* was taken as the basis for the LTE relaying solution. In LTE, the terms *backhaul link* and *access link* are often used to refer to the BS-relay connection and the relay-terminal connection, respectively. Since the relay communicates both with the macro cell and terminals served by the relay, interference between the access and backhaul links must be avoided. Depending on the spectrum used for access and backhaul links, LTE relaying can be classified into *outband* and *inband* types. Outband relaying implies than the backhaul operates in a spectrum separate from that of the access link, using the same radio interface as the access link. Provided the frequency separation between the backhaul and access links is sufficiently large, interference between the two links can be avoided. Inband relaying implies than the backhaul and access links operate in the same spectrum and therefore, additional interference avoidance mechanisms are required. For more details on LTE relaying, including the overall architecture and

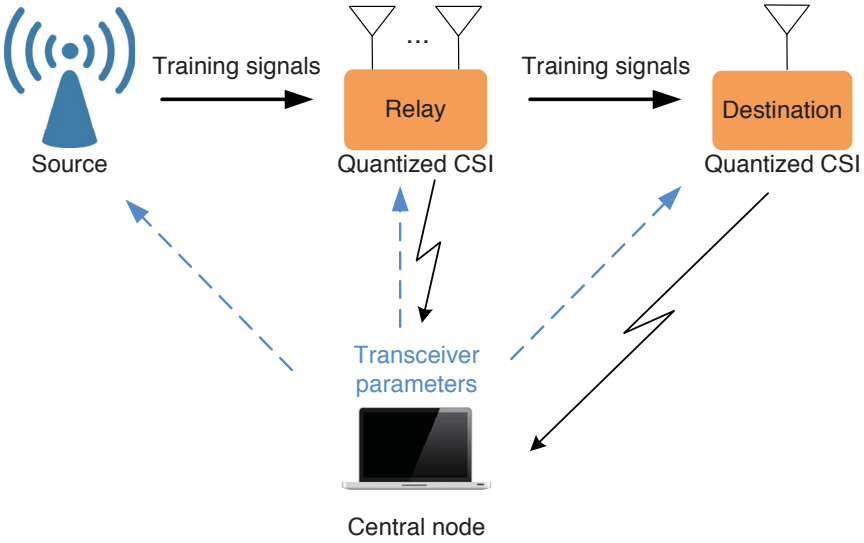



**Fig. 1.2** Illustration of imperfect CSI due to channel estimation and quantization errors.

backhaul design for inband relaying, interested readers are referred to [8].

## 1.2 Technical Challenges Posed by Wireless Relaying

Some prior contributions focus on the performance optimization of wireless cooperative relaying, for example, maximizing the communication performance of end-users, but without the consideration of some crucial practical constraints. These include the necessary acquisition of channel state information (CSI) and the resultant effect of possible CSI errors, the possibly high network energy overhead when multiple MIMO relays are involved in cooperative transmissions, and information security issues posed by malicious users who can eavesdrop the relay-assisted transmissions. Below, we further examine these challenges in the design of wireless MIMO relaying networks, which also largely motivate the new solutions and algorithms developed in this thesis.

It is well known that the efficacy of relay transceiver design largely relies on the availability and accuracy of the CSI for all radio links involved in the transmission, i.e., from source to relay as well as relay to destination. In practice, acquiring accurate CSI of these links is quite challenging due to processing complexity of the channel estimation (especially in the case of multi-antenna or multi-user MIMO relaying configurations) and the combined effects of various sources of estimation errors due to noise and radio interference,

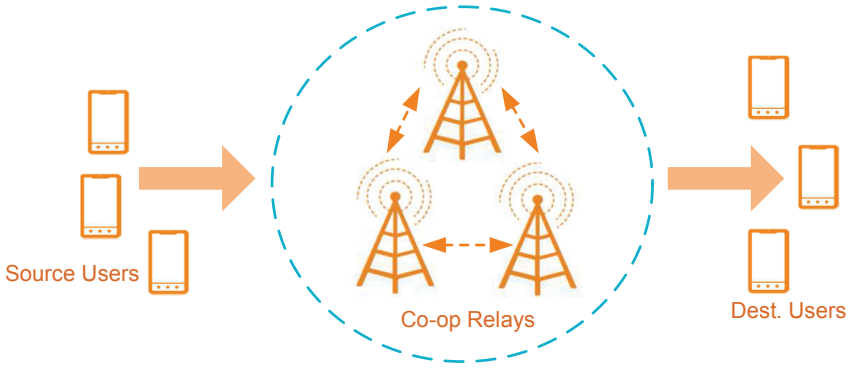



**Fig. 1.3** The concept of cooperative wireless relaying (the double-head arrows refer to the exchange of information between the cooperating relays via dedicated links).

the use of feedback with limited quantization, and in the case of mobile communications, the possible rapid changes in the channel condition that eventually render the available CSI obsolete. As an illustrative example, a typical scenario for collecting CSI of all the links in a relaying network is shown in Fig. 1.2. In the first step, the relay and destination nodes estimate their corresponding CSI based on the (known) transmitted training sequences. In the second step, the estimated CSI is further quantized using a limited number of bits and subsequently fed back to a central node via a dedicated control link. The central node then performs optimization of the transceivers of all the nodes based on the received CSI and finally forwards the optimized transceiver parameters to all the nodes as needed. In the literature, there exist various techniques for channel estimation and quantization. In general, a more accurate channel estimation scheme, possibly with higher signal processing complexity and a larger number of quantization levels, can provide better quality in CSI reporting, hence leading to a better performance for the relay-assisted transmission. Therefore, there is a tradeoff between the level of CSI errors and the resultant relaying performance. In view of this, robust relay transceiver design, which explicitly takes into consideration the effects of CSI errors, is an imperative yet challenging task.

Multi-cell cooperation has attracted increasing attention since it has been considered in current LTE-Advanced standard as a means to improve network capacity [9]. The concept of multi-cell cooperation has also been considered in the context of wireless relaying, where multiple relays, which can be BSs or access points serving different cells or RRHs, work cooperatively to forward the signals transmitted from multiple sources to their corresponding

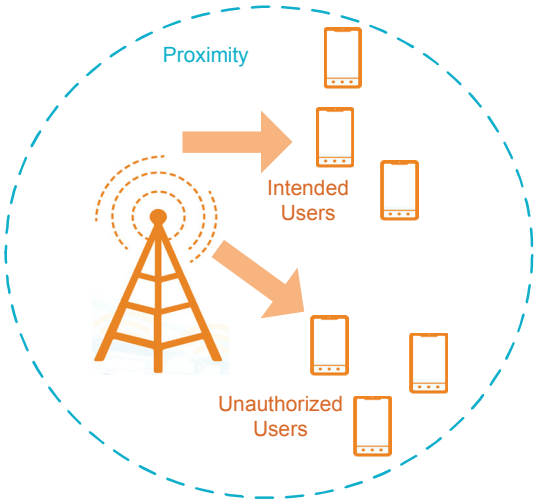



**Fig. 1.4** An illustrative example of the openness of the wireless medium.

destinations, as depicted in Fig. 1.3. In designing multi-relay networks, a widely-adopted approach is to maximize the overall performance of the end-users under relay power constraints. This approach may lead to a transmission strategy where all relays remain active, which deviates from the concept of energy efficiency. It is observed that information and communication technology (ICT) is sadly playing an important role in global greenhouse gas emissions. This is because the amount of energy consumed by ICT has significantly increased during the past decades due to the explosive growth in the mobile user population and service demand requirements. It is reported that the total energy consumed by cellular network infrastructures, wired communication networks and the Internet backbone currently takes up more than three percent of the worldwide electric energy consumption [10]. Against this background, designing energy efficient, or *green* communication networks is one of the main objectives in the development of future generation of wireless communication standards. For multi-relay networks, how to use the relays efficiently remains a critical problem, especially: how many relays are needed for transmission, which subset of relays shall remain active, and how their corresponding transceivers are optimized? It is therefore highly desirable to address these problems in a unified framework.

With the proliferation of smart phones storing more sensitive personal data ranging from social networking to online banking, wireless end-users have become vulnerable targets of malicious hackers owing to the broadcast nature of signal propagation, as illustrated in Fig. 1.4. According to a few reports on mobile cyber threats, the number of cyber at-

tacks to mobile users has been dramatically growing in the past few years [11]. Within this context, how to ensure information security is becoming a critical issue for wireless service providers. Although the classic bit-level encryption technique has been deemed to be most effective way of achieving this goal, a report by the Washington Post has drawn public attention to the potential security risks of wireless technologies, even when advanced encryption is used[2]. Against this background, physical layer security is emerging as a promising design alternative to complement classic encryption and to further enhance the security of wireless networks. Since Wyner opened this new avenue of security provision by introducing the notion of secrecy capacity [13], researchers have sought to enhance security for a wide range of communication channel models [14]. Recently, physical layer security has attracted increased interest, driven by new applications and network configurations such as cooperative wireless relaying. Although the diversity advantages gleaned from user cooperation have been recognized in the context of generic relay-assisted networks, ensuring secrecy in cooperative information relaying remains a key issue. Specifically, when additional intermediate nodes assist in forwarding the source messages to their destination, the information confidentiality may be more readily compromised, unless the relaying scheme is appropriately designed. In a number of works, the secrecy of relay-assisted communications has been investigated from physical layer perspective, showing that ingenious transceiver design can help improve the security level of relay-assisted communications [15]. Still much work remains to be done, and following this trend, it becomes important to investigate new signal processing techniques conceived for further improving wireless relaying security without compromising the QoS levels of the destination users.

## 1.3 Thesis Objective and Contributions

The primary objective of this thesis is to develop and investigate transceiver design approaches for applications in multi-user MIMO relay networks. To achieve this goal, we need to address the aforementioned research challenges by proposing new designs that can provide robustness to CSI errors, greater energy efficiency, and improved security against eavesdropping. In this regard, the main research contributions and findings of this thesis

[2]In [12], it is reported that two German researchers have demonstrated how to exploit the security flaws in the Signaling System 7 (SS7) to eavesdrop on all incoming and outgoing calls indefinitely from anywhere in the world. They have shown how to decode the messages by requesting each caller's carrier to release a temporary encryption key through the SS7.

are summarized as follows.

Chapter 3 focuses on linear transceiver design in a multi-user MIMO AF relay network. The objective of this chapter is to develop linear transceiver structures that maximizes the QoS at all end-users while making the solutions less sensitive to the errors incurred in the CSI reporting stage. We consider a scenario, where each source transmits multiple substreams to its corresponding destination with the assistance of multiple relays. Assuming realistic imperfect CSI of all the source-relay and relay-destination links, we propose a robust optimization framework for the joint design of the source transmit precoders (TPCs), relay AF matrices and destination receive filters. Specifically, two well-known CSI error models are considered, namely the *statistical* and *norm-bounded* error models. We first consider the design problem of minimizing the maximum per-stream mean square error (MSE) subject to the source and relay power constraints (*min-max* problem). Then, two different versions of this problem, which respectively take into account the statistical and norm-bounded CSI errors, are formulated. Under these two scenarios, algorithmic solutions having proven convergence and tractable complexity are proposed by resorting to the iterative block coordinate update approach along with matrix transformation and convex conic optimization techniques. Secondly, we consider the problem of minimizing the maximum per-relay power subject to the QoS constraints for each substream and the source power constraints (*QoS* problem). For this problem, an efficient initial feasibility search algorithm is proposed based on the relationship between the feasibility check and the min-max problems. We demonstrate through computer simulations that the proposed joint transceiver design algorithms can achieve an improved robustness against different types of CSI errors, when compared to non-robust approaches proposed earlier in the literature.

Having designed CSI error resilient relay transceivers that aim to improve the users' QoS, Chapter 4 addresses the design of a multi-user relaying network from an energy-efficient perspective. The objective of this chapter is to develop a relaying strategy that aims to minimize the total network energy consumption while still providing prescribed QoS levels to the destination users. We consider a scenario where all the relays are connected to a centralized processing node with the aid of high bandwidth, low latency fronthaul links (e.g., the so-called cloud radio access network (C-RAN) proposed for the next generation cellular standard). Exploiting the flexible centralized processing structure of the network, where relays can be adaptively activated/deactivated, we formulate the problem as a QoS-based network energy minimization problem that facilitates joint relay selection

and transceiver optimization. We propose an iterative solution based on the concept of the re-weighted $l_1$ norm along with a block-coordinate descent (BCD)-type algorithm. The active relays are then determined in a single attempt by thresholding a group sparsity pattern associated with the set of all relaying transceiver matrices. To circumvent a potentially undesirable condition, where the selected subset of relays fails to simultaneously satisfy all the destination users' prescribed QoS levels, we conceive a user admission control mechanism for overcoming the associated *infeasibility* problem. Simulation results demonstrate the explicit benefits of the proposed design approach, which results in a significantly lower energy consumption of the relaying network than that required by a conventional cooperative relaying scheme.

Chapter 5 then shifts the focus of relaying design from user performance and energy efficiency to communication security. Owing to the vulnerability of relay-assisted communications and its use cases such as D2D communications as mentioned earlier in this chapter, improving wireless security from a physical layer signal processing perspective is attracting increasing interest. Hereby we address the problem of secure transmission in a relay-assisted network, where a pair of legitimate users communicate with the aid of an MIMO relay in the presence of multiple eavesdroppers. Assuming imperfect knowledge of the eavesdroppers' channels, we jointly optimize the power of the source user, the relay AF transceiver matrix and the covariance of the artificial noise (AN) transmitted by the relay, in order to maximize the received signal-to-interference-plus-noise ratio (SINR) at the destination user, while imposing a set of *robust secrecy constraints*. A globally optimal solution based on a bi-level optimization framework is first proposed, but with high complexity. Hence, a low-complexity sub-optimal method relying on a newly proposed penalized difference-of-convex (DC) algorithmic framework is proposed. We show how this penalized DC framework can be invoked for solving the robust secure relaying problem with proven convergence. Our extensive simulation results show that both proposed solutions are capable of ensuring the secrecy of the relay-aided transmission and significantly improving the robustness to the eavesdroppers' channel uncertainties, as compared to the non-robust counterparts. It is also demonstrated that the advocated penalized DC-based method yields a performance close to the globally optimal solution.

## 1.4 Organization

The rest of the thesis is organized as follows. Chapter 2 provides a comprehensive literature survey of prior contributions on wireless MIMO relaying from various perspectives. Chapter 3 focuses on the design of robust multi-user MIMO relaying transceivers that can improve the QoSs of destination users in the presence of CSI errors. Chapter 4 investigates the design of an MIMO relay network from an energy-efficient perspective, where the aim is to minimize the total network energy consumption while providing predefined QoSs to the destination users. Chapter 5 considers joint MIMO relay transceiver design problem from a security perspective, seeking to maximize the QoS, measured by the SINR, at the destination user, subject to a set of robust secrecy constraints. Certain mathematical proofs and derivations are relegated to the Appendices.

The following notations are used throughout the thesis, unless otherwise noted. Boldface uppercase (lowercase) letters denote matrices (vectors), while regular font letters denote scalars; $\mathbb{C}$ denotes the set of complex numbers; the superscripts $(\cdot)^*$, $(\cdot)^T$, $(\cdot)^H$, and $(\cdot)^{-1}$ denote the conjugate, transpose, Hermitian transpose and inverse, respectively; $\|\cdot\|$ represents the Euclidean norm of a vector, while $\|\cdot\|_F$ denotes the Frobenius norm of a matrix; $\mathrm{Tr}(\cdot)$, $\mathtt{vec}(\cdot)$, and $\otimes$ stand for the matrix trace, vectorization and the Kronecker product, respectively; $\mathbb{C}^{N\times M}$ and $\mathbb{H}^M$ denotes the spaces of $(N \times M)$-element matrices having complex entries and $M \times M$ Hermitian matrices, respectively; $\mathrm{Re}\{\cdot\}$ denotes the real part of a complex number.

# Chapter 2

# A Survey of Transceiver Design for MIMO AF Relaying Networks

In this chapter, we provide a literature review of the state-of-the-art approaches available for the design of MIMO AF relaying networks[1]; these approaches serve as background material for the algorithm developments that appear in the subsequent chapters of the thesis. The presentation is organized as follows: In Section 2.1, we review the transceiver optimization techniques for one of the most fundamental relaying networks, i.e., the point-to-point MIMO AF relaying network. In Section 2.3, we extend the review to multi-user MIMO AF relaying systems. The issues posed by imperfect CSI in the design of relay transceivers are discussed in Section 2.4. Prior works addressing energy efficiency issues in designing wireless networks are reviewed in Section 2.5. Methods for improving relaying transmission secrecy from a physical layer perspective are discussed in Section 2.6. Finally, some concluding remarks are given in Section 2.7.

## 2.1 Point-to-Point AF Relaying Networks

The point-to-point MIMO AF relaying network, as shown in block diagram form in Fig. 2.1, is one of the most fundamental models studied in the MIMO relaying literature. In this model, a source node transmits its data to a corresponding destination node with the aid

[1]In practice, a "relaying network" is usually part of a larger wireless network, and as such, it might be referred to as a sub-network. Some authors also use the expression "relaying systems". In this thesis, for consistency, we adopt the former terminology.

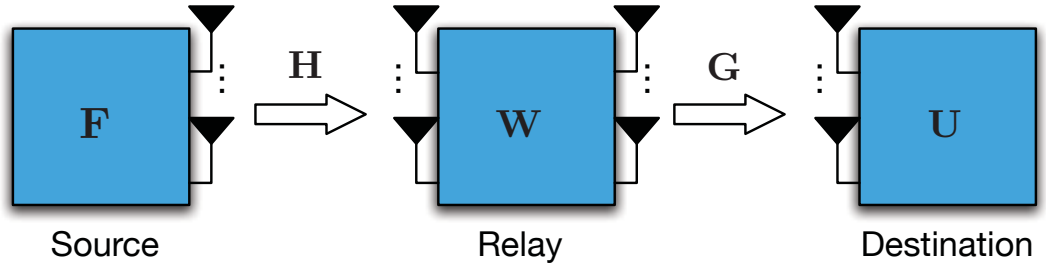



**Fig. 2.1** A point-to-point MIMO AF relaying network where **F**, **W** and **U** denotes the transceiver matrices at source, relay and destination, respectively.

of a relay node. All three nodes are equipped with multiple antennas in order to exploit additional *spatial degrees of freedom*, thus improving the relaying transmission reliability.

Suppose that the source, relay and destination employ $N_{\mathtt{S}}$, $N_{\mathtt{R}}$ and $N_{\mathtt{D}}$ antennas, respectively. The relay operates under the half-duplex protocol, where the data transmission is completed within two phases. In the first phase, the source transmits its signal to the relay, whereas in the second phase, the relay processes its received signal using an AF transceiver structure and forwards the resultant signal to the destination. As in most of the literature on MIMO AF relaying, the so-called narrow-band flat-fading radio propagation model is assumed throughout the thesis. In addition, no direct link is available between the source and destination due to a severe path loss.

Let $\mathbf{s} \in \mathbb{C}^{d\times 1}$ denote the vector of information symbols to be transmitted by the source. The entries of $\mathbf{s}$, denoted as $s_k$ for $k \in \{1, 2, \cdots, d\}$ and called substreams in this context, are modeled as random variables with zero mean and unit variance, i.e., $\mathbb{E}\{|s_k|^2\} = 1$. It is further assumed that the number of substreams satisfies $d \leq \min(N_{\mathtt{S}}, N_{\mathtt{R}}, N_{\mathtt{D}})$ to allow the decoding of the transmitted symbols at the destination. Following the application of a linear transmit precoder (TPC) matrix to $\mathbf{s}$, the signal transmitted by the source during the first phase is given by

$$\mathbf{x} = \mathbf{F}\mathbf{s}. \tag{2.1}$$

The corresponding signal received by the relay can be expressed as

$$\mathbf{z} = \mathbf{H}\,\mathbf{x} + \mathbf{n}_{\mathtt{R}} = \mathbf{H}\,\mathbf{F}\mathbf{s} + \mathbf{n}_{\mathtt{R}}, \tag{2.2}$$

where $\mathbf{H} \in \mathbb{C}^{N_{\mathtt{R}} \times N_{\mathtt{S}}}$ denotes the channel matrix between the source and relay antennas, and $\mathbf{n}_{\mathtt{R}} \in \mathbb{C}^{N_{\mathtt{R}} \times 1}$ is a spatially white additive noise vector with zero mean and covariance matrix $\mathbb{E}\{\mathbf{n}_{\mathtt{R}}\mathbf{n}_{\mathtt{R}}^H\} = \sigma_{\mathtt{R}}^2 \mathbf{I}_{N_{\mathtt{R}}}$.

Upon receiving $\mathbf{z}$, the relay applies a linear AF transformation represented by matrix $\mathbf{W} \in \mathbb{C}^{N_{\mathtt{R}} \times N_{\mathtt{R}}}$ to $\mathbf{z}$, and then forwards the resultant signal

$$\mathbf{r} = \mathbf{W}\mathbf{z} = \mathbf{W}\,\mathbf{H}\,\mathbf{F}\mathbf{s} + \mathbf{W}\mathbf{n}_{\mathtt{R}} \tag{2.3}$$

to the destination. Hence, during the second phase of the transmission, the received signal at the destination is given by

$$\begin{aligned} \mathbf{y} &= \mathbf{G}\,\mathbf{r} + \mathbf{n}_{\mathtt{D}} \\ &= \mathbf{G}\,\mathbf{W}\,\mathbf{H}\,\mathbf{F}\mathbf{s} + \mathbf{G}\,\mathbf{W}\mathbf{n}_{\mathtt{R}} + \mathbf{n}_{\mathtt{D}}, \end{aligned} \tag{2.4}$$

where $\mathbf{G} \in \mathbb{C}^{N_{\mathtt{D}} \times N_{\mathtt{R}}}$ denotes the channel matrix between the relay and destination antennas, and $\mathbf{n}_{\mathtt{D}}$ denotes a spatially white additive noise vector with zero mean and covariance matrix $\sigma_{\mathtt{D}}^2 \mathbf{I}_{N_{\mathtt{D}}}$. To estimate the signal received from the source, the destination applies a linear receive filter[2], represented by matrix $\mathbf{U} \in \mathbb{C}^{N_{\mathtt{D}} \times d}$ to the received signal, i.e., $\hat{\mathbf{s}} = \mathbf{U}^H\mathbf{y}$. In light of (2.4), the estimated symbols can further be written as

$$\hat{\mathbf{s}} = \mathbf{U}^H\,\mathbf{G}\,\mathbf{W}\,\mathbf{H}\,\mathbf{F}\mathbf{s} + \mathbf{U}^H\,\mathbf{G}\,\mathbf{W}\mathbf{n}_{\mathtt{R}} + \mathbf{U}^H\mathbf{n}_{\mathtt{D}}. \tag{2.5}$$

The general design problem for the transceiver matrices $\mathbf{F}$, $\mathbf{W}$ and $\mathbf{U}$, where the aim is to maximize the performance of the relaying transmission network in Fig. 2.1, can be mathematically expressed as

$$\min_{\mathbf{F},\mathbf{W},\mathbf{U}} \quad \mathcal{U}(\mathbf{F},\mathbf{W},\mathbf{U}) \tag{2.6a}$$

$$\text{s.t.} \quad \mathrm{Tr}(\mathbf{F}\mathbf{F}^H) \le P_{\mathtt{S}}^{\max} \tag{2.6b}$$

$$\mathrm{Tr}(\mathbf{W}\mathbf{R}_{\mathtt{z}}\mathbf{W}^H) \le P_{\mathtt{R}}^{\max}, \tag{2.6c}$$

where $\mathcal{U}(\cdot)$ denotes the chosen performance metrics or objective function, $P_{\mathtt{S}}^{\max}$ and $P_{\mathtt{R}}^{\max}$

[2]In general, a nonlinear receiver at the destination such as the decision-feedback receiver can also be adopted in the relaying network, see, e.g., [16], which leads to a different signal model. In this thesis, the focus will be on the linear receiver due to its simplicity.

denote the power budgets of the source and relay, respectively, and $\mathbf{R}_{\mathrm{z}} = \mathbb{E}\{\mathbf{z}\mathbf{z}^H\}$ denotes the covariance matrix of $\mathbf{z}$ as introduced in (2.2).

The transceiver optimization problem for the point-to-point MIMO AF relaying model in Fig. 3.1 has been extensively studied under various design objectives, including, e.g., the sum MSEs of all substreams, the mutual information between the source and destination, the worst substream MSE, the worst substream SINR and the product of all substream SINRs, see, e.g., [6, 17] and the references therein. In fact, the transceiver optimization problem in the above model is closely related to that in the point-to-point MIMO system [18, 19], with however some important differences: (1) received signal at the destination involves the product of source-relay and relay-destination channel matrices, i.e., $\mathbf{G}$ and $\mathbf{H}$ (compared to only one channel matrix in the point-to-point MIMO system); (2) the noise term received at the relay is amplified and forwarded to the destination along with the source signal; (3) the transceiver matrices at the various nodes are nonlinearly coupled [c.f. (2.5)] in the estimated signal and need to be jointly designed. These issues make the transceiver design problem in relaying systems substantially more difficult to solve than that in the point-to-point MIMO system. In the literature, we distinguish two main classes of solutions to the relay transceiver design problem (2.6), namely: the semi-closed form solutions and iterative algorithmic solutions.

The first class of solutions make use of special transceiver structures and subsequently optimize the underlying system parameters. These structures include zero-forcing (ZF) [20], minimum MSE (MMSE) [21], singular-value decomposition (SVD) [17], etc. Interestingly, it has been shown in [17] that the SVD relay transceiver structure, along with properly selected source TPC and destination receive filter becomes the optimal solution to the optimization problem (2.6) for a variety of objective functions. For example, for Schur-concave objective functions such as the sum MSEs, the optimal source TPC $\mathbf{F}$ and the relay AF transceiver matrix $\mathbf{W}$ jointly diagonalize the combined source-relay-destination channel matrix, as illustrated in Fig. 2.2. To gain further insight into this channel diagonalization scheme, denote by $\mathbf{H} \triangleq \mathbf{U}_{\mathrm{H}}\mathbf{\Lambda}_{\mathrm{H}}\mathbf{V}_{\mathrm{H}}^H$ and $\mathbf{G} \triangleq \mathbf{U}_{\mathrm{G}}\mathbf{\Lambda}_{\mathrm{G}}\mathbf{V}_{\mathrm{G}}^H$ the SVDs of $\mathbf{H}$ and $\mathbf{G}$, respectively, where $\mathbf{U}_{\mathrm{H}} \in \mathbb{C}^{N_{\mathrm{R}} \times N_{\mathrm{R}}}$ and $\mathbf{U}_{\mathrm{G}} \in \mathbb{C}^{N_{\mathrm{D}} \times N_{\mathrm{D}}}$ are unitary matrices, $\mathbf{\Lambda}_{\mathrm{H}} \in \mathbb{C}^{N_{\mathrm{R}} \times N_{\mathrm{S}}}$ and $\mathbf{\Lambda}_{\mathrm{G}} \in \mathbb{C}^{N_{\mathrm{D}} \times N_{\mathrm{R}}}$ are rectangle diagonal matrices, and $\mathbf{V}_{\mathrm{H}} \in \mathbb{C}^{N_{\mathrm{S}} \times N_{\mathrm{S}}}$ and $\mathbf{V}_{\mathrm{G}} \in \mathbb{C}^{N_{\mathrm{R}} \times N_{\mathrm{R}}}$ are unitary matrices. It is assumed that the main diagonal elements of $\mathbf{\Lambda}_{\mathrm{H}}$ and $\mathbf{\Lambda}_{\mathrm{G}}$, represented by $\{\lambda_{\mathrm{F},i}\}_{i=1}^d$ and $\{\lambda_{\mathrm{W},i}\}_{i=1}^d$, respective, are arranged in increasing order. The optimal source TPC and relay

transceiver matrix have the following structure:

$$\mathbf{F}^{\text{opt}} = \mathbf{V}_{\text{H},1}\mathbf{\Lambda}_{\text{F}} \tag{2.7}$$

$$\mathbf{W}^{\text{opt}} = \mathbf{V}_{\text{G},1}\mathbf{\Lambda}_{\text{W}}\mathbf{U}_{\text{H},1}^{H}, \tag{2.8}$$

where $\mathbf{\Lambda}_{\text{F}} = \texttt{diag}(\lambda_{\text{F},i}) \in \mathbb{C}^{d\times d}$ and $\mathbf{\Lambda}_{\text{W}} = \texttt{diag}(\lambda_{\text{W},i}) \in \mathbb{C}^{d\times d}$ while $\mathbf{V}_{\text{H},1}$, $\mathbf{V}_{\text{G},1}$ and $\mathbf{U}_{\text{H},1}$ contain the rightmost $d$ columns of $\mathbf{V}_{\text{H}}$, $\mathbf{V}_{\text{G}}$ and $\mathbf{U}_{\text{H}}$, respectively, i.e., those corresponding to the $d$ largest channel singular values . With the above transceiver structures, the original problem (2.6) with $\mathcal{U}(\cdot)$ chosen to be the sum MSEs reduces to the following scalar-variate problem:

$$\min_{\{\lambda_{\text{F},i}\}_{i=1}^{d},\{\lambda_{\text{W},i}\}_{i=1}^{d}} \quad \sum_{i=1}^{d}\left(1+\frac{(\lambda_{\text{G},i}\lambda_{\text{W},i}\lambda_{\text{H},i}\lambda_{\text{F},i})^2}{(\lambda_{\text{G},i}\lambda_{\text{W},i})^2+1}\right)^{-1} \tag{2.9a}$$

$$\text{s.t.} \quad \sum_{i=1}^{d}\lambda_{\text{F},i}^2 \le P_{\mathtt{S}}^{\max} \tag{2.9b}$$

$$\sum_{i=1}^{d}\lambda_{\text{W},i}^2\left(\lambda_{\text{H},i}^2\lambda_{\text{F},i}^2+1\right) \le P_{\mathtt{R}}^{\max} \tag{2.9c}$$

$$\lambda_{\text{F},i}, \lambda_{\text{W},i} \ge 0,\ i=1,\cdots,d. \tag{2.9d}$$

Solving the above problem is challenging since it is non-convex in $\{\lambda_{\text{F},i}\}_{i=1}^{d}, \{\lambda_{\text{W},i}\}_{i=1}^{d}$; therefore, a globally optimal solution cannot be efficiently found in general. The use of numerical methods such as grid-search algorithms based on primal and dual variables was proposed in [22] and [23], respectively. However, both methods become quite computationally intensive when an accurate approximation to the global minimum is desired. An alternating minimization algorithm that guarantees a good tradeoff between complexity and performance was proposed in [17]. This algorithm solves problem (2.9) in an iterative manner, i.e., with respect to $\{\lambda_{\text{F},i}\}_{i=1}^{d}$ and $\{\lambda_{\text{W},i}\}_{i=1}^{d}$, separately, one at a time with the other fixed. This leads to an iterative optimization procedure that may converge to a local optimum provided the solution sequence is properly initialized[3]. Although the concept of relay channel diagonalization has been widely adopted in the literature on MIMO

[3]Based on the theory of alternating minimization, due to the existence of the power constraint $\sum_{i=1}^{d}\lambda_{\text{W},i}^2\left(\lambda_{\text{H},i}^2\lambda_{\text{F},i}^2+1\right) \le P_{\mathtt{R}}^{\max}$, where $\{\lambda_{\text{F},i}\}, \{\lambda_{\text{W},i}\}$ are nonlinearly coupled, convergence to a local optimal point cannot be theoretically proved.

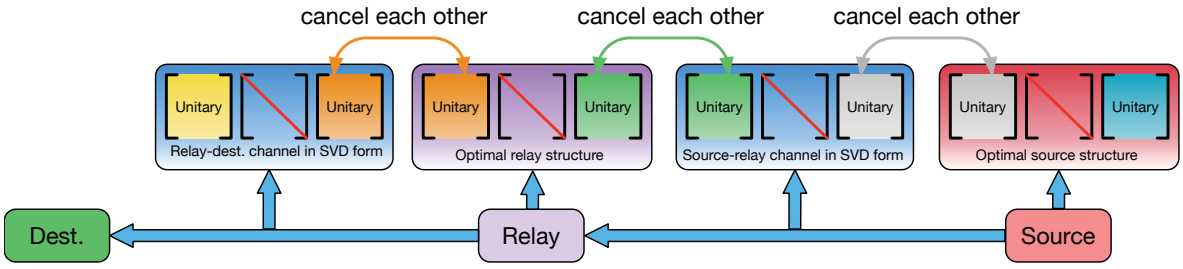



**Fig. 2.2** Illustration of the optimal source and relay transceiver structures jointly diagonalizing the combined source-relay-destination channel.

AF relaying optimization, it has its limitations. For instance, it is does not generalize easily to the case of multi-user MIMO relaying networks. Indeed, due to the existence of cross interference between multiple source-destination pairs, the channel matrix between each source-destination pair cannot in general be simultaneously diagonalized such as in Fig. 2.2. Furthermore, the extension of the method to practical scenarios where perfect CSI is not available for relay transceiver design (especially when the norm-bounded CSI error model is considered) poses additional difficulties.

The second class of solutions to (2.6), i.e., the iterative algorithmic solutions, aim to solve (2.6) numerically by relying on modern numerical optimization theory. Iterative algorithms of this type in fact have been adopted to solve a wide range of MIMO relaying problems, see, e.g., [24–30]. Exploiting the fact that (2.6) is convex in each one of $\mathbf{F}$, $\mathbf{W}$ and $\mathbf{U}$ (although it is in general not jointly convex in all its variables), the iterative algorithms rely on the BCD-type update rule, where $\mathbf{F}$, $\mathbf{W}$ and $\mathbf{U}$ are updated separately, one at a time with the others fixed, in a circular manner. Let $\mathbf{F}^{(n)}$, $\mathbf{W}^{(n)}$ and $\mathbf{U}^{(n)}$ denote the values of the corresponding variables after the $n$th iteration. Then the destination receive filter at the $n+1$ iteration can be computed by equating the gradient of $\mathcal{U}(\cdot)$ with respect to $\mathbf{U}^*$ to zero:

$$\nabla_{\mathbf{U}^*}\, \mathcal{U}(\mathbf{F}^{(n)}, \mathbf{W}^{(n)}, \mathbf{U}) = \mathbf{0}. \tag{2.10}$$

Once $\mathbf{U}^{(n+1)}$ is obtained, the source TPC $\mathbf{F}^{(n+1)}$ can be updated by solving the following

sub-problem:

$$\min_{\mathbf{F}} \quad \mathcal{U}(\mathbf{F}, \mathbf{W}^{(n)}, \mathbf{U}^{(n+1)}) \tag{2.11a}$$

$$\text{s.t.} \quad \text{Tr}(\mathbf{F}\mathbf{F}^H) \leq P_{\mathtt{S}}^{\max}. \tag{2.11b}$$

Finally, the relay AF transceiver matrix $\mathbf{W}^{(n+1)}$ can be updated by solving

$$\min_{\mathbf{W}} \quad \mathcal{U}(\mathbf{F}^{(n+1)}, \mathbf{W}, \mathbf{U}^{(n+1)}) \tag{2.12a}$$

$$\text{s.t.} \quad \text{Tr}(\mathbf{W}\mathbf{R}_{\text{z}}\mathbf{W}^H) \leq P_{\mathtt{R}}^{\max}. \tag{2.12b}$$

The generic iterative algorithmic solution is summarized in Algorithm 2.1. Different from the semi-closed form solution, it makes use of well-known results in the numerical optimization literature [31–35] and directly solves a set of simplified matrix-variate problems at each iteration. The algorithm needs to be properly initialized, for otherwise its convergence may be affected. In prior works, initialization strategies based on identity and random matrices have been employed [24], and lead to similar relaying performance and convergence rate. Regarding convergence, since each sub-problem can be solved with global optimality, the sequence of the objective values, i.e., $\{\mathcal{U}(\mathbf{F}^{(n)}, \mathbf{W}^{(n)}, \mathbf{U}^{(n)}), n = 0, 1, 2, \cdots\}$ is monotonically non-increasing with the iteration index $n$. Hence, assuming the objective function is lower bounded, the sequence of the objective values converges by invoking the monotonic convergence theorem. However, it is important to note that the convergence of the objective function does not imply the convergence of the obtained solution, i.e., $\{(\mathbf{F}^{(n)}, \mathbf{W}^{(n)}, \mathbf{U}^{(n)}), n = 0, 1, 2, \cdots\}$. The resultant solution may converge to a local optima depending on the initialization. As pointed out recently in the context of optimization theory [35], for problems in the form of (2.6) where some of the optimization variables are nonlinearly coupled, the obtained solution sequence by an iterative algorithm does not necessarily converge to a local optimum. Instead, under some mathematical conditions, that sequence converges to a so-called Nash point [35], which in effect is a weaker result.

## 2.2 Point-to-Point Cooperative AF Relaying Networks

In past years, the study of transceiver optimization in point-to-point MIMO AF relaying systems has been extended to cooperative relaying networks, where a set of relays collabo-

**Algorithm 2.1** Iterative Algorithm for Solving (2.6)

Initialization: $\mathbf{F}^{(0)}$ and $\mathbf{W}^{(0)}$.
**repeat**
1. Update $\mathbf{U}^{(n+1)}$ via (2.10);
2. Update $\mathbf{F}^{(n+1)}$ via (2.11);
3. Update $\mathbf{W}^{(n+1)}$ via (2.12);
4. Update iteration index: $n \leftarrow n+1$.

**until** Convergence.

rate in forwarding the source signals to the destination, as shown in Fig. 2.3. Let $M$ denote the total number of relays, and $\mathbf{H}_m$ and $\mathbf{G}_m$ denote the channel matrices between the source and $m^{\text{th}}$ relay, and between the $m^{\text{th}}$ relay and destination, respectively. In this case, the input-output relationship between the source and destination can still be expressed in the form of (2.4) with two key differences: the point-to-point MIMO channel matrices $\mathbf{H}$ and $\mathbf{G}$ in (2.4) now take the form of $\mathbf{H} = [\mathbf{H}_1^T, \mathbf{H}_2^T, \cdots, \mathbf{H}_M^T]^T$ and $\mathbf{G} = [\mathbf{G}_1, \mathbf{G}_2, \cdots, \mathbf{G}_M]$, respectively, and $\mathbf{W}$ becomes a block diagonal matrix, i.e., $\mathbf{W} = \texttt{blkdiag}\{\mathbf{W}_1, \mathbf{W}_2, \cdots, \mathbf{W}_M\}$. In order to fully exploit the so-called distributed array gain provided by the multiple relays, the signals forwarded by different relays should ideally be coherently combined at the destination.

The joint transceiver optimization problem in cooperative relaying networks becomes more difficult than its counterpart in single-relay networks. For the semi-closed form solution, the SVD-based method [17] may not be readily extended to the case of multiple physically separated relays due to the fact that their combined transceiver matrix $\mathbf{W}$ now exhibits a block-diagonal structure. In [36], by relaxing the set of per-relay power constraints into a single total relay power constraint, it has been shown that for the problem of minimizing the sum MSEs, the optimal source TPC and relay AF transceiver matrices jointly diagonalize the overall channel matrix, i.e., $\boldsymbol{\mathcal{H}} = \mathbf{GWHF}$, up to an arbitrary unitary matrix; however, optimizing this unitary matrix remains a challenging task. In [37], the relay transceiver matrices were jointly optimized to minimize the total relay transmission power subject to a set of QoS constraints on each transmitted substream. Similar to the method presented in [36], a suboptimal solution was obtained by replacing the objective function with a lower bound, which facilitates the diagonalization of the overall channel matrix. Other semi-closed form solutions assuming specific relay transceiver structures have been proposed, including e.g., the linear MMSE [38,39], the QR decomposition [40,41], etc.

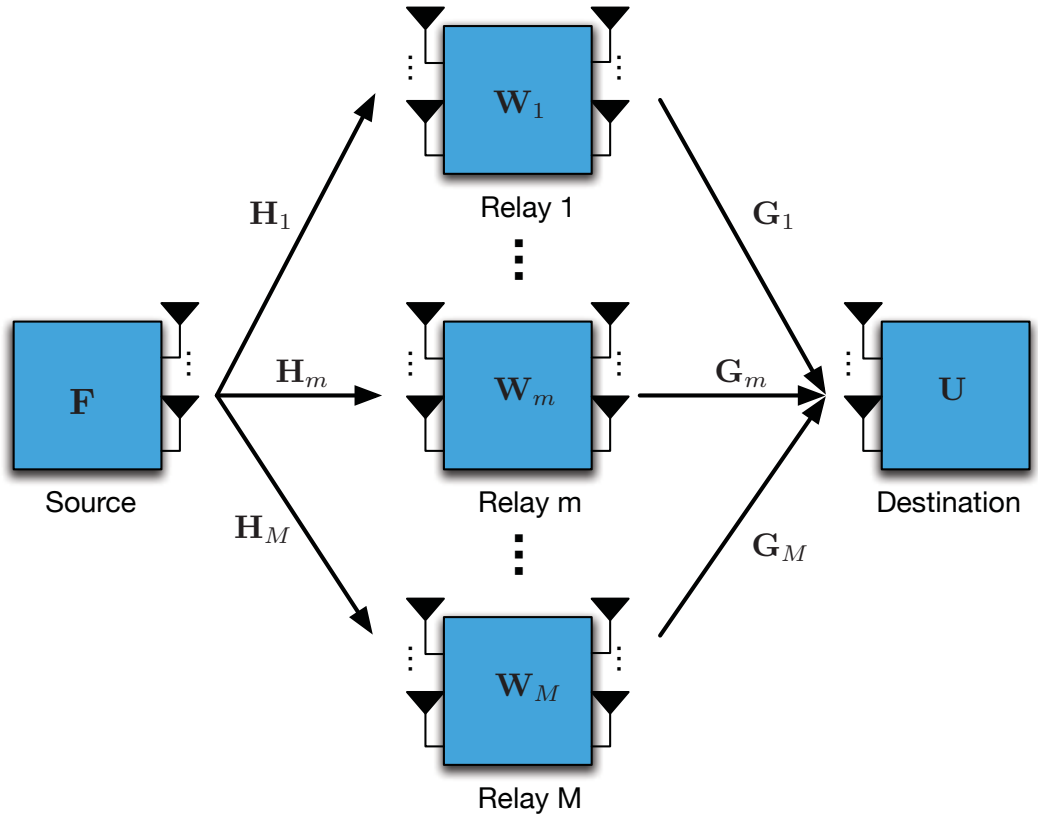



**Fig. 2.3** A point-to-point cooperative MIMO AF relaying network.

In particular, the authors in [39] derived a closed-form solution to the optimal relay AF transceiver matrix based on the MMSE criterion under a global relay power constraint. A brief summary of joint transceiver optimization in cooperative relaying networks can also be found in [6].

Various iterative algorithmic solutions have also been proposed to solve the joint transceiver optimization problem for the cooperative MIMO AF relaying network presented in Fig. 2.3, see, e.g., [26, 42–44]. These methods do not assume specific transceiver structures; instead, they aim to solve for the optimal transceiver matrices directly by relying on various numerical optimization techniques. Numerical algorithms of this kind have been proposed based on different optimization techniques including: BCD [26], gradient descent [42], gradient projection [43] and other related approaches such as the combination of BCD and gradient descent [44].

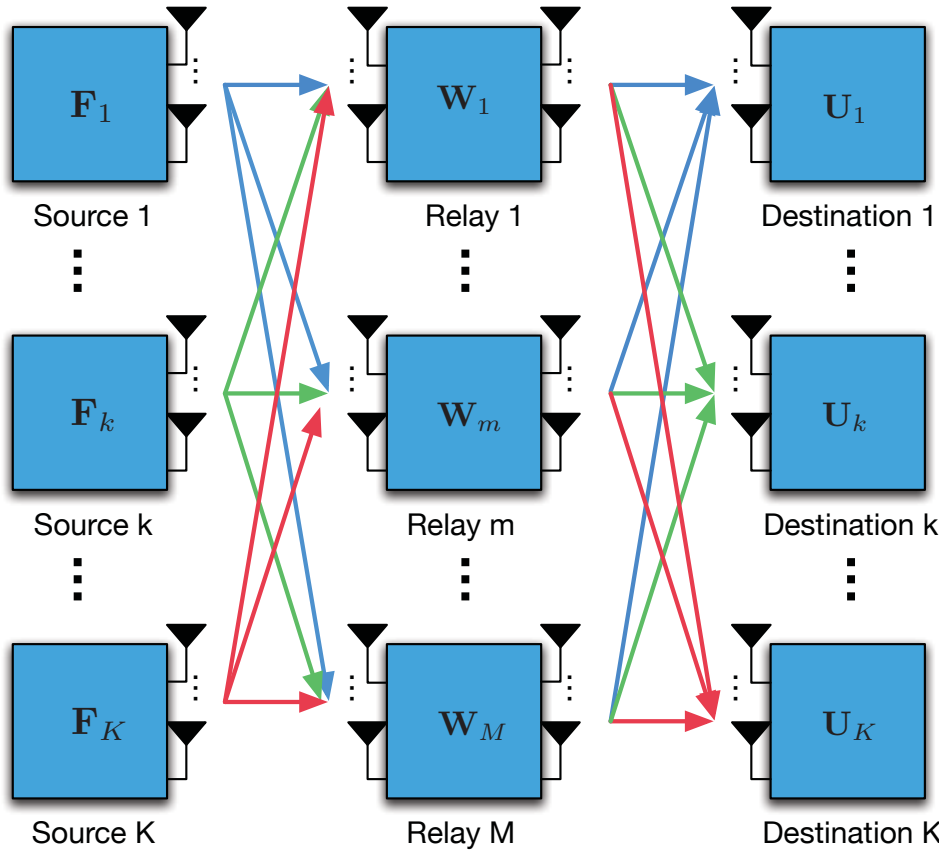



**Fig. 2.4** A multi-user cooperative MIMO AF relaying network.

## 2.3 Multi-User Cooperative AF Relaying Networks

In a typical multi-user wireless network, the amount of spectral resources available to each user decreases with an increase in the density of users sharing the resources, hence leading to a degradation in the QoS for each user. MIMO AF relaying is emerging as a promising technique for mitigating this fundamental limitation [45]. By exploiting the distributed array gain at the MIMO-aided relays, it becomes possible for multiple source-destination pairs to communicate concurrently with an acceptable QoS over the same bandwidth.

In a simplified multi-user framework where each source and destination are equipped with a single antenna, the relay transceiver matrix optimization has been extensively studied under different design criteria [46–50]. In general, designing the optimal relay AF matrix in these configurations is deemed challenging, because the resultant optimization problems

are generally non-convex. Hence, existing algorithms have mostly relied on convex approximation techniques, e.g., semidefinite relaxation (SDR) [49, 50] and second-order cone programming (SOCP) approximation [47, 48], in order to obtain approximate solutions to the original design problems.

When multiple antennas are simultaneously employed by the sources, relays and destinations, as illustrated in Fig. 2.4, the joint transceiver design problem becomes even more challenging. In particular, the results in [46–52], which focus on sources and destinations with single antennas are not readily extendable to this more general case. The literature on joint transceiver design for multi-user cooperative relaying networks is still relatively limited in spite of their potential additional performance benefits.

To be specific, in [28], global objective functions such as the sum power of the interference received at all the destinations and the sum MSEs of all the estimated substreams, were minimized by adopting the alternating minimization approach of [53], where only a single design variable is updated at each iteration based on the SDR technique of [54]. In [29], a similar type of iterative algorithms was proposed to solve the worst substream MSE minimization problem while in [55], a simplified model, where only one MIMO relay is used in the transmission, was considered. The authors of [25] proposed an alternative problem formulation, where the objective is to minimize the total source and relay power subject to a minimum SINR requirement for each source-destination link. To this end, a two-level iterative algorithm was proposed, which also involves SDR. Since the main goal of [25] was to achieve high spatial diversity gain to improve the attainable transmission integrity, the number of substreams transmitted by each source in this setting is limited to one. In [56], the authors formulated the joint transceiver optimization problem as a linear MMSE problem subject to relay power constraints. They showed that the resultant optimization problem belongs to the class of so-called quadratic programming problems, and proposed an algorithmic framework based on solving a sequence of quadratic matrix programs. This framework is in fact applicable to a wider range of transceiver design problems such as multi-cell coordinated systems and multi-user MIMO systems.

For completeness, it is worth noting that in [47], a different formulation was proposed where the aim is to minimize the total power of received interference plus noise at all destinations subject to linear distortionless constraints on the desired signals. An adaptive relaying algorithm based on the Kalman filter was proposed to solve this problem recursively over time. However, this algorithm suffers from some limitations including the use of single-

antenna sources and destinations, and use of a suboptimal power control.

## 2.4 Modeling of CSI Errors

Joint source TPC and relay transceiver matrix optimization is considered as a closed-loop transmission technique, which often needs to be performed at a central processing node (e.g., a relay cluster head or the (baseband unit) BBU pool in a C-RAN [57]). Hence, the CSI of all radio links involved in the transmission (all source-relay and relay-destination links) is required at the central node. However, as pointed out in Chapter 1 and also illustrated in Fig. 1.2, the available CSI at the central node is often inaccurate due to the combined effects of various sources of imperfections. To address this issue, the availability of robust transceiver designs, which explicitly take into account the effects of CSI errors, becomes crucial. Depending on the assumptions concerning the CSI errors, robust designs fall into two major categories, namely, the *statistically* robust [58] and *worst-case* robust designs [59]. The former class models the CSI errors as random variables with certain statistical distributions (e.g., Gaussian distributions), and robustness is achieved by optimizing the average performance over all the CSI error realizations. The latter class assumes that the CSI errors belong to some predefined bounded uncertainty region in a multi-dimensional space, such as a norm-bounded region, and optimizes the worst-case performance for all the possible CSI errors within the region.

Let us represent a general flat fading channel by $\mathbf{H} \in \mathbb{C}^{M\times N}$, where $M$ and $N$ denote the number of receive and transmit antennas, respectively. When CSI errors are present, we can express $\mathbf{H}$ using the following additive model:

$$\mathbf{H} = \hat{\mathbf{H}} + \Delta\mathbf{H}, \tag{2.13}$$

where $\hat{\mathbf{H}}$ denotes the estimated CSI while $\Delta\mathbf{H}$ captures the CSI errors [48, 49]. Below we briefly review the two aforementioned CSI error models of interest.

**(1) Statistical Error Model:** In this model, we assume that the elements of $\Delta\mathbf{H}$ are zero-mean complex Gaussian random variables. Specifically, based on the Kronecker model

[58, 60], they can in general be written as

$$\Delta\mathbf{H} = \mathbf{\Sigma}_{\mathrm{H}}^{1/2}\Delta\mathbf{H}^{\mathrm{W}}\mathbf{\Psi}_{\mathrm{H}}^{1/2}, \tag{2.14}$$

where $\mathbf{\Sigma}_{\mathrm{H}} \in \mathbb{C}^{M\times M}$ denotes the row correlation matrix, while $\mathbf{\Psi}_{\mathrm{H}} \in \mathbb{C}^{N\times N}$ is the column correlation matrix, both being positive definite. The entries of $\Delta\mathbf{H}^{\mathrm{W}}$ are independently and identically distributed (i.i.d.) complex Gaussian random variables with a zero mean and unit variance[4]. Using the Kronecker property, it can be shown that $\mathtt{vec}(\Delta\mathbf{H}^T) \sim \mathcal{CN}(\mathbf{0}_{MN\times 1}, \mathbf{\Sigma}_{\mathrm{H}} \otimes \mathbf{\Psi}_{\mathrm{H}})$. In this case, $\Delta\mathbf{H}$ is said to have a matrix-variate complex Gaussian distribution with its probability density function (PDF) given by

$$f(\Delta\mathbf{H}) = \frac{\exp(-\operatorname{Tr}(\Delta\mathbf{H}^H\mathbf{\Sigma}_{\mathrm{H}}^{-1}\Delta\mathbf{H}\mathbf{\Psi}_{\mathrm{H}}^{-1}))}{\pi^{MN}\det^M(\mathbf{\Sigma}_{\mathrm{H}})\det^N(\mathbf{\Psi}_{\mathrm{H}})}. \tag{2.15}$$

The expressions of $\mathbf{\Sigma}_{\mathrm{H}}$ and $\mathbf{\Psi}_{\mathrm{H}}$ in (2.14) depend on the specific channel estimation algorithm being used. For example, when the algorithm in [61] is adopted, we have $\mathbf{\Sigma}_{\mathrm{H}} = \mathbf{R}_T$ and $\mathbf{\Psi}_{\mathrm{H}} = \sigma_e^2\mathbf{R}_R$ with $\mathbf{R}_T$ and $\mathbf{R}_R$ denoting the transmit and receive antenna correlation matrices, respectively, and $\sigma_e^2$ denoting the estimation error variance. When channels are estimated using other algorithms, such as the one in [62], $\mathbf{\Sigma}_{\mathrm{H}}$ and $\mathbf{\Psi}_{\mathrm{H}}$ can have different structures. However, since $\mathbf{\Sigma}_{\mathrm{H}}$ and $\mathbf{\Psi}_{\mathrm{H}}$ represent the second-order statistics of the CSI, they are generally assumed to be known and remain approximately constant over several transmission blocks.

Channel quantization errors can also be modeled as the above, assuming however that $\Delta\mathbf{H}$ has i.i.d. elements, i.e., $\mathbf{\Sigma}_{\mathrm{H}} = \mathbf{I}_M$ and $\mathbf{\Psi}_{\mathrm{H}} = \sigma_e^2\mathbf{I}_N$. It has been shown in [63] that using rate distortion theory, $\mathbf{H}$ can be quantized by $b$ bits with a quantization error variance $\sigma_e^2 = 2^{-b/MN}$. Hence, the feedback overhead associated with $\mathbf{H}$ is given by

$$b = MN\log_2\left(\frac{1}{\sigma_e^2}\right). \tag{2.16}$$

**(2) Norm-Bounded Error Model:** An alternative way to model the CSI error is to assume that $\Delta\mathbf{H}$ is bounded in its Euclidean norm by some known constant [59], as repre-

[4]The superscript "W" simply refers to the spatially white or uncorrelated nature of these random variables.

sented by the set

$$\mathcal{H} \triangleq \{\Delta\mathbf{H} : \|\Delta\mathbf{H}\|_F \leq \eta\}, \tag{2.17}$$

where $\eta > 0$ specifies the radius of the uncertainty region, thus reflecting the degree of uncertainties. It should be emphasized that the actual error $\Delta\mathbf{H}$ is assumed to be unknown while the corresponding upper bound $\eta$ can be obtained using the preliminary knowledge of the type of imperfection and/or coarse knowledge of the channel type and its main characteristic. Interestingly, the determination of the radius $\eta$ can also be linked to the above statistical error model. For instance, given the PDF in (2.15), one can compute $\eta$ such that $\Pr(\|\Delta\mathbf{H}\|_F \leq \eta) \geq \alpha$, where $\alpha$ provides a measure of confidence (e.g., 95%). The reader is referred to [59] for additional discussions regarding the choice of $\eta$. The benefits of the norm-bounded error model have been well justified in the literature on robust relay optimization (see, e.g., [48, 49, 64]).

## 2.5 Energy Efficiency in Emerging Networks

A tradeoff between the relaying performance and network energy consumption has been observed in many prior works, see, e.g., [25, 48, 49, 51, 65–67]. That is, to simultaneously support multiple destination users with enhanced QoS requirements, more relays and/or higher relay transmission power are required, which is undesired from the perspective of energy efficient wireless networking. Understanding how to reduce network energy consumption while providing acceptable QoSs to the destination users therefore becomes an important challenge in the design of next-generation radio access networks [10].

In the literature, most research efforts devoted to designing energy-efficient relaying networks have been focusing on minimizing the total relay transmission power while meeting specific QoS requirements on the destination users. For instance, in [65], the source TPC and relay transceiver matrix were jointly optimized to minimize the relay transmission power while satisfying the QoS requirements of each substream in terms of the MSE. Similar to the results of [17], this approach leads to the diagonalization of the source-relay-destination channel up to a unitary matrix. The idea of channel diagonalization was then extended to the case of multiple cooperative relaying in [37]. In [66], the authors considered a single-antenna multi-user relay network and jointly selected the source power levels and

relay weights to minimize the total transmission power subject to a set of predetermined QoS constraints. The optimization problem is transformed into a DC program, which is then solved with local optimality using the constrained concave convex procedure (CCCP). The CCCP algorithm was subsequently adapted to solve the joint source TPC and relay transceiver optimization in a downlink (DL) relay scenario [67]. An extension of the model in [66] was considered in [51], where all the relays are now equipped with multiple antennas. An algorithmic solution relying on SOCP along with grid search was proposed, which however, leads to higher computational complexity. In [48] and [49], the relay transmission power minimization problem was considered under the assumption of imperfect CSI; robust algorithm were developed to solve the formulated non-convex nonlinear optimization problems. A more general multi-user MIMO relay network, where multiple antennas are employed by all the nodes, was considered in [25]. A two-tier iterative algorithm, which solves a sequence of semidefinite programs (SDPs), was proposed to minimize the total source and relay power subject to a minimum SINR requirement for each source-to-destination link.

Recently, a novel centralized radio access network architecture, called the C-RAN, has been proposed for 5G networks in order meet the increasing demand for high data-rate applications [57, 68]. In this new architecture, traditional BS functionalities are apportioned between a centralized BBU pool and RRHs. The BBU handles the baseband signal processing functions while the RRHs provide wireless connectivity to the mobile subscribers. Besides reducing the deployment and maintenance costs, C-RAN can improve network spectral efficiency by exploiting cloud computing to jointly process user data and perform interference coordination. To this end, low-latency and high-bandwidth optimal transport links are required to enable the exchange of large amounts of user traffic and control signals between the BBU pool and RRHs. The use of powerful BBU, multiple RRHs and high-speed transport links inevitably introduces additional power consumption [57], which has motivated various research efforts devoted to designing energy-efficient C-RAN.

Energy-efficient transmit solutions for C-RAN have been extensively studied for the scenarios of DL and uplink (UL) scenarios in cellular networks. For instance, a so-called *group sparse* beamforming framework has been proposed for network energy minimization in a C-RAN DL multicast scenario [69]. Capitalizing on the compressive sensing theory, a so-called group sparsity pattern is associated with the set of beamforming vectors across all the RRHs and used for obtaining a sparse solution, where only a subset of the RRHs

are activated at any instant. The effect of imperfect CSI has been considered in [70] under the same network setup and a robust version of the group sparse beamforming solution has also been proposed with the aid of semidefinite relaxation. The joint DL and UL energy minimization problem for C-RAN has been studied in [71], where the UL receive-beamforming problem is equivalently translated into an equivalent virtual DL problem by invoking the well-established UL-DL duality. In [72], sparse beamforming has been also applied in solving the network utility maximization problem under the nonlinear constraint posed by finite-capacity fronthaul links. Large-scale optimization methods for solving various transceiver design problems in C-RAN are surveyed in [73].

While most of the prior contributions focus on coordinated beamforming in both DL and UL C-RAN scenarios, the exploitation of RRHs as relays for further improving the network capacity and coverage has remained largely undisclosed. Specifically, in addition to the relay transmission power, the static power consumption from the RF circuitry, A/D conversion and optical fronthaul associated with each relay needs to be considered in the system design. That is, the total power consumption of a relay can now be expressed as $P_{\text{total}} = P_c + P_t$, where $P_c$ denotes the static power level, which becomes non-negligible compared to the transmission power $P_t$, as reported in [57]. Hence, when a relay can be switched off, both its static and transmission power can be saved. The consideration of the static power motivates a new design approach of joint relay selection and transceiver optimization, where the objective is to adaptively select a subset of active relays while satisfying the QoS constraints at all destination users.

## 2.6 Physical Layer Security for Wireless Relaying

Relay transceiver structures can be optimized from the perspectives of link performance and energy efficiency as seen in the previous sections. Another yet important aspect of wireless relaying design is the secrecy of relayed transmissions. As mentioned in Chapter 1, physical layer security is emerging as a promising alternative to complement the conventional encryption provided by the higher layer in the protocol stack. Physical layer security refers to signal processing techniques that exploit underlying wireless channel characteristics in order to reduce the ability of malicious attackers, e.g., eavesdroppers, to intercept sensitive communications, thereby improving security. In this thesis, emphasis will be placed on the use of physical layer techniques for improving the secrecy of relayed transmissions in the

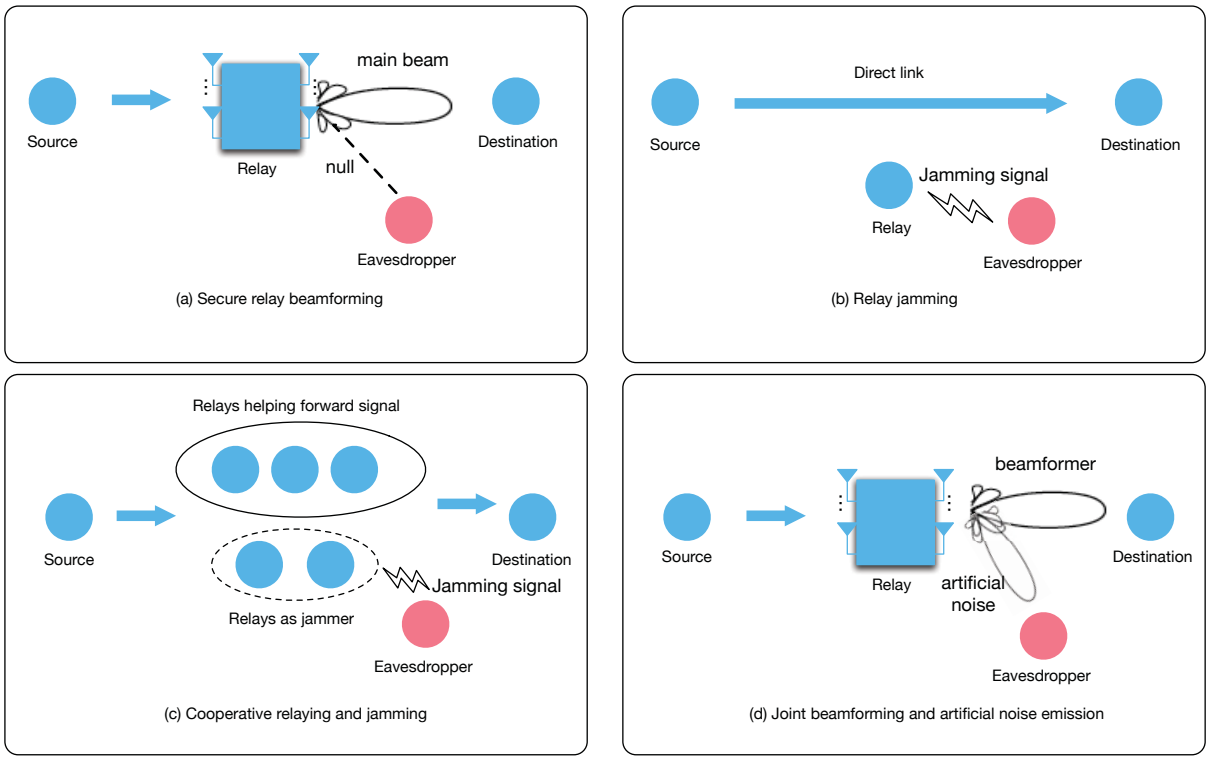



**Fig. 2.5** Relay-aided signal processing approaches for physical layer security enhancement in the presence of an eavesdropper.

presence of eavesdroppers. Below we review related works on this research topic.

In the literature, depending on the different roles played by relays and their corresponding signal processing strategies, prior works can be classified into four general classes.

- **Secure Beamforming:** In this category, one or more multi-antenna relays forward the signal transmitted from the source to the legitimate destination using a beamforming-aided approach, as sketched in Fig. 2.5 (a). Usually it is assumed that there is no direct link between the source and destination due to a high path loss. As in conventional relaying, a relay can adopt either the AF or DF strategy for forwarding the source message. For DF relaying, the optimal beamforming weights achieving the maximum secrecy capacity were derived in [74,75]. Compared to DF, AF relaying offers its inherent advantages of lower signal processing complexity and latency, and hence, will be the focus of our attention in this thesis. Intuitively, the motivation behind the use of AF relaying for improving transmission secrecy is the possibility to design the AF transceiver matrix such that its main transmit beam is focused

towards the legitimate destination as much as possible, while steering nulls towards the eavesdroppers. For instance, the optimal AF relaying weights maximizing the secrecy capacity of a single-antenna relay network were derived in [76], but without considering the information leakage from the source to eavesdroppers.

In the case where multiple antennas are employed at both the source and relay, joint transmit precoding and power allocation relying on the generalized SVD was proposed in [77]. Joint source precoding and AF relaying was investigated in [78] assuming an untrusted relay node. In [79], the secure AF relaying was designed for a two-way MIMO relaying network, as an extension to the one-way MIMO relaying network. In [80], the relay AF transceiver matrix was optimized to maximize the transmission reliability, e.g., SINR at the legitimate destination subject to a set of constraints on the leakage capacity, where a norm-bounded CSI error model is considered for all eavesdropper channels. In [81], intercept probability-constrained approach was considered for designing the MIMO AF relaying, where the CSI errors associated with the eavesdropper channels are assumed to be Gaussian distributed.

- **Relay-Aided Jamming:** When there exists a direct link between the source and legitimate destination, relays do not necessarily need to forward the information. Instead, they can cooperatively transmit AN, which is assumed to be independent of the confidential information, in order to jam the eavesdroppers (see Fig. 2.5 (b)). A widely-used jamming strategy is called null space jamming, which aims to transmit AN such that the latter resides in the null space of the relay-to-destination channel. By doing so, it can be ensured that the legitimate channel will not be affected by the AN, which may degrade the quality of the eavesdroppers' received signals. Null space jamming is relatively easy to implement but may be suboptimal. In [75] and [82], the authors considered a single-antenna relay network and developed a one-dimensional search method to compute an optimal distributed relay beamformer for jamming under total relay power constraint and per-relay power constraints, respectively.

- **Mixed Beamforming and Jamming:** This strategy simply combines the above two approaches by allowing a subset of the relays to cooperatively forward the confidential data to the legitimate destination while the remaining ones act as jammers. The allocation of relays between beamforming and jamming relays can be either predetermined or optimized based on the conditions affecting the different relays such

as: channel quality, relay locations and transmission power budgets. When the relay allocation is predetermined, the aforementioned techniques can be applied across the group of relays as in, e.g., [83, 84]. When the relay allocation is performed jointly with the beamforming and jamming design, the formulated problem is non-convex and quite challenging. Some heuristic relay selection methods such as exhaustive search [85] and convex relaxation [86] have been proposed. For instance, in [86], a one source, multiple relay and one destination model was considered, where all nodes are assumed to be single-antenna. The weights for relaying and jamming, and the relay allocation were jointly optimized to maximize the secrecy rate subject to a total relay power constraint. To this end, an SDR-based approach was proposed, which transforms the original problem into a convex SDP by relaxing an underlying rank-one constraint.

- **Joint Beamforming and Artificial Noise Emission:** This strategy is the most general one since it allows all the relays to simultaneously forward legitimate information and transmit AN. This approach fully exploits the available degrees of freedom at each relay when the latter are equipped with multiple antennas. However, owing to the coupling of the beamformer signal and AN generation at each relay, it is in general challenging to optimize the beamforming matrix and AN jointly. The literature on this type of secure relaying design methods is still limited. For example, [87] considered the problem of secrecy rate maximization by simultaneously optimizing the AF matrices and the AN covariance matrices at multiple relays. A globally optimal solution was obtained by resorting to a bi-level optimization framework, where the upper-level problem is tackled by one-dimensional search, while the inner-level problem is solved by the SDR. One of our focus in this thesis will be on the joint beamforming and AN design for secure communications in relaying networks.

Many prior works on physical layer security assume perfect knowledge of the eavesdropper's channel state information (ECSI) at the legitimate nodes. In practice, due to the lack of explicit cooperation between the latter and the eavesdroppers, at best an inaccurate estimate of the ECSI may be available. In the context of this thesis, this fact motivates the study of secure relaying designs, which also offers robustness to uncertainties in the eavesdropper channels. In [80], assuming that the ECSI errors lie in a predefined norm-bounded region, a secure AF relaying approach that is robust to the ECSI errors was

proposed to maximize reception quality at the legitimate destination while In [81], knowledge of the distribution of the ECSI errors (e.g., Gaussian) was assumed, and an intercept probability constrained maximum SINR beamforming scheme was proposed for an MIMO AF relay network. Assuming that the ECSI errors lie in a predefined norm-bounded region, a secure AF relaying scheme that is robust to the ECSI errors was proposed to maximize the reception quality at the legitimate destination while enforcing the information leakage to the eavesdroppers to fall below a threshold [80]. Under the same ECSI error model, joint relay beamforming and jamming signal design in a single-antenna relay network was developed in [86, 88], with the aim of maximizing the worst-case secrecy rate. Extension of this approach to a more general model where multiple antennas are employed at the relay was considered in [87, 89]. Minimization of the MSE of the received signal at the destination, subject to a set of SINR-based secrecy constraints, was considered in [79]. Using the same uncertainty model, the problem of total relaying power minimization was studied in [90,91] by simultaneously guaranteeing a predefined QoS level at the destination and a certain secrecy level against eavesdropping. Finally, a distributionally robust design approach was investigated in [92] under the assumption that the ECSI errors follow some unknown distributions with the knowledge of the first and second-order statistics.

## 2.7 Concluding Remarks

In this chapter, we have presented a review of important design approaches in the optimization of MIMO AF relay transceivers from the perspectives of user QoS, energy efficiency and transmission secrecy. In addition, modeling of the CSI errors within this framework has been addressed. This literature review provides motivations and serves as basis for the research contributions presented in subsequent chapters of this thesis. More specifically, Chapter 3 develops robust optimization algorithms for a general multi-user MIMO cooperative relaying network. Chapter 4 studies energy-efficient MIMO relaying algorithms in a C-RAN. Chapter 5 focuses on the robust secure relaying design in the presence of multiple eavesdroppers. It is also worth noting that in this thesis, we make excessive use of concepts and techniques from modern optimization theory. Interested readers are referred to [93] for the fundamentals of convex optimization and to [94–96] for more specific methods which are closely related to the algorithm development and analysis in this thesis.

# Chapter 3

# Robust Joint MIMO Relay Transceiver Optimization

As mentioned in Chapters 1 and 2, the efficacy of the joint transceiver design in MIMO relaying networks relies on the availability and accuracy of the CSI for all radio links involved in the relaying transmission. In practice, the CSI available for transceiver optimization at the central node is often subject to errors. The performance of the conventional relay transceiver optimization methods may hence be substantially degraded in the presence of realistic CSI errors. Robust transceiver designs, which explicitly take into account the effects of CSI errors, are highly desirable.

In this chapter, we study the joint transceiver design in a general MIMO multi-user relay network, where multiple source (S)-destination (D) pairs communicate with the assistance of multiple relays (R) and each source transmits multiple parallel data streams to its corresponding destination. Under the realistic assumption of imperfect CSI for all the S-R and R-D links, we propose a new robust optimization framework for minimizing the maximum per-stream MSE subject to the source and relay power constraints, which is termed as the *min-max* problem. In the proposed framework, we aim for solving both the *statistically* robust and *worst-case* robust versions of the min-max problem, which take into account either the statistical CSI errors or the norm-bounded CSI errors, respectively, while maintaining a tractable computational complexity. Furthermore, to strictly satisfy

Parts of the materials in this chapter have been presented at the 2014 IEEE 80th Vehicular Technology Conference in Vancouver, Canada [97], and published in the IEEE Transactions on Vehicular Technology [30].

the QoS specifications of all the data streams, we subsequently consider the problem of minimizing the maximum per-relay power, subject to the QoS constraints of all the data streams and to the source power constraints, which is referred to as the *QoS* problem. The main contributions of this chapter are threefold:

- The statistically robust min-max problem being non-convex, an algorithmic solution having a proven convergence is proposed by invoking the iterative block coordinate update approach of [98] while relying on both matrix transformation and convex conic optimization techniques. The proposed iterative algorithm successively solves in a circular manner three sub-problems corresponding to the source TPCs, relay AF matrices and receive filters, respectively. We show that the receive filter sub-problem yields a closed-form solution, while the other two sub-problems can be transformed to convex quadratically-constrained linear programs (QCLPs). Each QCLP can subsequently be reformulated as an efficiently solvable SOCP.

- The worst-case robust min-max problem is both non-convex and semi-infinite[1]. To overcome these challenges, we first present a generalized version of the so-called $\mathcal{S}$-lemma given in [99], based on which, each sub-problem can be exactly reformulated as an SDP with only linear matrix inequality (LMI) constraints. This results in an iterative algorithmic solution involving several SDPs.

- The QoS-based transceiver optimization is more challenging than that of the min-max problem, because it is difficult to find a feasible initialization. Hence, our major contribution here is to propose an efficient procedure for finding a feasible starting point for the iterative QoS-based optimization algorithm, provided that there exits one; otherwise, the procedure returns a certificate of infeasibility.

The rest of the chapter is organized as follows. Section 3.1 introduces the relaying system and CSI error models. The robust joint transceiver design problems are also formulated in this section. In Section 3.2 and 3.3, iterative algorithms are proposed for solving the min-max problem both under the statistical and the norm-bounded CSI error models, respectively. The QoS problem is dealt with in Section 3.4. Numerical results are reported in Section 3.5, followed by conclusions in Section 3.6.

[1]In optimization theory, semi-infinite programming is an optimization problem with a finite number of variables and an infinite number of constraints or an infinite number of variables and a finite number of constraints.

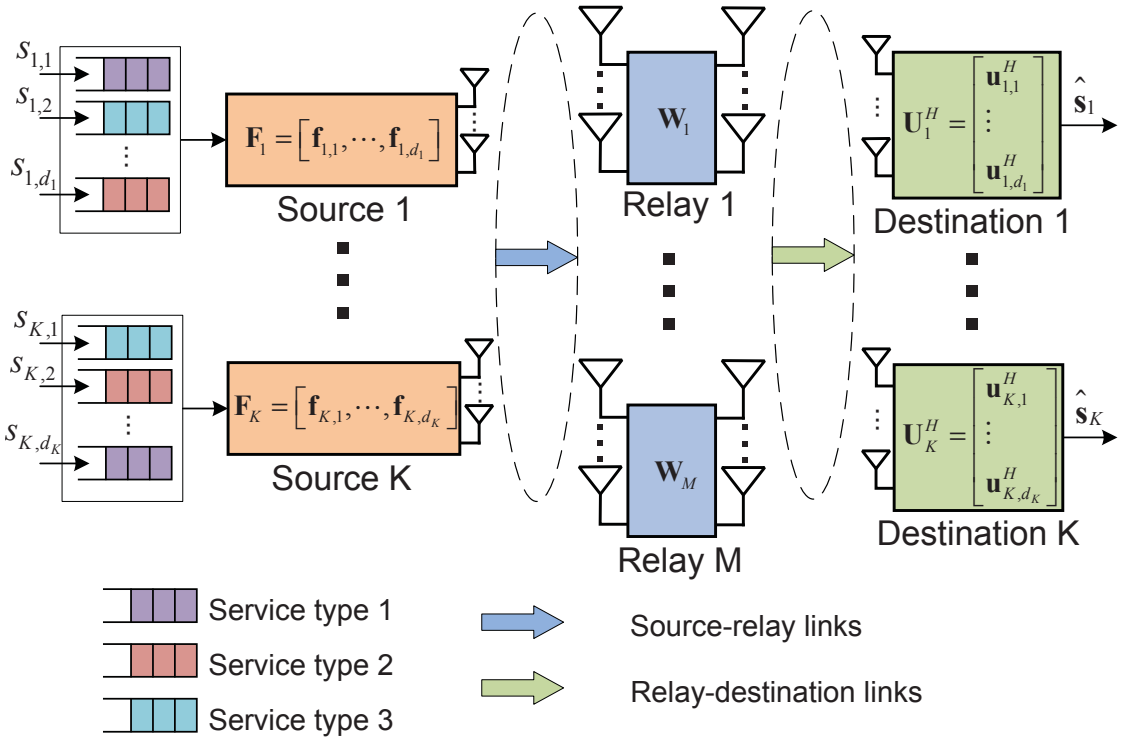



**Fig. 3.1** MIMO multi-user, multi-relay network with each source transmitting multiple data streams to its corresponding destination.

## 3.1 System Model and Problem Formulation

We consider an MIMO multi-user relaying network, where $M$ AF relay nodes assist the one-way communication between $K$ S-D pairs, as depicted in Fig. 3.1, where all the nodes are equipped with multiple antennas. Specifically, the $k^{\text{th}}$ S and D, respectively, employ $N_{\mathtt{S},k}$ and $N_{\mathtt{D},k}$ antennas for $k \in \mathcal{K} \triangleq \{1, 2, \cdots, K\}$, while the $m^{\text{th}}$ R employs $N_{\mathtt{R},m}$ antennas for $m \in \mathcal{M} \triangleq \{1, \cdots, M\}$. All the relays operate under the half-duplex AF protocol, where the data transmission from the sources to their destinations is completed in two stages. In the first stage, all the sources transmit their signals to the relays concurrently, while in the second stage, the relays apply linear processing to the received signals and forward the resultant signals to all the destinations. We assume that no direct links are available between the sources and destinations due to severe attenuation.

A narrowband flat-fading radio propagation model is considered, where we denote the channel matrix between the $k^{\text{th}}$ S and the $m^{\text{th}}$ R by $\mathbf{H}_{m,k} \in \mathbb{C}^{N_{\mathtt{R},m} \times N_{\mathtt{S},k}}$ and the channel

matrix between the $m^{\text{th}}$ R and the $k^{\text{th}}$ D by $\mathbf{G}_{k,m} \in \mathbb{C}^{N_{\mathtt{D},k} \times N_{\mathtt{R},m}}$. Let $\mathbf{s}_k \triangleq [s_{k,1}, \cdots, s_{k,d_k}]^T$ denote the information symbols to be transmitted by the $k^{\text{th}}$ S at a given time instant, where $d_k \leq \min\{N_{\mathtt{S},k}, N_{\mathtt{D},k}\}$ is the number of independent data streams. The symbols are modeled as independent random variables with a zero mean and unit variance, hence $\mathbb{E}\left\{\mathbf{s}_k \mathbf{s}_k^H\right\} = \mathbf{I}_{d_k}$. The $k^{\text{th}}$ S applies a linear vector of $\mathbf{f}_{k,l} \in \mathbb{C}^{N_{\mathtt{S},k} \times 1}$ for mapping the $l^{\text{th}}$ data stream to its $N_{\mathtt{S},k}$ antennas for $l \in \mathcal{D}_k \triangleq \{1, \cdots, d_k\}$, thus forming a linear TPC of $\mathbf{F}_k = [\mathbf{f}_{k,1}, \cdots, \mathbf{f}_{k,d_k}] \in \mathbb{C}^{N_{\mathtt{S},k} \times d_k}$. The transmit power is thus given by $\text{Tr}\left(\mathbf{F}_k \mathbf{F}_k^H\right) \leq P_{\mathtt{S},k}^{\max}$, where $P_{\mathtt{S},k}^{\max}$ is the maximum affordable power of the $k^{\text{th}}$ S. Let $\mathbf{n}_{\mathtt{R},m} \in \mathbb{C}^{N_{\mathtt{R},m} \times 1}$ be the spatially white, additive noise vector at the $m^{\text{th}}$ R, with a zero mean and covariance matrix of $\mathbb{E}\left\{\mathbf{n}_{\mathtt{R},m} \mathbf{n}_{\mathtt{R},m}^H\right\} = \sigma_{\mathtt{R},m}^2 \mathbf{I}_{N_{\mathtt{R},m}}$.

After the first stage of transmission, the signal received at the $m^{\text{th}}$ R is given by

$$\mathbf{z}_{\mathtt{R},m} = \sum_{k=1}^{K} \mathbf{H}_{m,k} \mathbf{F}_k \mathbf{s}_k + \mathbf{n}_{\mathtt{R},m}. \tag{3.1}$$

Each R applies a linear matrix $\mathbf{W}_m \in \mathbb{C}^{N_{\mathtt{R},m} \times N_{\mathtt{R},m}}$ to $\mathbf{z}_{\mathtt{R},m}$ and forwards the resultant signal

$$\mathbf{r}_{\mathtt{R},m} = \mathbf{W}_m \mathbf{z}_{\mathtt{R},m} = \sum_{k=1}^{K} \mathbf{W}_m \mathbf{H}_{m,k} \mathbf{F}_k \mathbf{s}_k + \mathbf{W}_m \mathbf{n}_{\mathtt{R},m} \tag{3.2}$$

to all the destinations at a power of

$$P_{\mathtt{R},m} = \sum_{k=1}^{K} \|\mathbf{W}_m \mathbf{H}_{m,k} \mathbf{F}_k\|_F^2 + \sigma_{\mathtt{R},m}^2 \|\mathbf{W}_m\|_F^2. \tag{3.3}$$

Let $\mathbf{n}_{\mathtt{D},k}$ denote the spatially white, additive noise vector at the $k^{\text{th}}$ D with a zero mean and covariance matrix of $\mathbb{E}\left\{\mathbf{n}_{\mathtt{D},k} \mathbf{n}_{\mathtt{D},k}^H\right\} = \sigma_{\mathtt{D},k}^2 \mathbf{I}_{N_{\mathtt{D},k}}$. The $k^{\text{th}}$ D observes the following signal after the second stage of transmission

$$\mathbf{y}_k = \sum_{q=1}^{K} \sum_{m=1}^{M} \mathbf{G}_{k,m} \mathbf{W}_m \mathbf{H}_{m,q} \mathbf{F}_q \mathbf{s}_q + \sum_{m=1}^{M} \mathbf{G}_{k,m} \mathbf{W}_m \mathbf{n}_{\mathtt{R},m} + \mathbf{n}_{\mathtt{D},k}, \tag{3.4}$$

where the subscript $q$ is now used for indexing the sources. To estimate the $l^{\text{th}}$ data stream received from its corresponding source, the $k^{\text{th}}$ D applies a linear vector $\mathbf{u}_{k,l}$ to the received signal, thus forming a receive filter $\mathbf{U}_k = [\mathbf{u}_{k,1}, \cdots, \mathbf{u}_{k,d_k}] \in \mathbb{C}^{N_{\mathtt{D},k} \times d_k}$. Specifically, the

estimated information symbols are given by $\hat{s}_{k,l} = \mathbf{u}_{k,l}^H \mathbf{y}_k$, which can be expressed as

$$\hat{s}_{k,l} = \underbrace{\mathbf{u}_{k,l}^H \sum_{m=1}^{M} \mathbf{G}_{k,m}\mathbf{W}_m\mathbf{H}_{m,k}\mathbf{f}_{k,l}s_{k,l}}_{\text{desired data stream}} + \underbrace{\mathbf{u}_{k,l}^H \sum_{m=1}^{M} \mathbf{G}_{k,m}\mathbf{W}_m\mathbf{H}_{m,k} \sum_{p=1,p\neq l}^{d_k} \mathbf{f}_{k,p}s_{k,p}}_{\text{inter-stream interference}}$$
$$+ \underbrace{\sum_{q=1,q\neq k}^{K} \mathbf{u}_{k,l}^H \sum_{m=1}^{M} \mathbf{G}_{k,m}\mathbf{W}_m\mathbf{H}_{m,q}\mathbf{F}_q\mathbf{s}_q}_{\text{inter-user interference}} + \underbrace{\sum_{m=1}^{M} \mathbf{u}_{k,l}^H\mathbf{G}_{k,m}\mathbf{W}_m\mathbf{n}_{\mathtt{R},m}}_{\text{enhanced noise from relays}} + \underbrace{\mathbf{u}_{k,l}^H\mathbf{n}_{\mathtt{D},k}}_{\text{receiver noise}} . \quad (3.5)$$

In this chapter, we make the following common assumptions concerning the statistical properties of the signals: (1) the information symbols transmitted from different S are uncorrelated, i.e., we have $\mathbb{E}\left\{\mathbf{s}_k\mathbf{s}_m^H\right\} = \mathbf{0}$, $\forall k, m \in \mathcal{K}$ and $k \neq m$; (2) the information symbols, the relay noise and the receiver noise, $\mathbf{s}_k$, $\mathbf{n}_{\mathtt{R},m}$ and $\mathbf{n}_{\mathtt{D},l}$, are mutually statistically independent $\forall k, l \in \mathcal{K}$ and $m \in \mathcal{M}$.

### 3.1.1 QoS Metric

We adopt the MSE as the QoS metric for each estimated data stream. The major advantage of using the MSE is to make our design problem tractable, which has been well justified in the AF relay matrix design literature [17,65] and in the references therein. In fact, the links between the MSE and other classic criteria such as the bit error rate (BER) and the SINR have been well established in [17, 18]. Specifically, it has been shown that an improvement in MSE will naturally lead to a reduced BER.

The MSE of the $l^{\text{th}}$ estimated data stream received at the $k^{\text{th}}$ D is defined as

$$\varepsilon_{k,l} = \mathbb{E}\left\{|\hat{s}_{k,l} - s_{k,l}|^2\right\}. \quad (3.6)$$

Substituting (3.5) into (3.6), and using the assumptions A1) and A2), we obtain

$$\varepsilon_{k,l} = \left\|\mathbf{u}_{k,l}^H \sum_{m=1}^{M} \mathbf{G}_{k,m}\mathbf{W}_m\mathbf{H}_{m,k}\mathbf{F}_k - \mathbf{e}_{k,l}^T\right\|^2 + \sum_{q=1,q\neq k}^{K} \left\|\mathbf{u}_{k,l}^H \sum_{m=1}^{M} \mathbf{G}_{k,m}\mathbf{W}_m\mathbf{H}_{m,q}\mathbf{F}_q\right\|^2$$
$$+ \sum_{m=1}^{M} \sigma_{\mathtt{R},m}^2 \left\|\mathbf{u}_{k,l}^H\mathbf{G}_{k,m}\mathbf{W}_m\right\|^2 + \sigma_{\mathtt{D},k}^2 \left\|\mathbf{u}_{k,l}\right\|^2, \quad (3.7)$$

where $\mathbf{e}_{k,l} \in \mathbb{R}^{d_k \times 1}$ is a vector with all zero entries but the $l^{\text{th}}$ entry, which is equal to one.

### 3.1.2 CSI Error Model

The two main CSI error models have been introduced in Chapter 2.4. Here, for ease of reference, we reintroduce these models for the specific model considered in this chapter. Let us characterize the true but unknown channels as

$$\mathbf{H}_{m,k} = \hat{\mathbf{H}}_{m,k} + \Delta\mathbf{H}_{m,k}, \ \mathbf{G}_{k,m} = \hat{\mathbf{G}}_{k,m} + \Delta\mathbf{G}_{k,m}, \tag{3.8}$$

where $\hat{\mathbf{H}}_{m,k}$ and $\hat{\mathbf{G}}_{k,m}$, respectively, denote the estimated S-R and R-D channels, while $\Delta\mathbf{H}_{m,k}$ and $\Delta\mathbf{G}_{k,m}$ capture the corresponding channel uncertainties [48, 49]. In what follows, we consider two popular techniques of modeling the channel uncertainties:

*1) Statistical Error Model*: In this model, we assume that the elements of $\Delta\mathbf{H}_{m,k}$ and $\Delta\mathbf{G}_{k,m}$ are zero-mean complex Gaussian random variables. Specifically, based on the Kronecker model [58, 60], they can in general be written as

$$\Delta\mathbf{H}_{m,k} = \mathbf{\Sigma}_{\mathrm{H}_{m,k}}^{1/2} \Delta\mathbf{H}_{m,k}^{\mathrm{W}} \mathbf{\Psi}_{\mathrm{H}_{m,k}}^{1/2} \tag{3.9}$$

$$\Delta\mathbf{G}_{k,m} = \mathbf{\Sigma}_{\mathrm{G}_{k,m}}^{1/2} \Delta\mathbf{G}_{k,m}^{\mathrm{W}} \mathbf{\Psi}_{\mathrm{G}_{k,m}}^{1/2}, \tag{3.10}$$

where $\mathbf{\Sigma}_{\mathrm{H}_{m,k}}$ and $\mathbf{\Sigma}_{\mathrm{G}_{k,m}}$ are the row correlation matrices, while $\mathbf{\Psi}_{\mathrm{H}_{m,k}}$ and $\mathbf{\Psi}_{\mathrm{G}_{k,m}}$ the are column correlation matrices, all being positive definite. The entries of $\Delta\mathbf{H}_{m,k}^{\mathrm{W}}$ and $\Delta\mathbf{G}_{k,m}^{\mathrm{W}}$ are i.i.d. complex Gaussian random variables with a zero mean and unit variance[2]. This model is suitable, when the CSI errors are dominated by the channel estimation errors.

*2) Norm-Bounded Error Model*: When the CSI is subject to quantization errors due to the limited-rate feedback, it can no longer be accurately characterized by the above statistical model. Instead, $\Delta\mathbf{H}_{m,k}$ and $\Delta\mathbf{G}_{k,m}$ are considered to assume values from the following norm-bounded sets [59]

$$\mathcal{H}_{m,k} \triangleq \{\Delta\mathbf{H}_{m,k} : \|\Delta\mathbf{H}_{m,k}\|_F \leq \eta_{m,k}\} \tag{3.11}$$

$$\mathcal{G}_{k,m} \triangleq \{\Delta\mathbf{G}_{k,m} : \|\Delta\mathbf{G}_{k,m}\|_F \leq \xi_{k,m}\}, \tag{3.12}$$

[2]The superscript "W" simply refers to the spatially white or uncorrelated nature of these random variables.

where $\eta_{m,k} > 0$ and $\xi_{k,m} > 0$ specify the radii of the uncertainty regions, thus reflecting the degree of uncertainties. The benefits of such an error model have been well justified in the literature of robust relay optimization (see, e.g., [48, 49, 64]). The determination of the radii of the uncertainty regions has also been discussed in [59].

Throughout the chapter, we assume that the magnitudes of the CSI errors are significantly lower than those of the channel estimates, and therefore, terms containing the cross product of CSI errors, i.e., $\Delta\mathbf{G}_{k,m}\mathbf{W}_m\Delta\mathbf{H}_{m,k}$ and $\Delta\mathbf{G}_{k,m}\mathbf{W}_m\Delta\mathbf{H}_{m,q}$ are neglected in our subsequent analysis since these terms have small norm relative to the other terms. We also introduce in Table 3.1 some useful notations to simplify our exposition.

Substituting (3.8) into (3.7), and applying the above-mentioned assumptions, the per-stream MSE in the presence of CSI errors can be approximated as (3.13). We now observe that the per-stream MSE becomes uncertain in $\Delta\mathbf{H}_{m,k}$ $\forall(m,k) \in \mathcal{M}\times\mathcal{K}$ and $\Delta\mathbf{G}_{k,m}$ $\forall m \in \mathcal{M}$. Therefore, we introduce the following compact notations for convenience:

$$\Delta\mathbf{G}_k \triangleq (\Delta\mathbf{G}_{k,1}, \cdots, \Delta\mathbf{G}_{k,M}) \in \mathcal{G}_k \triangleq \mathcal{G}_{k,1}\times\cdots\times\mathcal{G}_{k,M}$$
$$\Delta\mathbf{H} \triangleq (\Delta\mathbf{H}_{1,1}, \cdots, \Delta\mathbf{H}_{M,K}) \in \mathcal{H} \triangleq \mathcal{H}_{1,1}\times\cdots\times\mathcal{H}_{M,K}.$$

For subsequent derivations, the dependence of $\varepsilon_{k,l}$ on $\Delta\mathbf{H}$ and $\Delta\mathbf{G}_k$ is made explicit in (3.13).

The $k^{\text{th}}$ relay's transmit power in the presence of CSI errors can also be explicitly expressed as $P_{\mathtt{R},m}(\Delta\mathbf{H}_m)$, where $\Delta\mathbf{H}_m \triangleq (\Delta\mathbf{H}_{m,1}, \cdots, \Delta\mathbf{H}_{m,K}) \in \mathcal{H}_m \triangleq \mathcal{H}_{m,1}\times\cdots\times\mathcal{H}_{m,K}$.

$$\begin{aligned}\varepsilon_{k,l}(\Delta\mathbf{H},\Delta\mathbf{G}_k) \approx &\left\|\mathbf{u}_{k,l}^H\boldsymbol{\mathcal{T}}_{k,k} + \sum_{m=1}^{M}\mathbf{u}_{k,l}^H\Delta\mathbf{G}_{k,m}\boldsymbol{\mathcal{W}}_{m,k}\mathbf{F}_k + \sum_{m=1}^{M}\mathbf{u}_{k,l}^H\boldsymbol{\mathcal{G}}_{k,m}\Delta\mathbf{H}_{m,k}\mathbf{F}_k - \mathbf{e}_{k,l}^T\right\|^2 \\ &+ \sum_{q=1,q\neq k}^{K}\left\|\mathbf{u}_{k,l}^H\boldsymbol{\mathcal{T}}_{k,q} + \sum_{m=1}^{M}\mathbf{u}_{k,l}^H\Delta\mathbf{G}_{k,m}\boldsymbol{\mathcal{W}}_{m,q}\mathbf{F}_q + \sum_{m=1}^{M}\mathbf{u}_{k,l}^H\boldsymbol{\mathcal{G}}_{k,m}\Delta\mathbf{H}_{m,q}\mathbf{F}_q\right\|^2 \\ &+ \sum_{m=1}^{M}\sigma_{\mathtt{R},m}^2\left\|\mathbf{u}_{k,l}^H\boldsymbol{\mathcal{G}}_{k,m} + \mathbf{u}_{k,l}^H\Delta\mathbf{G}_{k,m}\mathbf{W}_m\right\|^2 + \sigma_{\mathtt{D},k}^2\left\|\mathbf{u}_{k,l}\right\|^2. \qquad (3.13)\end{aligned}$$

**Table 3.1** Equivalent Notations Used in the Subsequent Analysis

| Notations | Definitions |
|---|---|
| $\boldsymbol{\mathcal{G}}_{k,m}$ | $\hat{\mathbf{G}}_{k,m}\mathbf{W}_m$ |
| $\boldsymbol{\mathcal{W}}_{m,k}$ | $\mathbf{W}_m\hat{\mathbf{H}}_{m,k}$ |
| $\boldsymbol{\mathcal{U}}_{k,m}$ | $\mathbf{U}_k^H\hat{\mathbf{G}}_{k,m}$ |
| $\boldsymbol{\mathcal{H}}_{m,k}$ | $\hat{\mathbf{H}}_{m,k}\mathbf{F}_k$ |
| $\boldsymbol{\mathcal{T}}_{k,q}$ | $\sum_{m=1}^{M}\hat{\mathbf{G}}_{k,m}\mathbf{W}_m\hat{\mathbf{H}}_{m,q}\mathbf{F}_q$ |

### 3.1.3 Problem Formulation

In contrast to the prior advances [17, 28, 46–48] found in the relay optimization literature, where certain global objective functions are minimized subject to power constraints at the sources and relays, we formulate the following robust design problems under the explicit consideration of QoS. Let us commence by introducing the following unified operation:

$$\mathcal{U}\left\{f\left(\Delta\mathbf{X}\right)\right\} = \begin{cases} \mathbb{E}_{\Delta\mathbf{X}}\, f\left(\Delta\mathbf{X}\right) & \Delta\mathbf{X} \text{ is random} \\ \max\limits_{\Delta\mathbf{X}\in\mathcal{X}} f\left(\Delta\mathbf{X}\right) & \Delta\mathbf{X} \text{ is deterministic,} \end{cases} \tag{3.14}$$

where $\Delta\mathbf{X} \in \mathbb{C}^{M\times N}$ and $f\left(\cdot\right) : \mathbb{C}^{M\times N} \to \mathbb{R}$. Depending on the specific assumptions concerning $\Delta\mathbf{X}$, $\mathcal{U}\left\{\cdot\right\}$ either computes the expectation of $f\left(\Delta\mathbf{X}\right)$ over the ensemble of realizations $\Delta\mathbf{X}$ or maximizes $f\left(\Delta\mathbf{X}\right)$ for all $\Delta\mathbf{X}$ within some bounded set $\mathcal{X}$. This notation will be useful and convenient for characterizing the per-stream MSE of (3.13) and the relay's power $P_{\mathrm{R},m}\left(\Delta\mathbf{H}_m\right)$ for different types of CSI errors in a unified form in our subsequent analysis.

#### Min-Max Problem

For notational convenience, we define $\mathbf{F} \triangleq (\mathbf{F}_1, \cdots, \mathbf{F}_K)$, $\mathbf{W} \triangleq (\mathbf{W}_1, \cdots, \mathbf{W}_M)$, and $\mathbf{U} \triangleq (\mathbf{U}_1, \cdots, \mathbf{U}_K)$, which collect the corresponding design variables. In this problem, we jointly design $\{\mathbf{F}, \mathbf{W}, \mathbf{U}\}$ with the goal of minimizing the maximum per-stream MSE subject to the source and relay power constraints. This problem pertains to the design of energy-efficient relay networks, where there is a strict constraint on the affordable power consumption. Based on the notation in (3.14), it can be expressed in the following unified form, denoted

as $\mathcal{M}(P_{\mathtt{R}})$:

$$\min_{\mathbf{F},\mathbf{W},\mathbf{U}} \quad \max_{\forall k\in\mathcal{K}, l\in\mathcal{D}_k} \kappa_{k,l}\mathcal{U}\left\{\varepsilon_{k,l}\left(\Delta\mathbf{H},\Delta\mathbf{G}_k\right)\right\} \tag{3.15a}$$

$$\text{s.t.} \quad \mathcal{U}\left\{P_{\mathtt{R},m}\left(\Delta\mathbf{H}_m\right)\right\} \le \rho_m P_{\mathtt{R}},\ \forall m\in\mathcal{M} \tag{3.15b}$$

$$\mathrm{Tr}\left(\mathbf{F}_k^H\mathbf{F}_k\right) \le P_{\mathtt{S},k}^{\max},\ \forall k\in\mathcal{K}, \tag{3.15c}$$

where $\{\kappa_{k,l}>0:\forall k\in\mathcal{K}, l\in\mathcal{D}_k\}$ is a set of weights assigned to the different data streams for maintaining fairness amongst them, $P_{\mathtt{R}}$ is the common maximum affordable transmit power of all the relays and $\{\rho_m>0:\forall m\in\mathcal{M}\}$ is a set of coefficients specifying the individual power of each relay.

### QoS Problem

The second strategy, which serves as a complement to the above min-max problem, aims for minimizing the maximum per-relay power, while strictly satisfying the QoS constraints for all the data streams as well as all the source power constraints[3].

Specifically, this problem, denoted as $\mathcal{Q}(\gamma)$, can be formulated as

$$\min_{\mathbf{F},\mathbf{W},\mathbf{U}} \quad \max_{m\in\mathcal{M}} \quad \frac{1}{\rho_m}\mathcal{U}\left\{P_{\mathtt{R},m}(\Delta\mathbf{H}_m)\right\} \tag{3.16a}$$

$$\text{s.t.} \quad \mathcal{U}\left\{\varepsilon_{k,l}\left(\Delta\mathbf{H},\Delta\mathbf{G}_k\right)\right\} \le \frac{\gamma}{\kappa_{k,l}},\ \ \forall k\in\mathcal{K}, l\in\mathcal{D}_k \tag{3.16b}$$

$$\mathrm{Tr}\left(\mathbf{F}_k^H\mathbf{F}_k\right) \le P_{\mathtt{S},k}^{\max},\ \forall k\in\mathcal{K}, \tag{3.16c}$$

where $\gamma$ denotes a common QoS target for all the data streams.

The following remark is of interest.

**Remark 3.1** (On the problem formulation)**.** The major difference between the min-max and QoS problems is that solving the QoS problem is not always feasible. This is because the per-stream MSE imposed by the inter-stream and inter-user interference [c.f. (3.13)] cannot

[3]In fact, the min-max problem $\mathcal{M}(P_{\mathtt{R}})$ and the QoS problem $\mathcal{Q}(\gamma)$ are the so-called *inverse problems*, i.e., we have $\gamma=\mathcal{M}\left[\mathcal{Q}(\gamma)\right]$ and $P_{\mathtt{R}}=\mathcal{Q}\left[\mathcal{M}(P_{\mathtt{R}})\right]$. The proof follows a similar argument to that of Theorem 3 in [19]. However, as shown in the subsequent analysis, the proposed algorithm cannot guarantee finding the global optimum of the design problems. Therefore, monotonic convergence cannot be guaranteed which is formally stated as $P_{\mathtt{R}}\ge P'_{\mathtt{R}} \nRightarrow \mathcal{M}(P_{\mathtt{R}})\le\mathcal{M}(P'_{\mathtt{R}})$ and $\gamma\ge\gamma' \nRightarrow \mathcal{Q}(\gamma)\le\mathcal{Q}(\gamma')$. Due to the lack of the monotonicity, a one-dimensional binary search algorithm is unable to solve $\mathcal{Q}(\gamma)$ via a sequence of $\mathcal{M}(P_{\mathtt{R}})$ evaluations. Consequently, a formal inverse problem definition is not stated in the chapter.

be made arbitrarily small by simply increasing the transmit power. By contrast, solving the min-max problem is always feasible, since it relies on its "best effort" to improve the QoS for all the data streams at a limited power consumption. Both problem formulations are non-convex and in general NP-hard. These issues motivate the pursuit of a tractable, but suboptimal solution to the design problems considered.

## 3.2 Statistically Robust Transceiver Design for the Min-Max Problem

In this section, we propose an algorithmic solution to the min-max problem of (3.15) in the presence of the statistical CSI errors of Section 3.1. The corresponding statistically robust version of (3.15) can be formulated as

$$\min_{\mathbf{F},\mathbf{W},\mathbf{U}} \quad \max_{\forall k\in\mathcal{K}, l\in\mathcal{D}_k} \kappa_{k,l}\overline{\varepsilon}_{k,l} \tag{3.17a}$$

$$\text{s.t.} \quad \overline{P}_{\mathtt{R},m} \leq \rho_m P_{\mathtt{R}},\ \forall m \in \mathcal{M} \tag{3.17b}$$

$$\mathrm{Tr}\left(\mathbf{F}_k^H \mathbf{F}_k\right) \leq P_{\mathtt{S},k}^{\max},\ \forall k \in \mathcal{K}, \tag{3.17c}$$

where we have

$$\begin{aligned} \overline{\varepsilon}_{k,l} &\triangleq \mathbb{E}_{\Delta\mathbf{H},\Delta\mathbf{G}_k}\left\{\varepsilon_{k,l}\left(\Delta\mathbf{H},\Delta\mathbf{G}_k\right)\right\} \\ \overline{P}_{\mathtt{R},m} &\triangleq \mathbb{E}_{\Delta\mathbf{H}_m}\left\{P_{\mathtt{R},m}(\Delta\mathbf{H}_m)\right\}. \end{aligned} \tag{3.18}$$

To further exploit the structure of (3.17), we have to compute the expectations in (3.18), which we refer to as the averaged MSE and relay power, respectively. By exploiting the independence of $\Delta\mathbf{H}_{m,k}$ and $\Delta\mathbf{G}_{k,m}$ in (3.13), the per-stream MSE averaged over the channel

uncertainties can be expanded as

$$\begin{aligned}\bar{\varepsilon}_{k,l} = \mathbf{u}_{k,l}^H \left(\boldsymbol{\mathcal{T}}_{k,k}\boldsymbol{\mathcal{T}}_{k,k}^H + \mathbf{R}_k\right)\mathbf{u}_{k,l} - 2\Re\left\{\mathbf{u}_{k,l}^H \boldsymbol{\mathcal{T}}_{k,k}\,\mathbf{e}_{k,l}\right\} + 1 \\ + \sum_{q=1}^{K}\sum_{m=1}^{M} \underbrace{\mathbb{E}\left\{\mathbf{u}_{k,l}^H \Delta\mathbf{G}_{k,m}\boldsymbol{\mathcal{W}}_{m,q}\mathbf{F}_q\mathbf{F}_q^H\boldsymbol{\mathcal{W}}_{m,q}^H\Delta\mathbf{G}_{k,m}^H\mathbf{u}_{k,l}\right\}}_{\mathcal{I}_1} \\ + \sum_{q=1}^{K}\sum_{m=1}^{M} \underbrace{\mathbb{E}\left\{\mathbf{u}_{k,l}^H \boldsymbol{\mathcal{G}}_{k,m}\Delta\mathbf{H}_{m,q}\mathbf{F}_q\mathbf{F}_q^H\boldsymbol{\Delta}\mathbf{H}_{m,q}^H\boldsymbol{\mathcal{G}}_{k,m}^H\mathbf{u}_{k,l}\right\}}_{\mathcal{I}_2} \\ + \sum_{m=1}^{M}\sigma_{\mathtt{R},m}^2 \underbrace{\mathbb{E}\left\{\mathbf{u}_{k,l}^H\Delta\mathbf{G}_{k,m}\mathbf{W}_m\mathbf{W}_m^H\Delta\mathbf{G}_{k,m}^H\mathbf{u}_{k,l}\right\}}_{\mathcal{I}_3},\end{aligned} \tag{3.19}$$

where we have

$$\mathbf{R}_k = \sum_{q=1,q\neq k}^{K}\boldsymbol{\mathcal{T}}_{k,q}\boldsymbol{\mathcal{T}}_{k,q}^H + \sum_{m=1}^{M}\sigma_{\mathtt{R},m}^2\boldsymbol{\mathcal{G}}_{k,m}\boldsymbol{\mathcal{G}}_{k,m}^H + \sigma_{\mathtt{D},k}^2\mathbf{I}_{d_k}. \tag{3.20}$$

To compute the expectations in (3.19), we use the property of the Kronecker product $\mathtt{vec}(\mathbf{ABC}) = (\mathbf{C}^T \otimes \mathbf{A})\,\mathtt{vec}(\mathbf{B})$ and obtain

$$\mathcal{I}_1 = \mathbb{E}\left\{\mathtt{vec}^H(\Delta\mathbf{G}_{k,m}^H)(\mathbf{u}_{k,l}^* \otimes \boldsymbol{\mathcal{W}}_{m,q}\mathbf{F}_q)(\mathbf{u}_{k,l}^T \otimes \mathbf{F}_q^H\boldsymbol{\mathcal{W}}_{m,q}^H)\,\mathtt{vec}(\Delta\mathbf{G}_{k,m}^H)\right\}. \tag{3.21}$$

Invoking properties of the $\mathrm{Tr}(\cdot)$ operator, the above can be rewritten as

$$\mathcal{I}_1 = \mathrm{Tr}\left((\mathbf{u}_{k,l}^T \otimes \mathbf{F}_q^H\boldsymbol{\mathcal{W}}_{m,q}^H)\,\mathbb{E}\left\{\mathtt{vec}(\Delta\mathbf{G}_{k,m}^H)\,\mathtt{vec}^H(\Delta\mathbf{G}_{k,m}^H)\right\}(\mathbf{u}_{k,l}^* \otimes \boldsymbol{\mathcal{W}}_{m,q}\mathbf{F}_q)\right). \tag{3.22}$$

Based on the Kronecker model (3.10), i.e., $\mathbb{E}\left\{\mathtt{vec}(\Delta\mathbf{G}_{k,m}^H)\,\mathtt{vec}^H(\Delta\mathbf{G}_{k,m}^H)\right\} = \boldsymbol{\Sigma}_{\mathrm{G}_{k,m}} \otimes \boldsymbol{\Psi}_{\mathrm{G}_{k,m}}$, $\mathcal{I}_1$ can further be expressed as

$$\mathcal{I}_1 = \mathrm{Tr}\left((\mathbf{u}_{k,l}^T \otimes \mathbf{F}_q^H\boldsymbol{\mathcal{W}}_{m,q}^H)(\boldsymbol{\Sigma}_{\mathrm{G}_{k,m}} \otimes \boldsymbol{\Psi}_{\mathrm{G}_{k,m}})(\mathbf{u}_{k,l}^* \otimes \boldsymbol{\mathcal{W}}_{m,q}\mathbf{F}_q)\right). \tag{3.23}$$

Finally, using the mixed property of the Kronecker product, i.e, $(\mathbf{A}\otimes\mathbf{B})(\mathbf{C}\otimes\mathbf{D})=(\mathbf{AC}\otimes\mathbf{BD})$, after some simple manipulations we arrive at

$$\begin{aligned}\mathcal{I}_1 &= \mathrm{Tr}\left(\mathbf{u}_{k,l}^T\boldsymbol{\Sigma}_{\mathrm{G}_{k,m}}\mathbf{u}_{k,l}^*\otimes(\mathbf{F}_q^H\boldsymbol{\mathcal{W}}_{m,q}^H\boldsymbol{\mathcal{W}}_{m,q}\mathbf{F}_q)\right)\\ &= (\mathbf{u}_{k,l}^T\boldsymbol{\Sigma}_{\mathrm{G}_{k,m}}\mathbf{u}_{k,l}^*)\,\mathrm{Tr}(\mathbf{F}_q^H\boldsymbol{\mathcal{W}}_{m,q}^H\boldsymbol{\Psi}_{\mathrm{G}_{k,m}}\boldsymbol{\mathcal{W}}_{m,q}\mathbf{F}_q)\\ &= \mathrm{Tr}\left(\boldsymbol{\mathcal{W}}_{m,q}\,\mathbf{F}_q\mathbf{F}_q^H\,\boldsymbol{\mathcal{W}}_{m,q}^H\,\boldsymbol{\Psi}_{\mathrm{G}_{k,m}}\right)\mathbf{u}_{k,l}^H\boldsymbol{\Sigma}_{\mathrm{G}_{k,m}}\mathbf{u}_{k,l}.\end{aligned}\tag{3.24}$$

Similarly, $\mathcal{I}_2$ and $\mathcal{I}_3$ can be simplified to

$$\mathcal{I}_2 = \mathrm{Tr}\left(\mathbf{F}_q\mathbf{F}_q^H\boldsymbol{\Psi}_{\mathrm{H}_{m,q}}\right)\mathbf{u}_{k,l}^H\,\boldsymbol{\mathcal{G}}_{k,m}\,\boldsymbol{\Sigma}_{\mathrm{H}_{m,q}}\,\boldsymbol{\mathcal{G}}_{k,m}^H\,\mathbf{u}_{k,l}\tag{3.25}$$

$$\mathcal{I}_3 = \mathrm{Tr}\left(\mathbf{W}_m\mathbf{W}_m^H\boldsymbol{\Psi}_{\mathrm{G}_{k,m}}\right)\mathbf{u}_{k,l}^H\boldsymbol{\Sigma}_{\mathrm{G}_{k,m}}\mathbf{u}_{k,l}.\tag{3.26}$$

Based on (3.24)–(3.26), the averaged MSE in (3.19) is therefore equivalent to

$$\overline{\varepsilon}_{k,l} = \mathbf{u}_{k,l}^H\left(\boldsymbol{\mathcal{T}}_{k,k}\boldsymbol{\mathcal{T}}_{k,k}^H+\mathbf{R}_k+\boldsymbol{\Omega}_k\right)\mathbf{u}_{k,l}-2\Re\left\{\mathbf{u}_{k,l}^H\boldsymbol{\mathcal{T}}_{k,k}\,\mathbf{e}_{k,l}\right\}+1,\tag{3.27}$$

where

$$\begin{aligned}\boldsymbol{\Omega}_k = &\sum_{q=1}^{K}\sum_{m=1}^{M}\left(\mathrm{Tr}\left(\boldsymbol{\mathcal{W}}_{m,q}\,\mathbf{F}_q\mathbf{F}_q^H\,\boldsymbol{\mathcal{W}}_{m,q}^H\,\boldsymbol{\Psi}_{\mathrm{G}_{k,m}}\right)\boldsymbol{\Sigma}_{\mathrm{G}_{k,m}}+\mathrm{Tr}\left(\mathbf{F}_q\mathbf{F}_q^H\boldsymbol{\Psi}_{\mathrm{H}_{m,q}}\right)\boldsymbol{\mathcal{G}}_{k,m}\,\boldsymbol{\Sigma}_{\mathrm{H}_{m,q}}\,\boldsymbol{\mathcal{G}}_{k,m}^H\right)\\ &+\sum_{m=1}^{M}\sigma_{\mathtt{R},m}^2\,\mathrm{Tr}\left(\mathbf{W}_m\mathbf{W}_m^H\boldsymbol{\Psi}_{\mathrm{G}_{k,m}}\right)\boldsymbol{\Sigma}_{\mathrm{G}_{k,m}}.\end{aligned}\tag{3.28}$$

After careful inspection, it is interesting to find that $\overline{\varepsilon}_{k,l}$ is convex with respect to each block of its variables $\mathbf{F}$, $\mathbf{W}$ and $\mathbf{U}$, although not jointly convex in all the design variables.

The averaged relay power $\overline{P}_{\mathtt{R},m}$ can be derived as

$$\begin{aligned}\overline{P}_{\mathtt{R},m} = &\sum_{k=1}^{K}\left(\mathrm{Tr}\left(\mathbf{F}_k^H\hat{\mathbf{H}}_{m,k}^H\mathbf{W}_m^H\mathbf{W}_m\hat{\mathbf{H}}_{m,k}\mathbf{F}_k\right)+\mathrm{Tr}\left(\mathbf{F}_k\mathbf{F}_k^H\boldsymbol{\Psi}_{\mathrm{H}_{m,k}}\right)\mathrm{Tr}\left(\mathbf{W}_m^H\mathbf{W}_m\boldsymbol{\Sigma}_{\mathrm{H}_{m,k}}\right)\right)\\ &+\sigma_{\mathtt{R},m}^2\,\mathrm{Tr}\left(\mathbf{W}_m\mathbf{W}_m^H\right)\end{aligned}\tag{3.29}$$

and the convexity of $\overline{P}_{\mathtt{R},m}$ in each of $\mathbf{F}$ and $\mathbf{W}$ is immediate.

### 3.2.1 Iterative Joint Transceiver Optimization

It is worthwhile noting that the inner point-wise maximization in (3.17a) preserves the partial convexity of $\overline{\varepsilon}_{k,l}$. Substituting (3.27) and (3.29) back into (3.17), the latter is shown to possess a so-called *block multi-convex* structure [98], which, implies that the problem is convex in each block of variables, although in general not jointly convex in all the variables.

Motivated by the above property, we propose an algorithmic solution for the joint transceiver optimization based on the block coordinate update approach, which updates the three blocks of design variables, one at a time while fixing the values associated with the remaining blocks. In this way, three sub-problems can be derived from (3.17), with each updating $\mathbf{F}$, $\mathbf{W}$ and $\mathbf{U}$, respectively. Each sub-problem can be transformed into a convex one, which is computationally much simpler than directly finding the optimal solution to the original joint problem (if at all possible). Since solving for each block at the current iteration depends on the values of the other blocks gleaned from the previous iteration, this method in effect can be recognized as a joint optimization approach in terms of both the underlying theory [53, 98] and the related applications [25, 28]. We now proceed by analyzing each of these sub-problems.

#### Receive Filter Design

It can be observed in (3.19) that $\overline{\varepsilon}_{k,l}$ in (3.17a) only depends on the corresponding linear vector $\mathbf{u}_{k,l}$, while the constraints (3.17b) and (3.17c) do not involve $\mathbf{u}_{k,l}$. Hence, for a fixed $\mathbf{F}$ and $\mathbf{W}$, the optimal $\mathbf{u}_{k,l}$ can be obtained independently and in parallel for different $(k,l)$ values by equating the following complex gradient to zero:

$$\nabla_{\mathbf{u}_{k,l}^*}\overline{\varepsilon}_{k,l} = \mathbf{0}. \tag{3.30}$$

The resultant optimal solution of (3.30) is the Wiener filter

$$\mathbf{u}_{k,l} = \left(\boldsymbol{\mathcal{T}}_{k,k}\boldsymbol{\mathcal{T}}_{k,k}^H + \mathbf{R}_k + \boldsymbol{\Omega}_k\right)^{-1}\boldsymbol{\mathcal{T}}_{k,k}\,\mathbf{e}_{k,l}. \tag{3.31}$$

#### Source TPC Design

We then solve our problem for the TPC $\mathbf{F}$, while keeping $\mathbf{W}$ and $\mathbf{U}$ fixed. For better exposition of our solution, we can rewrite (3.17) after some matrix manipulations, explicitly

in terms of $\mathbf{F}$ as the following:

$$\min_{\mathbf{F}} \quad \max_{\forall k\in\mathcal{K}, l\in\mathcal{D}_k} \kappa_{k,l}\Bigg\{ \sum_{q=1}^{K}\sum_{m=1}^{M}\sum_{n=1}^{M} \mathrm{Tr}\left(\mathbf{F}_q^H \boldsymbol{\mathcal{W}}_{m,q}^H \boldsymbol{\mathcal{U}}_{k,m}^H \mathbf{E}_{k,l} \boldsymbol{\mathcal{U}}_{k,n} \boldsymbol{\mathcal{W}}_{n,q} \mathbf{F}_q\right)$$
$$- \sum_{m=1}^{M} 2\Re\left\{\mathrm{Tr}\left(\mathbf{E}_{k,l}\boldsymbol{\mathcal{U}}_{k,m}\boldsymbol{\mathcal{W}}_{m,k}\mathbf{F}_k\right)\right\} + a_3^{k,l}$$
$$+ \sum_{q=1}^{K}\sum_{m=1}^{M} \mathrm{Tr}\left(\mathbf{F}_q^H \boldsymbol{\mathcal{W}}_{m,k}^H \boldsymbol{\Psi}_{\mathrm{G}_{k,m}} \boldsymbol{\mathcal{W}}_{m,k} \mathbf{F}_q\right) \mathrm{Tr}\left(\mathbf{u}_{k,l}^{\mathbf{H}} \boldsymbol{\Sigma}_{\mathrm{G}_{k,m}} \mathbf{u}_{k,l}\right)$$
$$+ \sum_{q=1}^{K}\sum_{m=1}^{M} \mathrm{Tr}\left(\mathbf{F}_q^H \boldsymbol{\Psi}_{\mathrm{H}_{m,q}} \mathbf{F}_q\right) \mathrm{Tr}\left(\mathbf{u}_{k,l}^H \boldsymbol{\mathcal{G}}_{k,m} \boldsymbol{\Sigma}_{\mathrm{H}_{m,q}} \boldsymbol{\mathcal{G}}_{k,m}^H \mathbf{u}_{k,l}\right)\Bigg\} \tag{3.32a}$$

$$\text{s.t.} \quad \sum_{k=1}^{K} \mathrm{Tr}\left(\mathbf{F}_k^H \left(\hat{\mathbf{H}}_{m,k}^H \mathbf{W}_m^H \mathbf{W}_m \hat{\mathbf{H}}_{m,k} + \mathrm{Tr}\left(\mathbf{W}_m^H \mathbf{W}_m \boldsymbol{\Sigma}_{\mathrm{H}_{m,k}}\right) \boldsymbol{\Psi}_{\mathrm{H}_{m,k}}\right) \mathbf{F}_k\right) \leq \eta_{\mathtt{R},m}, \ \forall m \in \mathcal{M} \tag{3.32b}$$

$$\mathrm{Tr}\left(\mathbf{F}_k^{\mathbf{H}} \mathbf{F}_k\right) \leq P_{\mathtt{S},k}^{\max}, \ \forall k \in \mathcal{K}, \tag{3.32c}$$

where $\mathbf{E}_{k,l} \triangleq \mathbf{e}_{k,l}\mathbf{e}_{k,l}^T$, $\eta_{\mathtt{R},m} \triangleq \rho_m P_{\mathtt{R}} - \sigma_{\mathtt{R},m}^2 \mathrm{Tr}\left(\mathbf{W}_m \mathbf{W}_m^H\right)$ and

$$a_3^{k,l} \triangleq \mathbf{u}_{k,l}^H \Big[ \sum_{m=1}^{M} \sigma_{\mathtt{R},m}^2 \Big( \mathrm{Tr}\left(\mathbf{W}_m \mathbf{W}_m^H \boldsymbol{\Psi}_{\mathrm{G}_{k,m}}\right) \boldsymbol{\Sigma}_{\mathrm{G}_{k,m}} + \boldsymbol{\mathcal{G}}_{k,m} \boldsymbol{\mathcal{G}}_{k,m}^H \Big) + \sigma_{\mathtt{D},k}^2 \mathbf{I}_{N_{\mathtt{D},k}} \Big] \mathbf{u}_{k,l} + 1. \tag{3.33}$$

The solution to the problem (3.32) is not straightforward, hence we transform it into a more tractable form. To this end, we introduce the new variables of $\mathbf{f}_k \triangleq \mathtt{vec}\left(\mathbf{F}_k\right) \in \mathbb{C}^{N_{\mathtt{S},k} d_k \times 1}$ $\forall k \in \mathcal{K}$ and define the following quantities which are independent of $\mathbf{f}_k$ $\forall k \in \mathcal{K}$

$$\mathbf{A}_{1,q}^{k,l} \triangleq \sum_{m=1}^{M} \mathbf{I}_{d_k} \otimes \Big( \sum_{n=1}^{M} \boldsymbol{\mathcal{W}}_{m,q}^H \boldsymbol{\mathcal{U}}_{k,m}^H \mathbf{E}_{k,l} \boldsymbol{\mathcal{U}}_{k,n} \boldsymbol{\mathcal{W}}_{n,q}$$
$$+ \mathrm{Tr}\left(\mathbf{u}_{k,l}^H \boldsymbol{\Sigma}_{\mathrm{G}_{k,m}} \mathbf{u}_{k,l}\right) \boldsymbol{\mathcal{W}}_{m,k}^H \boldsymbol{\Psi}_{\mathrm{G}_{k,m}} \boldsymbol{\mathcal{W}}_{m,k}$$
$$+ \ \mathrm{Tr}\left(\mathbf{u}_{k,l}^H \boldsymbol{\mathcal{G}}_{k,m} \boldsymbol{\Sigma}_{\mathrm{H}_{m,q}} \boldsymbol{\mathcal{G}}_{k,m}^H \mathbf{u}_{k,l}\right) \boldsymbol{\Psi}_{\mathrm{H}_{m,q}} \Big) \tag{3.34}$$

$$\mathbf{a}_2^{k,l} = \mathtt{vec}\left( \sum_{m=1}^{M} \boldsymbol{\mathcal{W}}_{m,k}^H \boldsymbol{\mathcal{U}}_{k,m}^H \mathbf{E}_{k,l} \right) \tag{3.35}$$

$$\mathbf{A}_{4,k}^m = \mathbf{I}_{d_k} \otimes \left( \boldsymbol{\mathcal{W}}_{m,k}^H \boldsymbol{\mathcal{W}}_{m,k} + \mathrm{Tr}\left(\mathbf{W}_m^H \mathbf{W}_m \boldsymbol{\Sigma}_{\mathrm{H}_{m,k}}\right) \boldsymbol{\Psi}_{\mathrm{H}_{m,k}} \right). \tag{3.36}$$

It may be readily verified that $\mathbf{A}_{1,q}^{k,l}$ and $\mathbf{A}_{4,k}^{m}$ are positive definite matrices. Then we invoke the following identities, $\mathrm{Tr}\left(\mathbf{A}^H\mathbf{B}\mathbf{A}\right) = \mathtt{vec}\left(\mathbf{A}\right)^H\left(\mathbf{I}\otimes\mathbf{B}\right)\mathtt{vec}\left(\mathbf{A}\right)$ and $\mathrm{Tr}\left(\mathbf{A}^H\mathbf{B}\right) = \mathtt{vec}\left(\mathbf{B}\right)^H\mathtt{vec}\left(\mathbf{A}\right)$, for transforming both the objective (3.32a) and the constraints (3.32b)–(3.32c) into quadratic expressions of $\mathbf{f}_k$, and finally reach the following equivalent formulation:

$$\min_{\mathbf{f}_1,\cdots,\mathbf{f}_K,t} \quad t \tag{3.37a}$$

$$\text{s.t.} \quad \sum_{q=1}^{K}\mathbf{f}_q^H\mathbf{A}_{1,q}^{k,l}\mathbf{f}_q - 2\Re\left\{\mathbf{f}_k^H\mathbf{a}_2^{k,l}\right\} + a_3^{k,l} \le \frac{t}{\kappa_{k,l}}, \quad \forall k\in\mathcal{K}, l\in\mathcal{D}_k \tag{3.37b}$$

$$\sum_{k=1}^{K}\mathbf{f}_k^H\mathbf{A}_{4,k}^{m}\mathbf{f}_k \le \eta_{\mathtt{R},m}, \quad \forall m\in\mathcal{M} \tag{3.37c}$$

$$\mathbf{f}_k^H\mathbf{f}_k \le P_{\mathtt{S},k}^{\max}, \quad \forall k\in\mathcal{K}, \tag{3.37d}$$

where $t$ is an auxiliary variable. Problem (3.37) is a convex separable inhomogeneous QCLP [54]. This class of optimization problems can be handled by the recently developed parser-solvers, such as `CVX` [100] where the built-in parser is capable of verifying the convexity of the optimization problem (in user-specified forms) and then, of automatically transforming it into a standard form; the latter may then be forwarded to external optimization solvers, such as `SeduMi` [101] and `MOSEK` [102]. To gain further insights into this procedure, we show in Appendix A.1 that the problem (3.37) can be equivalently transformed into a standard SOCP that is directly solvable by a generic external optimization solver based on the interior-point method. Therefore, the SOCP form bypasses the tedious translation by the parser-solvers for every problem instance in real-time computation.

### Relay AF Matrix Design

In order to solve for the relay AF matrices, we follow a similar procedure to that used for the source TPC design. However, here we introduce a new variable, which vertically concatenates all the vectorized relay AF matrices, yielding:

$$\mathbf{w} \triangleq \begin{bmatrix}\mathbf{w}_1\\ \vdots \\ \mathbf{w}_M\end{bmatrix} \triangleq \begin{bmatrix}\mathtt{vec}\left(\mathbf{W}_1\right)\\ \vdots \\ \mathtt{vec}\left(\mathbf{W}_M\right)\end{bmatrix} \tag{3.38}$$

along with the following quantities, which are independent of $\mathbf{w}$:

$$\left[\mathbf{B}_1^{k,l}\right]_{m,n} = \sum_{q=1}^{K} \left[ \left( \boldsymbol{\mathcal{H}}_{m,q}^{*} \boldsymbol{\mathcal{H}}_{n,q}^{T} \right) \otimes \left( \boldsymbol{\mathcal{U}}_{k,m}^{H} \mathbf{E}_{k,l} \boldsymbol{\mathcal{U}}_{k,n} \right) \right] \tag{3.39}$$

$$\mathbf{b}_{2,m}^{k,l} \triangleq \mathtt{vec}\left( \boldsymbol{\mathcal{U}}_{k,m}^{H} \mathbf{E}_{k,l} \boldsymbol{\mathcal{H}}_{m,k}^{H} \right) \tag{3.40}$$

$$\begin{aligned} \mathbf{B}_{3,m}^{k,l} \triangleq \sum_{q=1}^{K} \Big[ & \operatorname{Tr}\left( \mathbf{u}_{k,l}^{H} \boldsymbol{\Sigma}_{\mathrm{G}_{k,m}} \mathbf{u}_{k,l} \right) \boldsymbol{\mathcal{H}}_{m,q}^{*} \boldsymbol{\mathcal{H}}_{m,q}^{T} \otimes \boldsymbol{\Psi}_{\mathrm{G}_{k,m}} + \operatorname{Tr}\left( \mathbf{F}_q^{H} \boldsymbol{\Psi}_{\mathrm{H}_{m,q}} \mathbf{F}_q \right) \boldsymbol{\Sigma}_{\mathrm{H}_{m,q}}^{T} \otimes \boldsymbol{\mathcal{U}}_{k,m}^{H} \mathbf{E}_{k,l} \boldsymbol{\mathcal{U}}_{k,m} \Big] \\ & + \sigma_{\mathtt{R},m}^2 \operatorname{Tr}\left( \mathbf{u}_{k,l}^{H} \boldsymbol{\Sigma}_{\mathrm{G}_{k,m}} \mathbf{u}_{k,l} \right) \mathbf{I}_{N_{\mathtt{R},m}} \otimes \boldsymbol{\Psi}_{\mathrm{G}_{k,m}} + \sigma_{\mathtt{R},m}^2 \mathbf{I}_{N_{\mathtt{R},m}} \otimes \left( \boldsymbol{\mathcal{U}}_{k,m}^{H} \mathbf{E}_{k,l} \boldsymbol{\mathcal{U}}_{k,m} \right) \end{aligned} \tag{3.41}$$

$$b_4^{k,l} \triangleq \sigma_{\mathtt{D},k}^2 \left\| \mathbf{u}_{k,l} \right\|^2 + 1 \tag{3.42}$$

$$\mathbf{B}_{5,m} \triangleq \left[ \sigma_{\mathtt{R},m}^2 \mathbf{I}_{N_{\mathtt{R},m}} + \sum_{k=1}^{K} \left( \boldsymbol{\mathcal{H}}_{m,k}^{*} \boldsymbol{\mathcal{H}}_{m,k}^{T} + \operatorname{Tr}\left( \mathbf{F}_k \mathbf{F}_k^{H} \boldsymbol{\Psi}_{\mathrm{H}_{m,k}} \right) \boldsymbol{\Sigma}_{\mathrm{H}_{m,k}}^{T} \right) \right] \otimes \mathbf{I}_{N_{\mathtt{R},m}} \tag{3.43}$$

where $\mathbf{B}_1^{k,l}$ is a block matrix with its $(m,n)^{\text{th}}$ block defined above. Then, using the identities $\operatorname{Tr}\left(\mathbf{A}^H \mathbf{B} \mathbf{C} \mathbf{D}^H\right) = \mathtt{vec}\left(\mathbf{A}\right)^H \left(\mathbf{D}^T \otimes \mathbf{B}\right) \mathtt{vec}\left(\mathbf{C}\right)$, $\operatorname{Tr}\left(\mathbf{A}^H \mathbf{B} \mathbf{A}\right) = \mathtt{vec}\left(\mathbf{A}\right)^H \left(\mathbf{I} \otimes \mathbf{B}\right) \mathtt{vec}\left(\mathbf{A}\right)$ and $\operatorname{Tr}\left(\mathbf{A}^H \mathbf{B}\right) = \mathtt{vec}\left(\mathbf{B}\right)^H \mathtt{vec}\left(\mathbf{A}\right)$, we can formulate the following optimization problem:

$$\min_{\mathbf{w},t} \quad t \tag{3.44a}$$

$$\text{s.t.} \quad \mathbf{w}^H \mathbf{B}_1^{k,l} \mathbf{w} - \sum_{m=1}^{M} 2\Re\left\{ \mathbf{w}_m^H \mathbf{b}_{2,m}^{k,l} \right\} + \sum_{m=1}^{M} \mathbf{w}_m^H \mathbf{B}_{3,m}^{k,l} \mathbf{w}_m + b_4^{k,l} \leq \frac{t}{\kappa_{k,l}}, \ \forall l \in \mathcal{D}_k, k \in \mathcal{K} \tag{3.44b}$$

$$\mathbf{w}_m^H \mathbf{B}_{5,m} \mathbf{w}_m \leq \rho_m P_{\mathtt{R}}, \ \forall m \in \mathcal{M}. \tag{3.44c}$$

It may be readily shown that $\mathbf{B}_1^{k,l}$, $\mathbf{B}_{3,m}^{k,l}$, and $\mathbf{B}_{5,m}$ are all positive definite matrices and that (3.44) is also a convex separable inhomogeneous QCLP. Using a similar approach to the one derived in Appendix A.1, the SOCP formulation of (3.44) can readily be obtained. The details of the transformation are therefore omitted for brevity.

### 3.2.2 Algorithm and Properties

We assume that there exists a central processing node, which, upon collecting the channel estimates $\{\hat{\mathbf{H}}_{m,k}, \hat{\mathbf{G}}_{k,m}, \forall m \in \mathcal{M}, k \in \mathcal{K}\}$, and the covariance matrices of the CSI errors

$\left\{\mathbf{\Sigma}_{\mathrm{H}_{m,k}}, \mathbf{\Sigma}_{\mathrm{G}_{k,m}}, \mathbf{\Psi}_{\mathrm{H}_{m,k}}, \mathbf{\Psi}_{\mathrm{G}_{k,m}}, \forall m \in \mathcal{M}, k \in \mathcal{K}\right\}$, optimizes all the design variables, and sends them back to the corresponding nodes. The iterative procedure listed in Algorithm 3.1 therefore should be implemented in a centralized manner, where $\left\{\mathbf{F}^{(i)}, \mathbf{W}^{(i)}, \mathbf{U}^{(i)}\right\}$ and $t^{(i)}$ represent the set of design variables and the objective value in (3.17a), respectively, at the $i^{\text{th}}$ iteration. A simple termination criterion can be $\left|t^{(i)} - t^{(i-1)}\right| < \epsilon$, where $\epsilon > 0$ is a predefined threshold. Below we shall analyze both the convergence properties and the complexity of the proposed algorithm.

**Algorithm 3.1** Iterative Algorithm for Statistically Robust Min-Max Problem

1: Initialization: Set the iteration index $i = 0$, $\mathbf{F}_k^{(0)} = \sqrt{P_{\mathrm{S},k}^{\max}}\mathbf{I}_{N_{\mathrm{S},k} \times d_k}$, $\forall k \in \mathcal{K}$ and $\mathbf{W}_m^{(0)} = \sqrt{\frac{\rho_m P_{\mathrm{R}}}{\mathrm{Tr}(\mathbf{B}_{5,m})}}\mathbf{I}_{N_{\mathrm{R},m}}$, $\forall m \in \mathcal{M}$
2: **repeat**
3: Compute $\mathbf{u}_{k,l}^{(i+1)}$ $\forall k \in \mathcal{K}$, $l \in \mathcal{D}_k$, using the Wiener filter (3.31) in parallel;
4: Compute $\mathbf{F}_k^{(i+1)}$ $\forall k \in \mathcal{K}$ by solving the SOCP (A.1);
5: Compute $\mathbf{W}_m^{(i+1)}$ $\forall m \in \mathcal{M}$ by solving the SOCP (3.44);
6: $i \leftarrow i + 1$;
7: **until** $\left|t^{(i)} - t^{(i-1)}\right| < \epsilon$

### Convergence

Given a feasible initialization for Algorithm 3.1, the solution to each subsequent subproblem is globally optimal. As a result, the sequence of the objective values in (3.17a) is monotonically non-increasing, as the iteration index $i$ increases. Since the maximum per-stream MSE is bounded from below (at least) by zero, the sequence of the objective values must converge by invoking the monotonic convergence theorem. It is worth noting that the convergence of the objective function does not necessarily imply global convergence of the obtained solution sequence, i.e., $\{\mathbf{U}^{(i)}, \mathbf{W}^{(i)}, \mathbf{F}^{(i)}\}$. In fact, the convergence of the solution sequence generated by the block coordinate descent type algorithm has been studied in the literature, see, e.g., [31–34, 98]. It has been shown that under some mild conditions, every limit point of the algorithm is either a stationary point or a *Nash* point. We study the convergence of a similar iterative algorithm in detail in Chapter 4, Section 4.2.2.

### Complexity

When the number of antennas at the sources and relays, i.e., $N_{\mathtt{S},k}$ and $N_{\mathtt{R},m}$, have the same order of magnitude, the complexity of Algorithm 3.1 is dominated by the SOCP of (A.1) detailed in Appendix A.1, as it involves all the constraints of the original problem (3.17). To simplify the complexity analysis, we assume that $N_{\mathtt{S},k} = N_{\mathtt{S}}$, and $d_k = d$, $\forall k \in \mathcal{K}$. In (A.1), the total number of design variables is $N_{\text{total}} = N_{\mathtt{S}}^2 K + 1 + K^2 d + KM$. The size of the second-order cones (SOCs) in the constraints (A.1b)–(A.1g) is given by $(N_{\mathtt{S}}^2+1)dK(K-1)$, $(N_{\mathtt{S}}^2+1)dK$, $(K+2)dK$, $(N_{\mathtt{S}}^2+1)KM$, $(K+1)M$ and $(N_{\mathtt{S}}^2+1)K$, respectively. Therefore, the total dimension of all the SOCs in these constraints can be shown to be $D_{\text{SOCP}} = \mathcal{O}(N_{\mathtt{S}}^2 dK^2 + N_{\mathtt{S}}^2 MK)$. It has been shown in [95] that problem (A.1) can be solved most efficiently using the primal-dual interior-point method at a worst-case complexity on the order of $\mathcal{O}(N_{\text{total}}^2 D_{\text{SOCP}})$ if no special structure in the problem data is exploited. The computational complexity of Algorithm 3.1 is therefore on the order of $\mathcal{O}(N_{\mathtt{S}}^6)$, $\mathcal{O}(K^6)$ and $\mathcal{O}(M^3)$ in the individual parameters $N_{\mathtt{S}}$, $K$ and $M$, respectively. In practice, however, we find that the matrices $\mathbf{A}_{1,q}^{k,l}$ and $\mathbf{A}_{4,k}^{m}$ in (3.34) and (3.36), respectively exhibit a significant level of sparsity, which allows solving the SOCP more efficiently. In our simulations, we therefore measured the CPU time required for solving (A.1) for different values of $N_{\mathtt{S}}$, $K$ and $M$ (the results are not reported) and found that the orders of complexity obtained empirically are significantly lower than those of the above worst-case analysis. Empirically, we found these to be around $\mathcal{O}(N_{\mathtt{S}}^{1.6})$, $\mathcal{O}(K^{1.7})$ and $\mathcal{O}(M^{1.3})$.

## 3.3 Worst-Case Robust Transceiver Design for the Min-Max Problem

In this section, we consider the joint transceiver design problem under min-max formulation of (3.15) and the norm-bounded CSI error model of Section 3.2). To this end, based on the notation in (3.14), we explicitly rewrite this problem as

$$\min_{\mathbf{F},\mathbf{W},\mathbf{U}} \quad \max_{\substack{\forall k\in\mathcal{K}, l\in\mathcal{D}_k,\\ \forall \Delta\mathbf{H}\in\mathcal{H}, \Delta\mathbf{G}_k\in\mathcal{G}_k}} \kappa_{k,l}\varepsilon_{k,l}\left(\Delta\mathbf{H}, \Delta\mathbf{G}_k\right) \tag{3.45a}$$

$$\text{s.t.} \quad P_{\mathtt{R},m}\left(\Delta\mathbf{H}_m\right) \leq \rho_m P_{\mathtt{R}},\ \forall m \in \mathcal{M}, \Delta\mathbf{H}_m \in \mathcal{H}_m \tag{3.45b}$$

$$\text{Tr}\left(\mathbf{F}_k^{\mathbf{H}}\mathbf{F}_k\right) \leq P_{\mathtt{S},k}^{\max},\ \forall k \in \mathcal{K}, \tag{3.45c}$$

whose epigraph form can be expressed as

$$\min_{\mathbf{F},\mathbf{W},\mathbf{U}} \quad t \tag{3.46a}$$

$$\text{s.t.} \quad \varepsilon_{k,l}\left(\Delta\mathbf{H}, \Delta\mathbf{G}_k\right) \leq \frac{t}{\kappa_{k,l}},\ \forall k \in \mathcal{K}, l \in \mathcal{D}_k, \Delta\mathbf{H} \in \mathcal{H}, \Delta\mathbf{G}_k \in \mathcal{G}_k \tag{3.46b}$$

$$P_{\mathtt{R},m}\left(\Delta\mathbf{H}_m\right) \leq \rho_m P_{\mathtt{R}},\ \forall m \in \mathcal{M}, \Delta\mathbf{H}_m \in \mathcal{H}_m \tag{3.46c}$$

$$\mathrm{Tr}\left(\mathbf{F}_k^{\mathbf{H}}\mathbf{F}_k\right) \leq P_{\mathtt{S},k}^{\max},\ \forall k \in \mathcal{K}, \tag{3.46d}$$

where $t$ is an auxiliary variable. As compared to the statistically robust version of (3.17), problem (3.46) now encounters two major challenges, namely the non-convexity and the *semi-infinite* nature of the constraints (3.46b) and (3.46c), which render the optimization problem mathematically intractable. In what follows, we derive a solution to address these calamities.

### 3.3.1 Iterative Joint Transceiver Optimization

To overcome the first difficulty, we still rely on the iterative block coordinate update approach described in Section 3.2; however, the three resultant sub-problems are semi-infinite due to the continuous but bounded channel uncertainties in (3.46b) and (3.46c). To handle the semi-infiniteness, an equivalent reformulation of these constraints as LMI will be derived by using certain matrix transformation techniques and by exploiting an extended version of the $\mathcal{S}$-lemma of [99]. In turn, such LMI will convert each of the sub-problems into an equivalent SDP [93] efficiently solvable by interior-point methods [103].

#### Receive Filter Design

In this sub-problem, we have to minimize $t$ in (3.46a) with respect to $\mathbf{u}_{k,l}$ subject to the constraint (3.46b). To transform this constraint into an equivalent LMI, the following lemma is presented, which is an extended version of the one in [99].

**Lemma 3.1** (Extension of $\mathcal{S}$-lemma [99])**.** Let $\mathbf{A}(\mathbf{x}) = \mathbf{A}^H(\mathbf{x})$, $\mathbf{\Sigma}(\mathbf{x}) = \mathbf{\Sigma}^H(\mathbf{x})$ and $\{\mathbf{D}_k(\mathbf{x})\}_{k=1}^N$, and $\{\mathbf{B}_k\}_{k=1}^N$ be matrices with appropriate dimensions, where $\mathbf{A}(\mathbf{x})$, $\mathbf{\Sigma}(\mathbf{x})$

and $\{\mathbf{D}_k(\mathbf{x})\}_{k=1}^N$ are affine functions of $\mathbf{x}$. The following semi-infinite matrix inequality

$$\left(\mathbf{A}(\mathbf{x})+\sum_{k=1}^{N}\mathbf{B}_k^H\mathbf{C}_k\mathbf{D}_k(\mathbf{x})\right)\left(\mathbf{A}(\mathbf{x})+\sum_{k=1}^{N}\mathbf{B}_k^H\mathbf{C}_k\mathbf{D}_k(\mathbf{x})\right)^H \preceq \mathbf{\Sigma}(\mathbf{x}) \tag{3.47}$$

holds for all $\|\mathbf{C}_k\|_S \leq \rho_k,\ k=1,\cdots,N$ if and only if there exist nonnegative scalars $\tau_1,\cdots,\tau_N$ satisfying the following:

$$\begin{bmatrix} \mathbf{\Sigma}(\mathbf{x})-\sum\limits_{k=1}^{N}\tau_k\mathbf{B}_k^H\mathbf{B}_k & \mathbf{A}(\mathbf{x}) & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{A}^H(\mathbf{x}) & \mathbf{I} & \rho_1\mathbf{D}_1^H(\mathbf{x}) & \cdots & \rho_N\mathbf{D}_N^H(\mathbf{x}) \\ \mathbf{0} & \rho_1\mathbf{D}_1(\mathbf{x}) & \tau_1\mathbf{I} & \cdots & \mathbf{0} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \rho_N\mathbf{D}_N(\mathbf{x}) & \mathbf{0} & \cdots & \tau_N\mathbf{I} \end{bmatrix} \succeq \mathbf{0}. \tag{3.48}$$

A simplified version of Lemma 3.1, which considers only a single uncertainty block, i.e., $N=1$, can be traced back to [104], while a further related corollary is derived in [99, Proposition 2]. Lemma 3.1 extends this result to the case of multiple uncertainty blocks, i.e., $N>1$; the proof follows similar steps as in [99] and therefore is omitted.

Upon using Lemma 3.1, the QoS constraint (3.46b) can equivalently be reformulated as follows.

**Proposition 3.1.** There exist $\boldsymbol{\tau}_{k,l}^{\mathrm{g}} \in \mathbb{R}_+^{M\times 1}$ and $\boldsymbol{\tau}_{k,l}^{\mathrm{h}} \in \mathbb{R}_+^{KM\times 1}$ capable of ensuring that the semi-infinite constraint (3.46b) is equivalent to the following matrix inequality:

$$\mathbf{Q}_{k,l} \triangleq \begin{bmatrix} \frac{t}{\kappa_{k,l}} - \mathbf{1}^T\boldsymbol{\tau}_{k,l}^{\mathrm{g}} - \mathbf{1}^T\boldsymbol{\tau}_{k,l}^{\mathrm{h}} & \boldsymbol{\theta}_{k,l} & \mathbf{0}_{1\times N_{\mathtt{D},k}N_{\mathtt{R}}} & \mathbf{0}_{1\times N_{\mathtt{S}}N_{\mathtt{R}}} \\ \boldsymbol{\theta}_{k,l}^{\mathbf{H}} & \mathbf{I}_{d+N_{\mathtt{R}}+N_{\mathtt{D},k}} & \overline{\mathbf{\Theta}}_{k,l}^{\mathbf{H}} & \overline{\mathbf{\Phi}}_{k,l}^{\mathbf{H}} \\ \mathbf{0}_{N_{\mathtt{D},k}N_{\mathtt{R}}\times 1} & \overline{\mathbf{\Theta}}_{k,l} & \mathtt{diag}\left(\boldsymbol{\tau}_{k,l}^{\mathrm{g}}\right) * \mathbf{I}_{N_{\mathtt{D},k}N_{\mathtt{R}}} & \mathbf{0}_{N_{\mathtt{D},k}N_{\mathtt{R}}\times N_{\mathtt{S}}N_{\mathtt{R}}} \\ \mathbf{0}_{N_{\mathtt{S}}N_{\mathtt{R}}\times 1} & \overline{\mathbf{\Phi}}_{k,l} & \mathbf{0}_{N_{\mathtt{S}}N_{\mathtt{R}}\times N_{\mathtt{D},k}N_{\mathtt{R}}} & \mathtt{diag}\left(\boldsymbol{\tau}_{k,l}^{\mathrm{h}}\right) * \mathbf{I}_{N_{\mathtt{S}}N_{\mathtt{R}}} \end{bmatrix} \succeq \mathbf{0}, \tag{3.49}$$

where we have $N_{\mathtt{R}} \triangleq \sum_{m=1}^{M} N_{\mathtt{R},m}$, $N_{\mathtt{S}} \triangleq \sum_{k=1}^{K} N_{\mathtt{S},k}$ and the operator $(*)$ denotes the Khatri-Rao product (block-wise Kronecker product) [105]. In (3.49), $\overline{\mathbf{\Theta}}_{k,l}$ and $\overline{\mathbf{\Phi}}_{k,l}$ are defined

as

$$\overline{\boldsymbol{\Theta}}_{k,l} \triangleq \begin{bmatrix} \xi_{k,1}\boldsymbol{\Theta}_1^{k,l} \\ \vdots \\ \xi_{k,M}\boldsymbol{\Theta}_M^{k,l} \end{bmatrix}, \ \overline{\boldsymbol{\Phi}}_{k,l} \triangleq \begin{bmatrix} \eta_{1,1}\boldsymbol{\Phi}_{1,1}^{k,l} \\ \vdots \\ \eta_{M,K}\boldsymbol{\Phi}_{M,K}^{k,l} \end{bmatrix}, \tag{3.50}$$

while $\boldsymbol{\Theta}_{k,l}$, $\boldsymbol{\Phi}_{k,l}$ and $\boldsymbol{\theta}_{k,l}$ are defined in Eq. (A.10) of Appendix A.2.

*Proof:* See Appendix A.2. ■

Using (3.49), the sub-problem formulated for $\mathbf{u}_{k,l}$ can be equivalently recast as

$$\min_{t,\mathbf{u}_{k,l},\boldsymbol{\tau}_{k,l}^{\mathrm{g}},\boldsymbol{\tau}_{k,l}^{\mathrm{h}}} t \quad \text{s.t.} \quad \mathbf{Q}_{k,l} \succeq \mathbf{0}. \tag{3.51}$$

With fixed $\mathbf{F}$ and $\mathbf{W}$, (3.49) depends affinely on the design variables $\left\{t, \mathbf{u}_{k,l}, \boldsymbol{\tau}_{k,l}^{\mathrm{g}}, \boldsymbol{\tau}_{k,l}^{\mathrm{h}}\right\}$. Therefore, (3.51) is a convex SDP of the LMI form [93], which is efficiently solvable by existing optimization tools based on the interior-point method. Since the $\mathbf{u}_{k,l}$ for different values of $(k,l)$ are independent of each other, they can be updated in parallel by solving (3.51) for different $k$ and $l$.

### Source TPC Design

We now have to solve problem (3.46) for $\mathbf{F}$ by fixing $\mathbf{U}$ and $\mathbf{W}$. The solution is formulated in the following proposition.

**Proposition 3.2.** The sub-problem of optimizing the TPCs $\mathbf{F}$ can be formulated as the following SDP:

$$\min_{t,\mathbf{F},\boldsymbol{\tau}_{k,l}^{\mathrm{g}},\boldsymbol{\tau}_{k,l}^{\mathrm{h}},\boldsymbol{\tau}_{m}^{\mathrm{p}}} \quad t \tag{3.52a}$$

$$\text{s.t.} \quad \mathbf{Q}_{k,l} \succeq \mathbf{0}, \ \forall k \in \mathcal{K}, \ l \in \mathcal{D}_k \tag{3.52b}$$

$$\mathbf{P}_m \succeq \mathbf{0}, \ \forall m \in \mathcal{M} \tag{3.52c}$$

$$\begin{bmatrix} P_{\mathrm{s},k}^{\max} & \mathbf{f}_k^H \\ \mathbf{f}_k & \mathbf{I}_{N_{\mathrm{s},k}d_k} \end{bmatrix} \succeq \mathbf{0}, \ \forall k \in \mathcal{K}, \tag{3.52d}$$

where we have

$$\mathbf{P}_m \triangleq \begin{bmatrix} \rho_m P_{\mathtt{R}} - \mathbf{1}^T \boldsymbol{\tau}_m^{\mathrm{p}} & \mathbf{t}_m^{\mathbf{H}} & \mathbf{0}_{1\times N_{\mathtt{S}} N_{\mathtt{R},m}} \\ \mathbf{t}_m & \mathbf{I} & \overline{\mathbf{T}}_m \\ \mathbf{0}_{N_{\mathtt{S}} N_{\mathtt{R},m}\times 1} & \overline{\mathbf{T}}_m^{\mathbf{H}} & \mathtt{diag}\left(\boldsymbol{\tau}_m^{\mathrm{p}}\right) * \mathbf{I} \end{bmatrix} \succeq \mathbf{0} \tag{3.53}$$

with $\boldsymbol{\tau}_m^{\mathrm{p}} \in \mathbb{R}^{K\times 1}$, $\overline{\mathbf{T}}_m(\mathbf{F}) \triangleq \left[\mathbf{T}_{m,1}^T, \cdots, \mathbf{T}_{m,K}^T\right]^T$ and

$$\mathbf{t}_m \triangleq \begin{bmatrix} \mathtt{vec}\left(\mathbf{W}_m \hat{\mathbf{H}}_{m,k} \mathbf{F}_1\right) \\ \vdots \\ \mathtt{vec}\left(\mathbf{W}_m \hat{\mathbf{H}}_{m,K} \mathbf{F}_K\right) \\ \sigma_{\mathtt{R},m}\, \mathtt{vec}\left(\mathbf{W}_m\right) \end{bmatrix}, \tag{3.54}$$

$$\mathbf{T}_{m,k} \triangleq \begin{bmatrix} \mathbf{0}_{\sum_{q=1}^{k-1} d_q N_{\mathtt{R},m} \times N_{\mathtt{S},k} N_{\mathtt{R},m}} \\ \mathbf{F}_k^T \otimes \mathbf{W}_m \\ \mathbf{0}_{\left(\sum_{q=k+1}^{K} d_q N_{\mathtt{R},m} + N_{\mathtt{R},m}^2\right) \times N_{\mathtt{S},k} N_{\mathtt{R},m}} \end{bmatrix}. \tag{3.55}$$

*Proof:* Since $\mathbf{F}$ is involved in all the constraints of the original problem (3.46), below we will transform each of these constraints into tractable forms.

First, note that (3.46b) has already been reformulated as (3.49), which is a trilinear function of $\mathbf{F}$, $\mathbf{W}$ and $\mathbf{U}$. By fixing the values of $\mathbf{W}$ and $\mathbf{U}$, it essentially becomes an LMI in $\mathbf{F}$.

Then, to deal with the semi-infinite constraint of the relay power (3.46c), we can express $P_{\mathtt{R},m}$ as follows based on the definitions in (3.54)

$$P_{\mathtt{R},m} = \left\| \mathbf{t}_m + \sum_{k=1}^{K} \mathbf{T}_{m,k} \mathbf{h}_{m,k} \right\|^2. \tag{3.56}$$

Substituting (3.56) into (3.46c) and again applying Lemma 3.1, (3.46c) can be equivalently recast as the matrix inequality (3.52c), whose left-hand side is bilinear in $\mathbf{W}_m$ and $\mathbf{F}$, which is hence an LMI in $\mathbf{F}$, when $\mathbf{W}_m$ is fixed.

Finally, (3.46d) can be expressed as $\|\mathbf{f}_k\|^2 \le P_{\mathtt{S},k}^{\max}$, which can be equivalently recast as (3.52d) by using the Schur complement rule of [93]. The SDP form (3.52) is then readily obtained. ■

### Relay AF Matrix Design

Since the constraint (3.52d) is independent of the relay AF matrices $\mathbf{W}$, this sub-problem is equivalent to

$$\min_{t,\mathbf{W},\boldsymbol{\tau}^{\mathrm{g}}_{k,l},\boldsymbol{\tau}^{\mathrm{h}}_{k,l},\boldsymbol{\tau}^{\mathrm{p}}_{m}} t \quad \text{s.t.} \quad (3.52\text{b}), (3.52\text{c}). \tag{3.57}$$

The above problem becomes a standard SDP in $\mathbf{W}$ by noting that $\mathbf{Q}_{k,l}$ and $\mathbf{P}_m$ in (3.52b) and (3.52c), respectively, are LMIs in $\mathbf{W}$, provided that the other design variables are kept fixed.

The convergence analysis of the overall iterative algorithm, which solves problems (3.51), (3.52) and (3.57) with the aid of the block coordinate approach, is similar to that in Section 3.2.B and therefore omitted for brevity. One slight difference from Algorithm 3.1 is that we initialize $\mathbf{F}_k^{(0)} = \sqrt{P_{\mathtt{S},k}^{\max}}\mathbf{I}_{N_{\mathtt{S},k}\times d_k}$, $\forall k \in \mathcal{K}$ and $\mathbf{U}_k^{(0)} = \mathbf{I}_{d_k\times N_{\mathtt{S},k}}$, $\forall k \in \mathcal{K}$, and the iterative algorithm will start by solving for the optimal $\mathbf{W}_m^{(1)}$. Solving (3.52) imposes a worst-case complexity on the order of $\mathcal{O}(N_{\text{total}}^2 D_{\text{SDP}})$, where $D_{\text{SDP}}$ represents the total dimensionality of the semidefinite cones in constraints (3.52b)–(3.52d). Similar to the complexity analysis of Algorithm 3.1, we assume that $N_{\mathtt{S},k} = N_{\mathtt{R},m} = N_{\mathtt{D},k} = N$, and $d_k = d, \forall k \in \mathcal{K},\ m \in \mathcal{M}$. In (3.52), the total number of design variables is $N_{\text{total}} = N^2K + 1 + K^2d + KM$. After some simplifications, we find that the computational complexity of the iterative algorithm for the norm-bounded CSI errors is on the order of $\mathcal{O}(N^8)$ and $\mathcal{O}(K^7)$ and $\mathcal{O}(M^5)$, respectively. Comparing the SDP formulation of (3.52) for the norm-bounded CSI errors and the SOCP formulation in (A.1) deduced for the statistical CSI errors, the total dimensionality of (3.52) is seen to be significantly larger than that of (A.1) as given in the complexity discussions in Section 3.2.

## 3.4 Transceiver Design for the QoS Problem

In this section, we turn our attention to the joint transceiver design for the QoS problem (3.16). Following the same approaches as in Sections 3.2 and 3.3, the solution to the QoS problem can also be obtained by adopting the block coordinate update method. Since the derivations of the corresponding subproblems and algorithms are similar to those in Sections 3.2 and 3.3 deduced for the min-max problem, we hereby only present the main results.

### 3.4.1 QoS Problem under Statistical CSI Errors

**Receive Filter Design:**

An optimal $\mathbf{u}_{k,l}$ can be obtained by minimizing $\overline{\varepsilon}_{k,l}\left(\Delta\mathbf{H}, \Delta\mathbf{G}_k\right)$ with respect to $\mathbf{u}_{k,l}$, which yields exactly the same solution as the Wiener filter in (3.31).

**Source TPC Design**

The specific sub-problem of finding the optimal $\mathbf{F}$ can be solved by the following QCLP:

$$\min_{\mathbf{f}_1,\cdots,\mathbf{f}_K,t} \quad t \tag{3.58a}$$

$$\text{s.t.} \quad \sum_{q=1}^{K} \mathbf{f}_q^H \mathbf{A}_{1,q}^{k,l} \mathbf{f}_q - 2\Re\left\{\mathbf{f}_k^H \mathbf{a}_2^{k,l}\right\} + a_3^{k,l} \leq \frac{\gamma}{\kappa_{k,l}}, \quad \forall k \in \mathcal{K},\ l \in \mathcal{D}_k \tag{3.58b}$$

$$\sum_{k=1}^{K} \mathbf{f}_k^H \mathbf{A}_{4,k}^{m} \mathbf{f}_k \leq \eta'_{\mathtt{R},m},\ \forall m \in \mathcal{M} \tag{3.58c}$$

$$\mathbf{f}_k^H \mathbf{f}_k \leq P_{\mathtt{S},k}^{\max},\ \forall k \in \mathcal{K}, \tag{3.58d}$$

where $\eta'_{\mathtt{R},m} \triangleq \rho_m t - \sigma_{\mathtt{R},m}^2 \operatorname{Tr}(\mathbf{W}_m \mathbf{W}_m^H)$.

**Relay AF Matrix Design**

The optimal $\mathbf{W}$ can be found by solving

$$\min_{\mathbf{w},t} \quad t \tag{3.59a}$$

$$\text{s.t.} \quad \mathbf{w}^H \mathbf{B}_1^{k,l} \mathbf{w} - \sum_{m=1}^{M} 2\Re\left\{\mathbf{w}_m^H \mathbf{b}_{2,m}^{k,l}\right\} + \sum_{m=1}^{M} \mathbf{w}_m^H \mathbf{B}_{3,m}^{k,l} \mathbf{w}_m + b_4^{k,l} \leq \frac{\gamma}{\kappa_{k,l}},\ \forall k,l \tag{3.59b}$$

$$\mathbf{w}_m^H \mathbf{B}_{5,m} \mathbf{w}_m \leq \rho_m t,\ \forall m \in \mathcal{M}. \tag{3.59c}$$

### 3.4.2 QoS Problem under Norm-Bounded CSI Errors

*1) Receive Filter Design:* The optimal $\mathbf{u}_{k,l}$ can be obtained from (3.51).

*2) Source TPC Design:* The optimal $\mathbf{F}$ can be obtained as the solution to the following

SDP:

$$\min_{t,\mathbf{F},\boldsymbol{\tau}^{\mathrm{g}}_{k,l},\boldsymbol{\tau}^{\mathrm{h}}_{k,l},\boldsymbol{\tau}^{\mathrm{p}}_{m}} \quad t \tag{3.60a}$$

$$\text{s.t.} \quad \mathbf{Q}'_{k,l} \succeq \mathbf{0},\ \forall k \in \mathcal{K},\ l \in \mathcal{D}_k \tag{3.60b}$$

$$\mathbf{P}'_m \succeq \mathbf{0},\ \forall m \in \mathcal{M} \tag{3.60c}$$

$$\begin{bmatrix} P^{\max}_{\mathtt{S},k} & \mathbf{f}_k^H \\ \mathbf{f}_k & \mathbf{I}_{N_{\mathtt{S},k}d_k} \end{bmatrix} \succeq \mathbf{0},\ \forall k \in \mathcal{K}, \tag{3.60d}$$

where $\mathbf{Q}'_{k,l}$ is obtained from $\mathbf{Q}_{k,l}$ in (3.49) upon replacing $t$ by $\gamma$ in the top-left entry $(1,1)$. Similarly, $\mathbf{P}'_m$ can be obtained by substituting $P_{\mathtt{R}}$ with $t$ in the $(1,1)^{\text{th}}$ entry of $\mathbf{P}_m$ in (3.53).

*3) Relay AF Matrix Design:* The optimal relay AF matrices are obtained by solving

$$\min_{t,\mathbf{W},\boldsymbol{\tau}^{\mathrm{g}}_{k,l},\boldsymbol{\tau}^{\mathrm{h}}_{k,l}} t \quad \text{s.t.} \quad (3.60\text{b}), (3.60\text{c}). \tag{3.61}$$

### 3.4.3 Initial Feasibility Search Algorithm

An important aspect of solving the above QoS problem is to find a feasible initial point. Indeed, it has been observed that if the iterative algorithm is initialized with a random (possibly infeasible) point, the algorithm may fail at the first iteration. Finding a feasible initial point of a non-convex problem, such as our QoS problem (3.16), is in general NP-hard. All these considerations motivate the study of an efficient initial feasibility search algorithm, which finds a reasonably "good" starting point for the QoS problem of (3.16).

Motivated by the "phase I" approach in general optimization theory [93], we formulate the feasibility check problem for the QoS problem as follows

$$\min_{\mathbf{F},\mathbf{W},\mathbf{U}} \quad s \tag{3.62a}$$

$$\text{s.t.} \quad \kappa_{k,l}\mathcal{U}\left\{\varepsilon_{k,l}\left(\Delta\mathbf{H}, \Delta\mathbf{G}_k\right)\right\} \leq s,\ \forall k \in \mathcal{K}, l \in \mathcal{D}_k \tag{3.62b}$$

$$\mathrm{Tr}\left(\mathbf{F}_k^{\mathbf{H}}\mathbf{F}_k\right) \leq P^{\max}_{\mathtt{S},k},\ \forall k \in \mathcal{K}, \tag{3.62c}$$

where $s$ is a slack variable, which represents an abstract measure for the violation of the constraint (3.16b). The above problem can be solved iteratively using the block coordinate approach, until the objective value $s$ converges or the maximum affordable number of

iterations is reached. If at the $(i+1)^{\text{th}}$ iteration, $s^{(i+1)}$ is less than the QoS target $\gamma$, then the procedure successfully finds a feasible initial point; otherwise, we claim that the QoS problem is infeasible. In this case, it is necessary to adjust $\gamma$ or to drop the services of certain users by incorporating an admission control procedure, which, however, is beyond the scope of this thesis.

Interestingly, (3.62) can be reformulated as

$$\min_{\mathbf{F},\mathbf{W},\mathbf{U}} \quad \max_{\forall k\in\mathcal{K},l\in\mathcal{D}_k} \kappa_{k,l}\mathcal{U}\left\{\varepsilon_{k,l}\left(\Delta\mathbf{H},\Delta\mathbf{G}_k\right)\right\} \tag{3.63a}$$

$$\text{s.t.} \quad \mathcal{U}\left\{P_{\mathtt{R},m}\left(\Delta\mathbf{H}_m\right)\right\} \leq \rho_m P_{\mathtt{R}}^{\infty},\ \forall m\in\mathcal{M} \tag{3.63b}$$

$$\text{Tr}\left(\mathbf{F}_k^{\mathbf{H}}\mathbf{F}_k\right) \leq P_{\mathtt{S},k}^{\max},\ \forall k\in\mathcal{K}, \tag{3.63c}$$

where we have $P_{\mathtt{R}}^{\infty}\to\infty$, which is equivalent to removing the constraint on the relay's transmit power. In fact, (3.63) becomes exactly the same as the min-max problem of (3.15) upon setting $P_{\mathtt{R}}=P_{\mathtt{R}}^{\infty}$. We therefore propose an efficient iterative feasibility search algorithm, which is listed as Algorithm 3.2, based on the connection between the feasibility check and the min-max problems.

**Algorithm 3.2** Iterative Initial Feasibility Search Algorithm for the QoS problems

1: **repeat**
2: Solve one cycle of the problem (3.63) and denote the current objective value by $\hat{\gamma}^{(i+1)}$;
3: Verify if $\hat{\gamma}^{(i+1)}\leq\gamma$, and if so, stop the algorithm;
4: $i\leftarrow i+1$;
5: **until** Termination criterion is satisfied, e.g., $\left|\hat{\gamma}^{(i)}-\hat{\gamma}^{(i-1)}\right|\leq\epsilon$; or the maximum allowed number of iteration is reached.

Based on the definition of $\mathcal{U}\{\cdot\}$ in (3.14), Algorithm 3.2 is applicable to the QoS problems associated with both types of CSI errors considered. Furthermore, Algorithm 3.2 indeed provides a feasible initial point for the QoS problem if it exists. Otherwise, it provides a certificate of infeasibility if $\hat{\gamma}^{(i+1)}>\gamma$ after the algorithm terminates. Then the QoS problem is deemed infeasible in this case and the admission control procedure may deny the access of certain users.

## 3.5 Simulation Experiments and Discussions

This section presents our Monte Carlo simulation results for verifying the resilience of the proposed transceiver optimization algorithms against CSI errors. In all simulations we assume that there are $K = 2$ S-D pairs, which communicate with the assistance of $M = 2$ relays. Each node is equipped with $N_{\mathtt{S},k} = N_{\mathtt{R},m} = N_{\mathtt{D},k} = 3$ antennas, $\forall k \in \mathcal{K}, m \in \mathcal{M}$. Each source transmits 2 independent quadrature phase-shift keying (QPSK) modulated data streams to its corresponding destination, i.e., $d_k = 2,\ \forall k \in \mathcal{K}$. Equal noise variances of $\sigma^2_{\mathtt{D},k} = \sigma^2_{\mathtt{R},m}$ are assumed. The maximum source and relay transmit power is normalized to one, i.e., we have $P^{\max}_{\mathtt{S},k} = 1, \forall k \in \mathcal{K}$ and $\rho_m P_{\mathtt{R}} = 1, \forall m \in \mathcal{M}$. Equal weights of $\kappa_{k,l}$ are assigned to the different data streams, unless otherwise stated. The channels are assumed to be flat-fading, with the coefficients given by i.i.d. zero-mean unit-variance complex Gaussian random variables. The signal-to-noise ratios (SNRs) at the relays and the destinations are defined as $\mathrm{SNR}_{\mathtt{R},m} \triangleq \frac{P^{\max}_{\mathtt{S}}}{N_{\mathtt{R},m}\sigma^2_{\mathtt{R},m}}$ and $\mathrm{SNR}_{\mathtt{D},k} \triangleq \frac{P^{\max}_{\mathtt{R}}}{N_{\mathtt{D},k}\sigma^2_{\mathtt{D},k}}$,respectively. The optimization solver MOSEK [102] is used for solving each optimization problem.

### 3.5.1 Performance Evaluation under Statistical CSI Errors

We first evaluate the performance of the iterative algorithm proposed in Section 3.2 under statistical CSI errors. The channel correlation matrices in (3.9) and (3.10) are obtained by the widely-employed exponential model of [61]. Specifically, their entries are given by $\left[\mathbf{\Sigma}_{\mathrm{H}_{m,k}}\right]_{i,j} = \left[\mathbf{\Sigma}_{\mathrm{G}_{k,m}}\right]_{i,j} = \alpha^{|i-j|}$ and $\left[\mathbf{\Psi}_{\mathrm{H}_{m,k}}\right]_{i,j} = \left[\mathbf{\Psi}_{\mathrm{G}_{k,m}}\right]_{i,j} = \sigma^2_e \beta^{|i-j|},\ i,j \in \{1,2,3\}$, where $\alpha$ and $\beta$ are the correlation coefficients, and $\sigma^2_e$ denotes the variance of the CSI errors. The available channel estimates $\hat{\mathbf{H}}_{m,k}$ and $\hat{\mathbf{G}}_{k,m}$ are generated according to $\hat{\mathbf{H}}_{m,k} \sim \mathcal{CN}\left(\mathbf{0}_{N_{\mathtt{R},m}\times N_{\mathtt{S},k}}, \frac{1-\sigma^2_e}{\sigma^2_e}\mathbf{\Sigma}_{\mathrm{H}_{m,k}} \otimes \mathbf{\Psi}^T_{\mathrm{H}_{m,k}}\right)$ and $\hat{\mathbf{G}}_{k,m} \sim \mathcal{CN}\left(\mathbf{0}_{N_{\mathtt{D},k}\times N_{\mathtt{R},m}}, \frac{1-\sigma^2_e}{\sigma^2_e}\mathbf{\Sigma}_{\mathrm{G}_{k,m}} \otimes \mathbf{\Psi}^T_{\mathrm{G}_{k,m}}\right)$, respectively, such that the entries of the true channel matrices have unit variances. We compare the robust transceiver design proposed in Algorithm 3.1 to the: (i) non-robust design, which differs from the robust design in that it assumes $\mathbf{\Sigma}_{\mathrm{H}_{m,k}} = \mathbf{\Sigma}_{\mathrm{G}_{k,m}} = \mathbf{0}$ and $\mathbf{\Psi}_{\mathrm{H}_{m,k}} = \mathbf{\Psi}_{\mathrm{G}_{k,m}} = \mathbf{0}$, i.e., it neglects the effects of the CSI errors; (ii) perfect CSI case, where the true channel matrices $\mathbf{H}_{m,k}$ and $\mathbf{G}_{k,m}$ are used instead of the estimates $\hat{\mathbf{H}}_{m,k}$ and $\hat{\mathbf{G}}_{k,m}$ in Algorithm 3.1 and there are no CSI errors, i.e., we have $\mathbf{\Sigma}_{\mathrm{H}_{m,k}} = \mathbf{\Sigma}_{\mathrm{G}_{k,m}} = \mathbf{0}$ and $\mathbf{\Psi}_{\mathrm{H}_{m,k}} = \mathbf{\Psi}_{\mathrm{G}_{k,m}} = \mathbf{0}$. The curves labeled "Optimal MSE" correspond to the value of the objective function in (3.17a) after optimization by Algorithm 3.1. In all the simulation figures, the MSEs of the different approaches are calculated by averaging the squared error

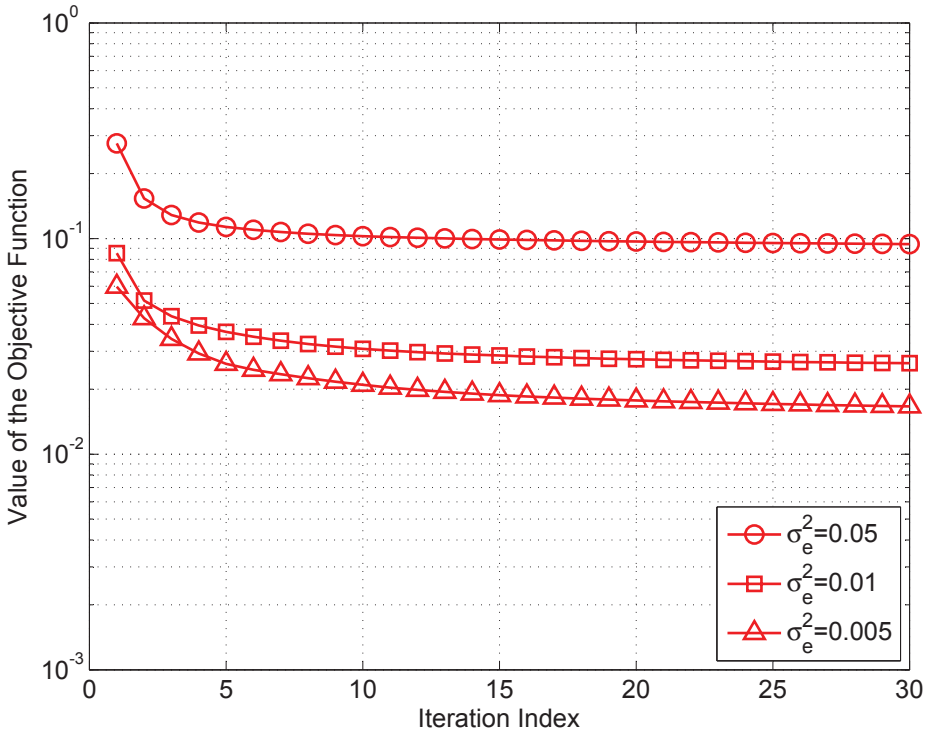
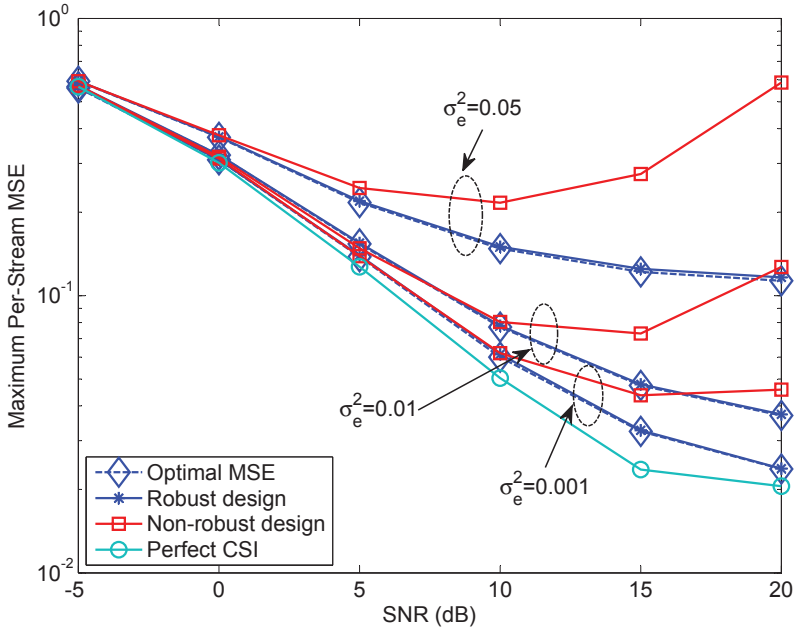



**Fig. 3.2** Convergence behavior of the proposed iterative algorithm with statistical CSI errors.

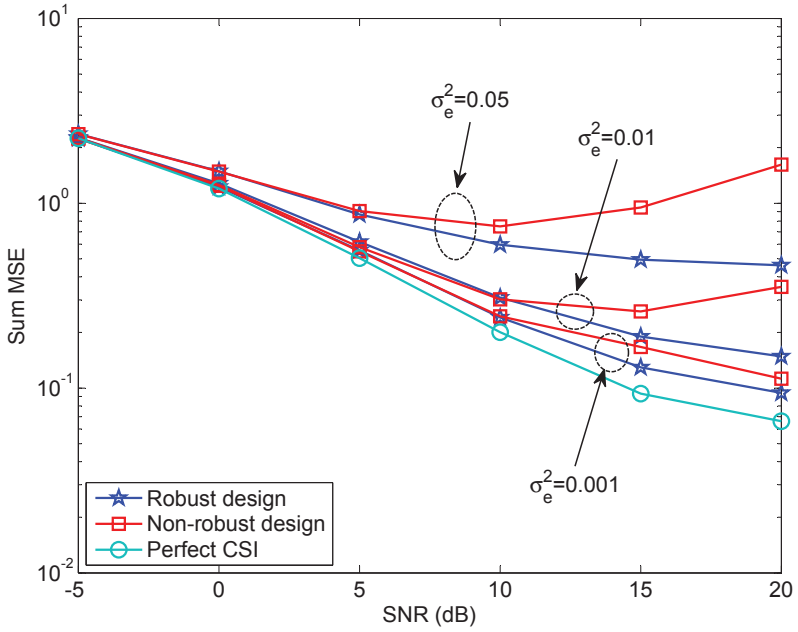
between the transmitted and estimated experimental data symbols over 1000 independent CSI error realizations and 10000 QPSK symbols for each realization.

As a prelude to the presentation of our main simulation results below, the convergence behavior of Algorithm 3.1 is presented for different CSI error variances. It can be observed in Fig. 3.2 that in all cases, the proposed Algorithm 3.1 can converge within a reasonable number of iterations. Therefore, in our experimental work we set the number of iterations to a fixed value of 5 and the resultant performance gains will be discussed below.

**Experiment A.1 (MSE/BER performance)**: In Fig. 3.3 (a), the maximum per-stream MSE among all the data streams is shown as a function of the SNR for different values of CSI error variance. It is observed that the proposed robust design approach achieves a better resilience against the CSI errors, than the non-robust design approach. The performance gains become more evident in the median to high SNR range. For the non-robust design, degradations are observed because the MSE obtained at high SNRs is




(a)




(b)

**Fig. 3.3** MSE performance of different design approaches versus SNR: (a) Maximum per-stream MSE; (b) Sum MSE ($\mathrm{SNR}_{\mathrm{R},m} = \mathrm{SNR}_{\mathrm{D},k} = \mathrm{SNR}$, $\alpha = \beta = 0.5$).

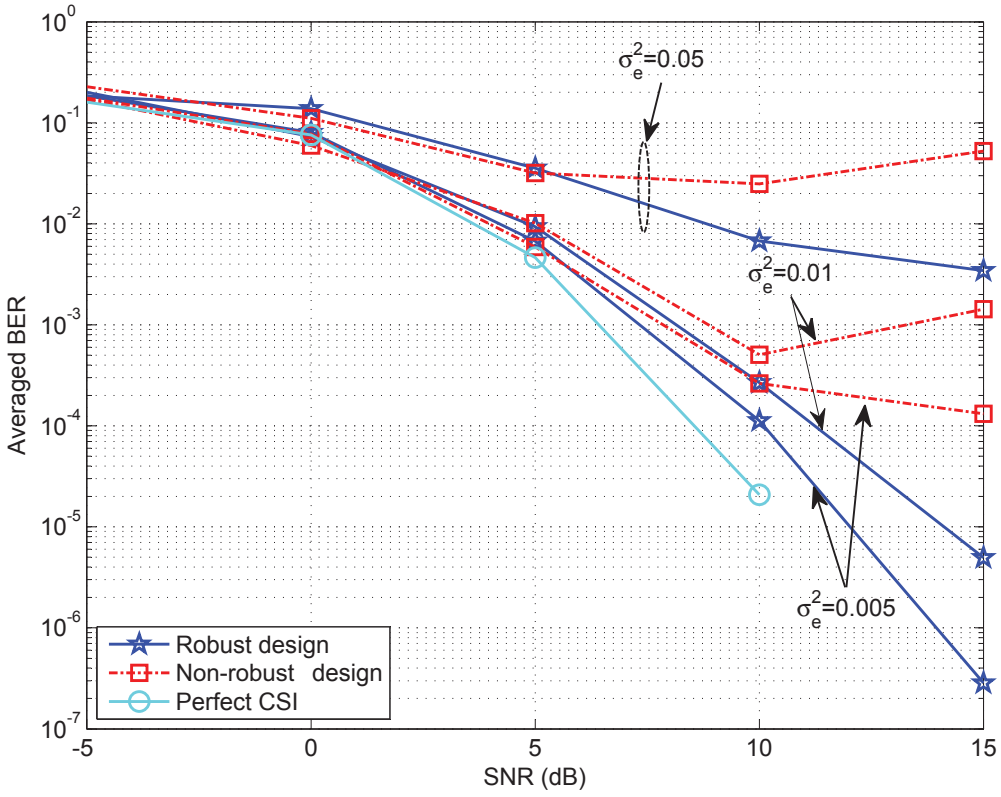



**Fig. 3.4** BER performance of different design approaches versus SNR

dominated by the interference, rather than by the noise. Therefore, the relays are confined to relatively low transmit powers in order to control the interference. This, in turn, leads to a performance degradation imposed by the CSI errors. In contrast, the proposed robust design is capable of compensating for the extra interference imposed by the CSI errors, thereby demonstrating its superiority over its non-robust counterpart. Furthermore, we observe that the "Optimal MSE" and our simulation results tally well, which justifies the approximations invoked in calculating the per-stream MSE in (3.13). In addition to the per-stream performance, the overall system performance [4] quantified in terms of the sum MSE of different approaches is examined in Fig. 3.3 (b), where a similar trend to that of Fig. 3.3 (a) can be observed.

[4]Please note that the objective of portraying the sum MSE performance is to validate whether the proposed robust design approach can also achieve a performance gain over the non-robust approach in terms of its overall performance. In fact, the sum MSE performance can be optimized by solving a design problem with the sum MSE being the objective function.

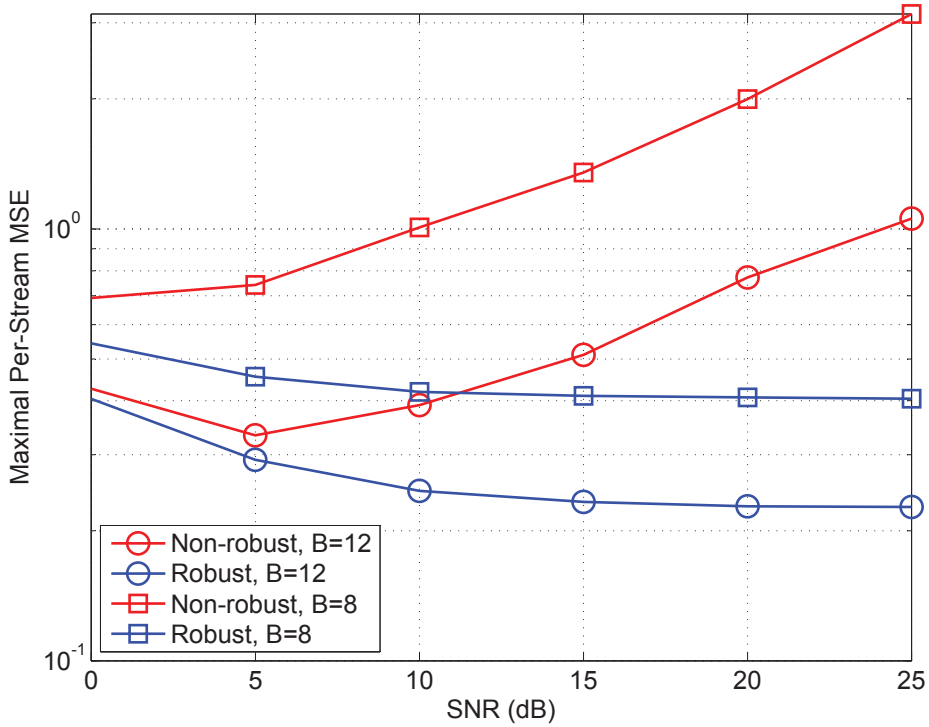



**Fig. 3.5** Per-stream MSE performance with the optimized codebook based on the GLA-VQ. ($B = 8$ corresponds to $\sigma_e^2 = 0.334$ and $B = 12$ corresponds to $\sigma_e^2 = 0.175$.)

The BER performance of different design methods is compared in Fig. 3.4 for different values of the CSI error variances. It is seen that similar to the results in Fig. 3.3, the proposed robust method significantly outperforms the non-robust one. It also becomes evident that the MSE and BER performance are positively correlated, i.e., a lower MSE leads to a lower BER.

The MSE performance associated with a limited number of feedback bits is also studied. To this end, we assume that each user is equipped with a codebook that is optimized using the generalized Lloyd algorithm of vector quantization (GLA-VQ) [63]. Each user then quantizes the channel vector and the corresponding codebook index is fed back to the central processing unit. The results presented in Fig. 3.5 illustrate that the proposed algorithm significantly outperformed the non-robust one for the different number of quantization bits considered.

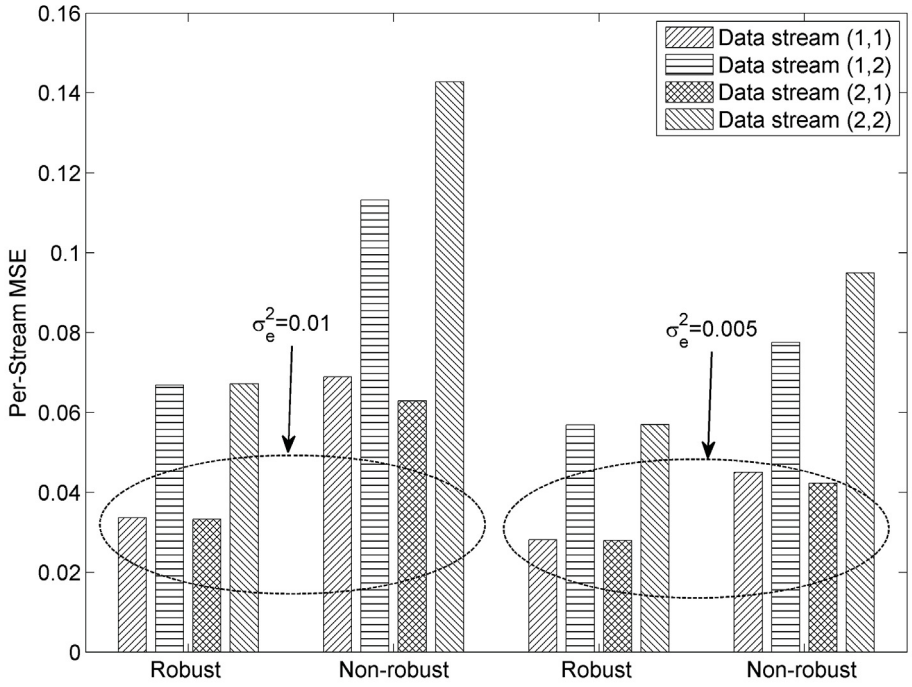
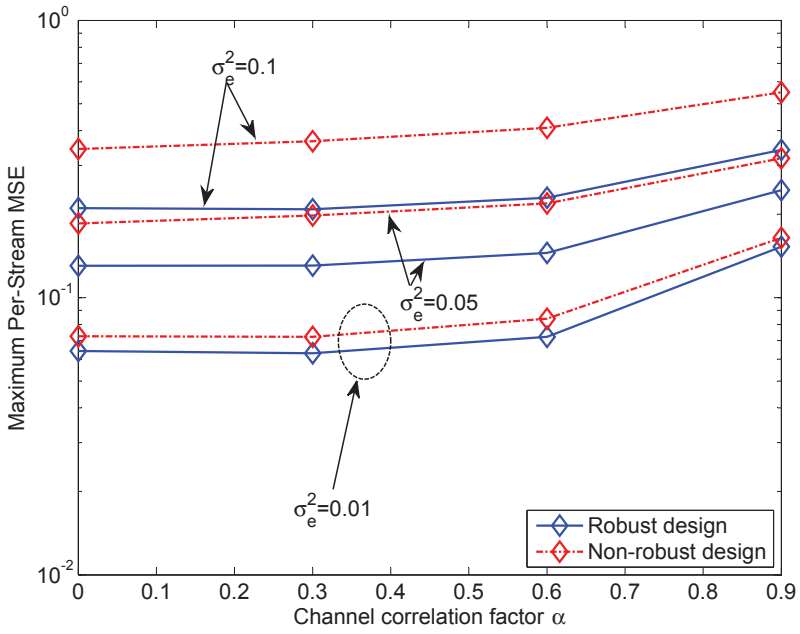


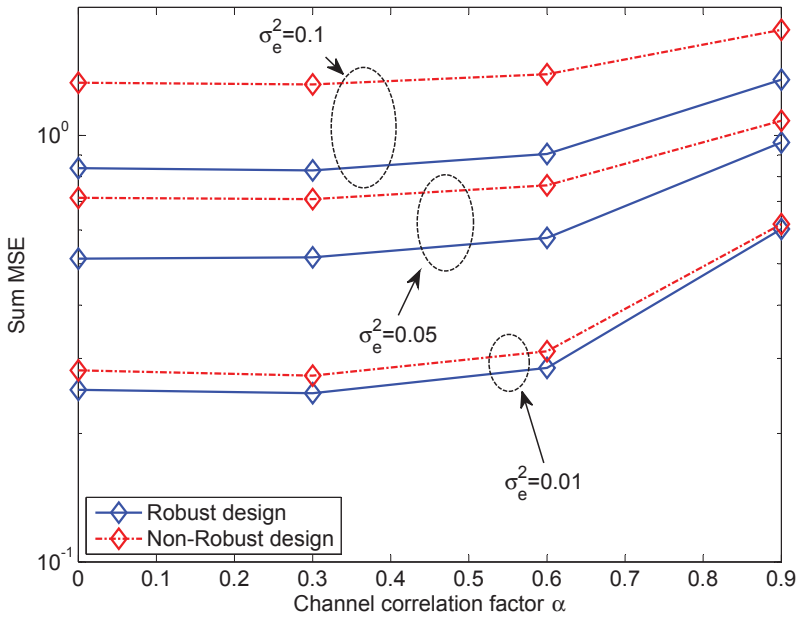

**Fig. 3.6** Comparison of the per-stream MSEs of the robust and non-robust design approaches ($\mathrm{SNR}_{\mathtt{R},m} = \mathrm{SNR}_{\mathtt{D},k} = 15\mathrm{dB}$, $\alpha = \beta = 0.5$).

**Experiment A.2 (Data stream fairness)**: Next we examine the accuracy of the proposed robust design in providing weighted fairness for the different data streams. To this end, we set the weights for the different data streams to be $\kappa_{1,1} = \kappa_{2,1} = \frac{1}{3}$ and $\kappa_{1,2} : \kappa_{2,2} = \frac{1}{6}$. Fig. (3.6) depicts the MSE of each data stream for different values of the error variance. Comparing the two methods, the robust design approach results in a significantly better weighted fairness than the non-robust one. In particular, the MSEs obtained are strictly inversely proportional to the predefined weights. This feature is specially desirable for multimedia communications, where the streams corresponding to different service types may have different priorities.

**Experiment A.3 (Effects of channel correlation)**: The effects of channel correlations on the MSE performance of the different approaches are investigated in Fig. 3.7. It can be observed that the performance of all the approaches is degraded, as the correlation factor $\alpha$ increases. While the robust design shows consistent performance gains




(a)




(b)

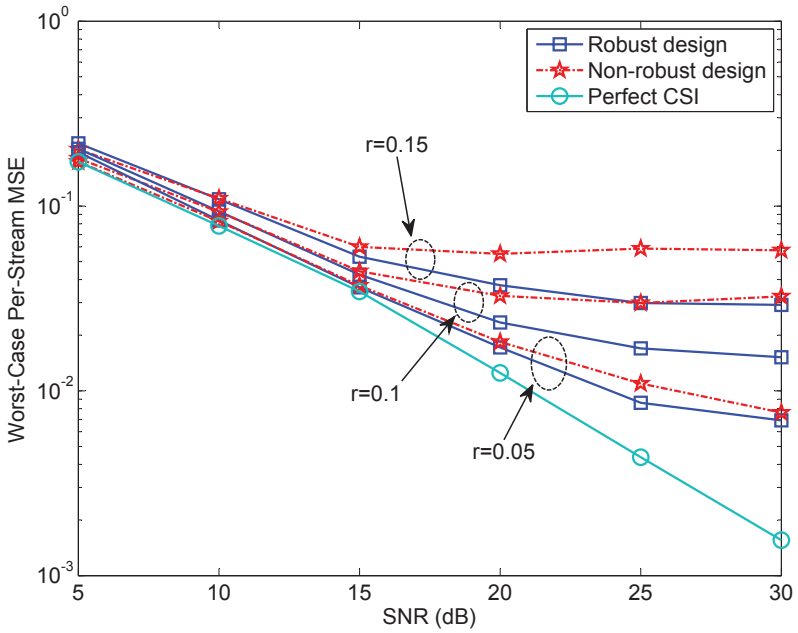
**Fig. 3.7** MSE performance of different design approaches versus correlation factor of the source-relay channels. (a) per-stream MSE; (b) sum MSE ($\mathrm{SNR}_{\mathtt{R},m} = \mathrm{SNR}_{\mathtt{D},k} = 10\mathrm{dB}$, $\beta = 0.45$).

over its non-robust one associated with different $\alpha$ and $\sigma_e^2$, the discrepancies between the two approaches tend to become less significant with an increase in $\alpha$. This is because the achievable spatial multiplexing gain is reduced by a higher channel correlation and therefore the robust design can only attain a limited performance improvement in the presence of high channel correlations.

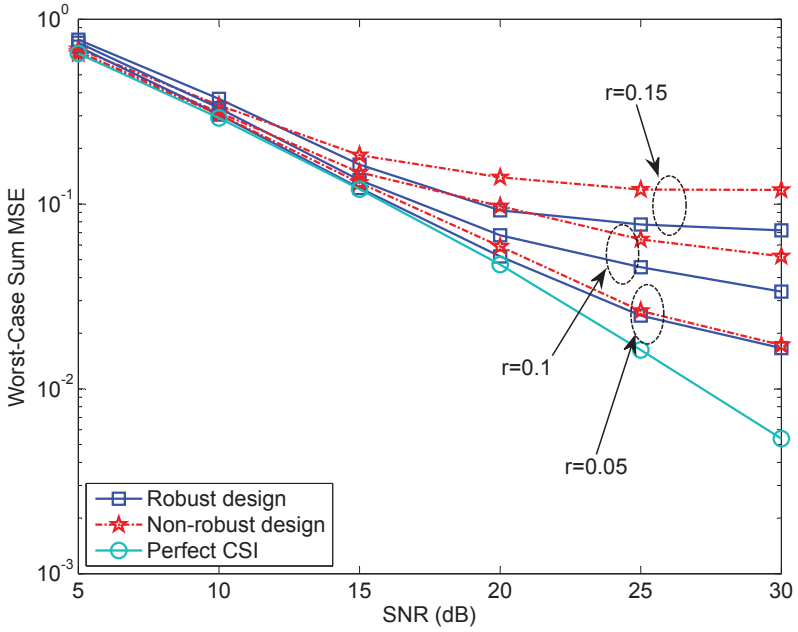
### 3.5.2 Performance Evaluation under Norm-Bounded CSI Errors

In this subsection, we evaluate the performance of the proposed worst-case design approach in Section 3.3 for the min-max problem under norm-bounded CSI errors. Similar to the previous subsection, we compare the proposed robust design approach both to the non-robust approach and to the perfect CSI scenario. We note that the power of each relay is a function of $\Delta\mathbf{H}_m$. According to the worst-case robust design philosophy, the maximum relay transmit power has to be bounded by the power budget, while the average relay transmit power may become significantly lower than that of the non-robust design. To facilitate a fair comparison of the different approaches, we therefore assume the absence of CSI errors for the S-R links, i.e., we have $\Delta\mathbf{H}_{m,k} = \mathbf{0}$. For the R-D links, we consider the uncertainty regions with equal radius, i.e., we have $\xi_{k,m} = r, \forall k \in \mathcal{K}, m \in \mathcal{M}$. In order to determine the worst-case per-stream MSE, we generate 5000 independent realizations of the norm-bounded CSI errors. In each realization, the entries of the CSI error matrix $\Delta\mathbf{G}_{k,m}$ are first generated using independent complex circular Gaussian random variables with zero mean and unit variance. In the case that $\|\Delta\mathbf{G}_{k,m}\|_F > r$, the entries are scaled down by a factor of $c$, i.e., $\frac{1}{c}\|\Delta\mathbf{G}_{k,m}\|_F = r$. For each realization, we evaluate the maximum per-stream MSE averaged over 1000 QPSK symbols and random Gaussian noise. Then the worst-case per-stream MSE is obtained by selecting the largest one among all the realizations.

**Experiment B.1 (MSE performance)**: The worst-case per-stream MSE and the worst-case sum MSE are reported in Fig. 3.8 as a function of the SNR. Three sizes of the uncertainty region are considered, that is: $r = 0.05$, $r = 0.1$ and $r = 0.15$. Focusing on the first case, it can be seen that the performance achieved by our robust design approach first monotonically decreases as the SNR increases, and then subsequently remains approximately constant at high SNR values. This is primarily because at low SNR, the main source of error in the estimation of the data streams is the channel noise. At high SNR, the channel noise is no longer a concern and the MSE is dominated by the CSI errors.




(a)




(b)

**Fig. 3.8** MSE performance of different design approaches versus SNR. (a) worst-case per-stream MSE; (b) worst-case sum MSE.

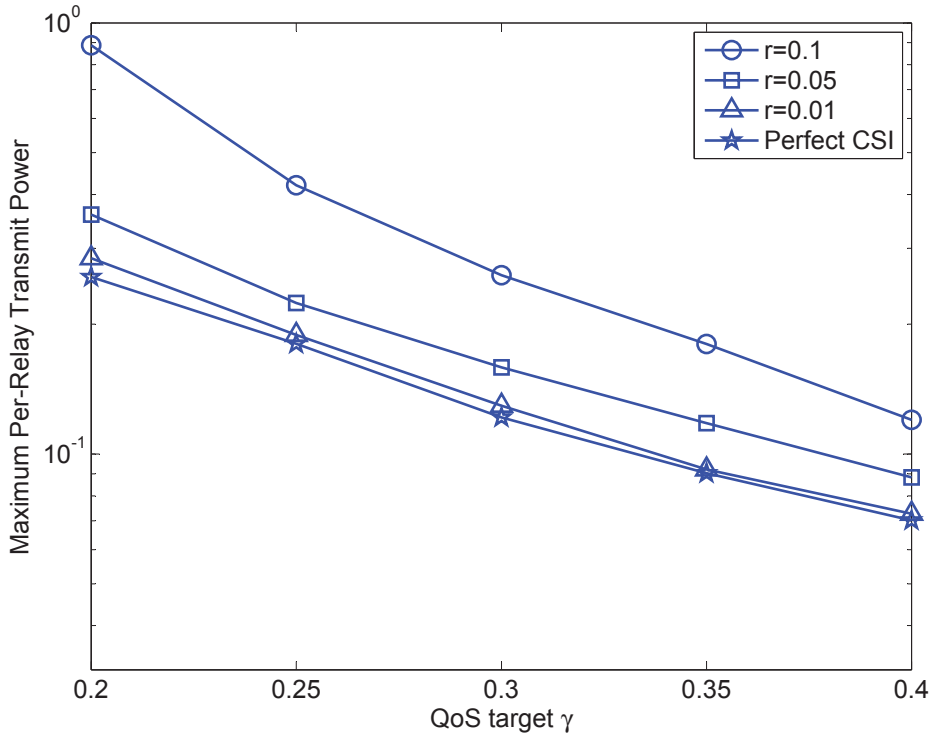



**Fig. 3.9** Maximum relay transmit power versus QoS targets with different uncertainty sizes of the CSI errors.

Observe also in Fig. 3.8 that for $r = 0.1$ and $r = 0.15$, the MSE is clearly higher, although it presents a similar trend to the case of $r = 0.5$. The performance gain achieved by the robust design also becomes more noticeable for these larger sizes of the uncertainty regions.

**Experiment B.2 (Relay power consumption)**: Next we investigate the performance of the approach proposed in Section 3.4 for the QoS problem under the norm-bounded CSI errors. The maximum per-relay transmit power is plotted in Fig. 3.9 as a function of the QoS target $\gamma$ for different sizes of uncertainty regions. As expected, it can be observed that the relay power for all cases decreases as the QoS target is relaxed. An important observation from this figure is that when the size of uncertainty region is large, the required relay transmit power becomes significantly higher than the perfect CSI case. From an energy-efficient design perspective, this is not desirable, which motivates the consideration of the min-max design in such applications.

**Experiment B.3 (Distribution of per-stream MSE)**: Finally, we evaluate how

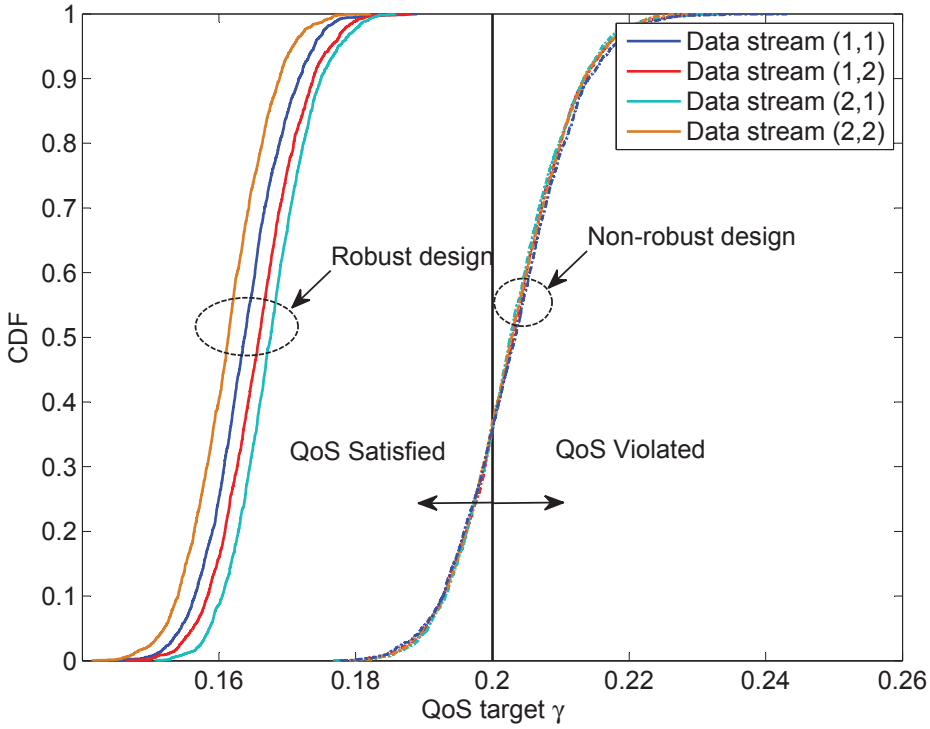



**Fig. 3.10** CDFs of per-stream MSEs using the robust and non-robust approaches for SNR=5dB.

consistently the QoS constraints of all the data streams can be satisfied by the proposed design approach for the QoS problem. In this experiment, the CSI errors of both the S-R and R-D links are taken into consideration and generated according to the i.i.d. zero-mean complex Gaussian distribution with a variance of $\sigma_e^2 = 0.001$. Then the probability that the CSI errors are bounded by the predefined radius $r$ can be formulated as [49, Section IV-C]

$$\Pr\left\{\|\mathbf{h}_{m,k}\|^2 \leq r^2\right\} = \Pr\left\{\|\mathbf{g}_{k,m}\|^2 \leq r^2\right\} = \frac{1}{\Gamma\left(\frac{N^2}{2}\right)}\gamma\left(\frac{N^2}{2}, \frac{r^2}{\sigma_e^2}\right), \tag{3.64}$$

where $\Gamma(\cdot)$ and $\gamma(\cdot,\cdot)$, respectively, denote the complete and lower incomplete Gamma functions. Given the required bounding probability of, say 90% in the simulation, the radius $r$ can be numerically determined from (3.64). Fig. 3.10 shows the cumulative distribution functions (CDFs) of the MSE of each data stream using both the robust and non-robust

design methods. As expected, the proposed robust method ensures that the MSE of each data stream never exceeds the QoS target shown as the vertical black solid line in Fig. 3.10. By contrast, for the non-robust design, the MSE frequently violates the QoS target, namely for more than 60% of the realizations. Based on these observations, we conclude that the proposed robust design approach outperforms its non-robust counterpart in satisfying the QoS constraints for all the data streams.

## 3.6 Conclusions

In this chapter, jointly optimized source TPCs, AF relay matrices and receive filters were designed by considering two different types of objective functions with specific QoS consideration in the presence of CSI errors in both the S-R and R-D links. To this end, a pair of practical CSI error models, namely, the statistical and the norm-bounded models were considered. Accordingly, the robust transceiver design approach was formulated to minimize the maximum per-stream MSE subject to the source and relay power constraints (min-max problem). In order to solve the non-convex optimization problems formulated, an iterative solution based on the block coordinate update algorithm was proposed, which involves a sequence of convex conic optimization problems. The proposed algorithm generated a convergent sequence of objective function values. The problem of relay power minimization subject to specific QoS constraints and to source power constraints was also studied. An efficient feasibility search algorithm was proposed by studying the link between the feasibility check and the min-max problems. Our simulation results demonstrate a significant enhancement in the performance of the proposed robust approaches over the conventional non-robust approaches.

# Chapter 4

# Centralized Energy Efficient Relaying in Next-Generation RANs

In this chapter, we investigate the minimization of energy consumption of a multi-user relaying network operating within a C-RAN cluster. Specifically, we consider a network, where multiple source and destination pairs (also termed UE in LTE) communicate with each others by relying on the assistance of multiple cooperative RRHs connected to the BBU pool. Taking into account both the static (fixed circuitry) and dynamic (RF transmission) components of the RRHs power [107], a joint RRH selection as well as AF matrix design problem is formulated, whereby the total power consumption is minimized over the relaying network while maintaining a predefined QoS level at each destination UE. The main contributions of this chapter are summarized as follows.

- To solve the resultant non-convex problem and to arrive at a sparse solution, an iterative approach relying on the concepts of re-weighted $l_1$ norm minimization [96] along with a BCD-type update [53] is proposed. Following this optimization, the subset of active RRHs selected for relaying the transmission is determined in a single instance by thresholding the recovered group sparsity pattern vector.

- We thoroughly analyze the convergence properties of the proposed algorithm. Specifically, we show that the proposed algorithm is equivalent to a modified block-type

Part of the materials in this chapter has been published in the IEEE Transactions on Vehicular Technology [106].

majorization minimization (MM) algorithm [108], and prove that the generated solution sequence converges to a so-called Nash point, which is a generalized concept of a stationary point.

- To overcome a potentially undesired scenario, where the relaying network fails to simultaneously satisfy all the destination UEs' QoS levels with the aid of only the selected subset of RRHs, we propose an iterative user admission control scheme. Explicitly, in this proposed scheme, specific destination UEs exhibiting a high "infeasibility indicator" will be excluded from the optimization procedure by the network for the sake of maintaining the QoS of all the other users.

- The quantitative benefits of the proposed joint RRH selection and transceiver optimization algorithm and the UE admission control scheme are demonstrated by computer simulations.

The rest of the chapter is organized as follows. The relaying system model under the C-RAN is introduced in Section 4.1. In Section 4.2, we formulate and solve the network energy minimization problem. Furthermore, an iterative UE admission control mechanism is proposed for overcoming the potential infeasibility issue as mentioned above. Simulation results quantifying the benefits of the proposed algorithms are provided in Section 4.3. Finally, the chapter is concluded in Section 4.4.

## 4.1 System Model

Consider a multi-user relaying network within a C-RAN, consisting of $L$ multi-antenna RRHs and $K$ pairs of single-antenna source UEs (SUEs) and destination UEs (DUEs), as depicted in Fig. 4.1 (a). The $L$ RRHs work collaboratively under the C-RAN umbrella, relaying the messages of the SUEs to their corresponding DUEs. It is assumed that each SUE only communicates with its paired DUE. The RRHs and SUE–DUE pairs are indexed by the sets $\mathcal{L} \triangleq \{1, 2, \cdots, L\}$ and $\mathcal{K} \triangleq \{1, 2, \cdots, K\}$, respectively. For each $l \in \mathcal{L}$, RRH-$l$ is equipped with $N_l$ antennas and operates in a half-duplex AF mode. It is assumed that no direct links are available between the SUEs and DUEs due to a high pathloss.

We consider a narrowband flat-fading channel model, where $\mathbf{h}_{l,k} \in \mathbb{C}^{N_l \times 1}$ specifies the channel between SUE-$k$ and RRH-$l$, while $\mathbf{g}_{k,m} \in \mathbb{C}^{N_l \times 1}$ denotes the Hermitian transpose of

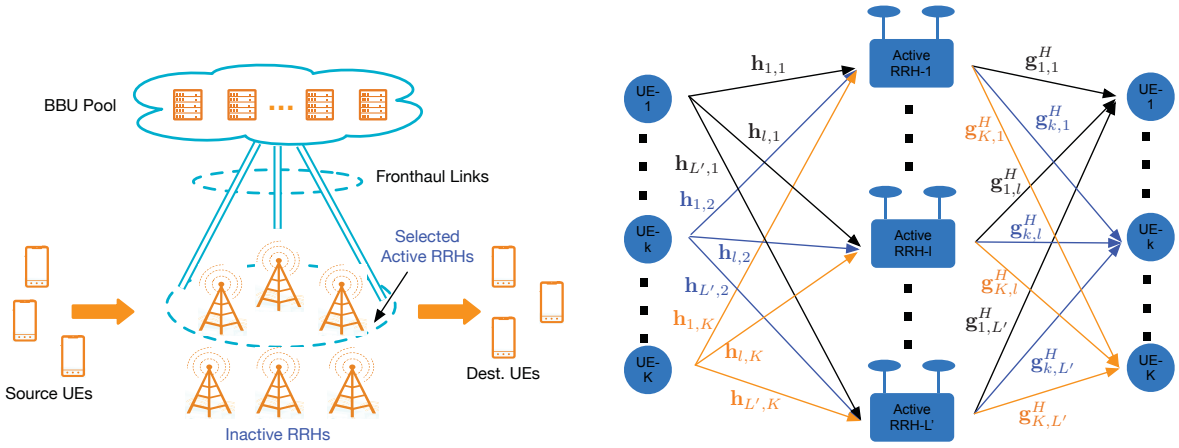



**Fig. 4.1** Multi-user multi-relay network within a C-RAN and illustration of the interference scenario within the relaying network.

the channel between RRH-$l$ and DUE-$k$. Let $s_k$ denote the information symbol transmitted by SUE-$k$ at a specific time instance, which is modeled as a zero-mean unit-variance complex random variable. During the first transmission phase, RRH-$l$ receives the following signal:

$$\mathbf{r}_l = \sum_{k=1}^{K} \mathbf{h}_{l,k} s_k + \mathbf{n}_{\mathrm{R},l}, \quad l \in \mathcal{L}, \tag{4.1}$$

where $\mathbf{n}_{\mathrm{R},l}(n)$ denotes the spatially white, additive noise vector at RRH-$l$ with zero mean and covariance matrix $\sigma^2_{\mathrm{R},l}\mathbf{I}_{N_l}$. During the second phase, for each $l \in \mathcal{L}$, RRH-$l$ applies a linear transformation matrix $\mathbf{W}_l \in \mathbb{C}^{N_l \times N_l}$ to $\mathbf{r}_l$ and forwards the resultant signal to all DUEs. The signal received by DUE-$k$ can be expressed as

$$\begin{aligned} y_k &= \sum_{l=1}^{L} \mathbf{g}_{k,l}^H \mathbf{W}_l \mathbf{r}_l + n_{\mathrm{D},k} \\ &= \underbrace{\sum_{l=1}^{L} \mathbf{g}_{k,l}^H \mathbf{W}_l \mathbf{h}_{l,k} s_k}_{\text{Desired Signal}} + \underbrace{\sum_{l=1}^{L} \sum_{\substack{m=1,\\ m\neq k}}^{K} \mathbf{g}_{k,l}^H \mathbf{W}_l \mathbf{h}_{l,m} s_m}_{\text{Multi-User Interference}} + \underbrace{\sum_{l=1}^{L} \mathbf{g}_{k,l}^H \mathbf{W}_l \mathbf{n}_{\mathrm{R},l} + n_{\mathrm{D},k}}_{\text{Noise Terms}}, \quad k \in \mathcal{K}, \end{aligned} \tag{4.2}$$

where $n_{\mathrm{D},k}$ denotes the additive white noise at DUE-$k$ with zero mean and a variance of $\sigma^2_{\mathrm{D},k}$.

The above expression indicates that the signal received at each DUE is a superposition of the desired signal component, the multi-user interference (co-channel interference) arriving from the other SUEs as well as the noise contributions from the RRHs and the DUE. This general interference scenario, which is illustrated in Fig. 4.1 (b), is often referred to as the relay-aided interference channel in the literature of relaying optimization [25,28,30,97,109] and of interference alignment [110,111].

We rely on the mean-square MSE as the QoS metric for the received signal of each DUE. After DUE-$k$ applies an equalizer gain $u_k$ to its received signal $y_k$, resulting in the soft estimate $\hat{s}$, i.e., $\hat{s}_k = u_k^* y_k$, the MSE at DUE-$k$ can be expressed as

$$\begin{aligned}\mathtt{MSE}_k(u_k, \{\mathbf{W}_l\}) &= \mathbb{E}\left\{|u_k^* y_k - s_k|^2\right\} \\ &= \sum_{q=1}^{K} \Big| \sum_{l=1}^{L} \mathbf{g}_{k,l}^H \mathbf{W}_l \mathbf{h}_{l,q} \Big|^2 |u_k|^2 - \sum_{l=1}^{L} 2\,\mathrm{Re}\left\{u_k^* \mathbf{g}_{k,l}^H \mathbf{W}_l \mathbf{h}_{l,k}\right\} \\ &\quad + \sum_{l=1}^{L} \sigma_{\mathtt{R},l}^2 \|\mathbf{g}_{k,l}^H \mathbf{W}_l\|_2^2 |u_k|^2 + \sigma_{\mathtt{D},k}^2 |u_k|^2 + 1. \end{aligned} \tag{4.3}$$

The transmission power required for AF relaying at RRH-$l$ is given by

$$P_{t,l} = \mathrm{Tr}\left(\mathbf{W}_l \big(\sigma_{\mathtt{R},l}^2 \mathbf{I}_{N_l} + \sum_{k=1}^{K} \mathbf{h}_{l,k} \mathbf{h}_{l,k}^H\big) \mathbf{W}_l^H\right). \tag{4.4}$$

In addition to the transmission power, we also consider the static power consumption associated with each RRH, including that of the RF circuitry, A/D converion and optical fronthaul. This power, which is non-negligible, can be saved when the associated RRH is switched off [69–71]. Denoting the static power level of RRH-$l$ by $P_{c,l}$, the total relaying network power consumption is given by

$$P = \sum_{l=1}^{L} \mathcal{I}\left(\|\mathbf{W}_l\|_F^2\right) P_{c,l} + \sum_{l=1}^{L} P_{t,l}, \tag{4.5}$$

where $\mathcal{I}(\cdot)$ denotes the indicator function, i.e., $\mathcal{I}(x) = 1$ if $x \neq 0$ and $\mathcal{I}(x) = 0$, otherwise. Based on (4.3) and (4.5), we then formulate the energy minimization problem with QoS constraints in the next section.

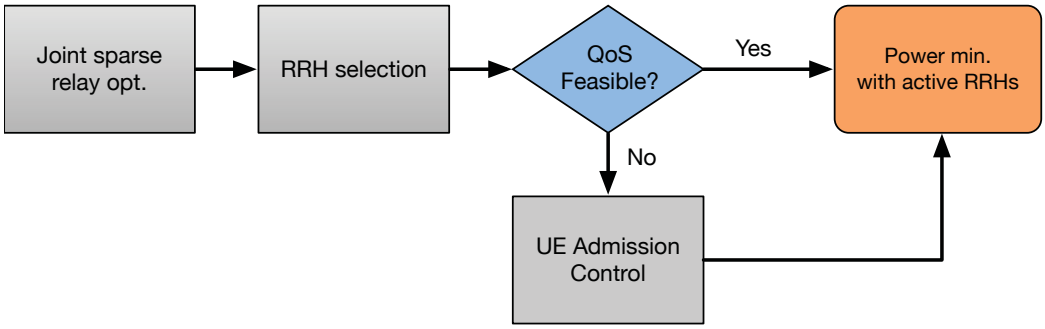



**Fig. 4.2** The overall flow of the proposed algorithm.

## 4.2 Energy Minimization Design and UE Admission Control

In this section, we first develop an energy-efficient multi-user relaying solution for C-RAN by jointly optimizing the relay AF matrices $\{\mathbf{W}_l\}$ at all RRHs and performing RRH selection. To circumvent the undesirable condition where the DUEs' QoS levels cannot all be simultaneously satisfied by the selected subset of RRHs, we will also propose a UE admission control mechanism for overcoming the associated *infeasibility* problem. The overall flow of the proposed solution is depicted in Fig. 4.2, which consists of the following main steps:

*1)* Optimize the AF relaying matrices across all RRHs by introducing a sparsity-inducing $l_1$-norm.

*2)* Perform RRH selection by exploiting the group sparsity pattern obtained from the first step.

*3)* If the above-mentioned infeasibility problem is observed, involve the UE admission control algorithm.

### 4.2.1 Iterative algorithm

We first aim for minimizing the total energy consumption of the relaying network by simultaneously designing the AF matrices $\{\mathbf{W}_l\}_{l\in\mathcal{L}}$ for all the RRHs and the equalizer gains $\{u_k\}$ for all the DUEs, while still maintaining a predefined QoS level at each DUE. To connect this problem to the vast body of literature on sparse signal recovery [112], we express the

indicator function $\mathcal{I}(\cdot)$ in (4.5) in terms of the $l_0$-norm $\|\cdot\|_0$. The design problem can then be mathematically expressed as [1]

$$\min_{\mathbf{u},\{\mathbf{W}_l\}} \quad \sum_{l=1}^{L} \left\| \|\mathbf{W}_l\|_F^2 \right\|_0 P_{c,l} + \sum_{l=1}^{L} P_{t,l} + \|\mathbf{u}\|_2^2 \tag{4.6a}$$

$$\text{s.t.} \quad \mathtt{MSE}_k(u_k, \{\mathbf{W}_l\}) \leq \gamma_k, \ \forall k \in \mathcal{K} \tag{4.6b}$$

$$P_{t,l} \leq P_{l,\max}, \ \forall l \in \mathcal{L}, \tag{4.6c}$$

where $\gamma_k$ denotes the predefined target MSE of DUE-$k$ and $\mathbf{u} = [u_1, \cdots, u_K]^T$.

To overcome the non-convexity issues of (4.6a) due to the $l_0$-norm, we follow a similar approach as in [96] and use the so-called re-weighted $l_1$-norm minimization for approximating the non-convex $l_0$-norm. Specifically, this approach yields $\left\| \|\mathbf{W}_l\|_F^2 \right\|_0 \approx \frac{\mu}{\|\mathbf{W}_l^{(n-1)}\|_F^2 + \varepsilon} \|\mathbf{W}_l\|_F^2$, where $\mathbf{W}_l^{(n-1)}$ denotes the value of $\mathbf{W}_l$ obtained from the previous iteration (with index $n-1$), $\mu$ is a positive constant (to be specified later) and $\varepsilon > 0$ is a small positive constant ensuring numerical stability. Using the latter approximation, (4.6) can be reformulated as:

$$\min_{\mathbf{u},\{\mathbf{W}_l\}} \quad \sum_{l=1}^{L} \frac{\mu}{\|\mathbf{W}_l^{(n-1)}\|_F^2 + \varepsilon} \|\mathbf{W}_l\|_F^2 P_{c,l} + \sum_{l=1}^{L} P_{t,l} + \|\mathbf{u}\|_2^2 \tag{4.7a}$$

$$\text{s.t.} \quad \mathtt{MSE}_k(u_k, \{\mathbf{W}_l\}) \leq \gamma_k, \ \forall k \in \mathcal{K} \tag{4.7b}$$

$$P_{t,l} \leq P_{l,\max}, \ \forall l \in \mathcal{L}\,. \tag{4.7c}$$

However, after this approximation, the above problem still remains non-convex due to the fact that the variables $\{u_k\}$ and $\{\mathbf{W}_l\}$ are nonlinearly coupled in the QoS constraints (4.7b). Then, a further inspection of the MSE expression (4.3) reveals that it has a so-called bi-convex structure, i.e., it is convex with respect to one block of variables when the other is fixed. Based on this property, we can solve the problem by employing the BCD-type algorithm of [98], which makes it possible to update the two blocks of variables one at a time while fixing the values of the other. In this way, (4.7) can be solved iteratively with

[1] The addition of the term $\|\mathbf{u}\|_2^2$ in the objective function (4.6a) makes the latter strongly convex in $\mathbf{u}$. This modification does not affect the feasibility of the problem, i.e., the feasible set $\mathcal{F} = \{(\mathbf{u}, \{\mathbf{W}_l\}) : (4.6b), (4.6c)\}$ remains unchanged. More importantly, the strong convexity of (4.6a) is useful in ensuring the convergence of the algorithm derived subsequently.

respect to $\mathbf{u}$ and $\{\mathbf{W}_l\}$ in Gauss-Seidel fashion. Below, we formulate the two sub-problems and derive their corresponding solutions.

### Updating u

With fixed AF relaying matrices $\{\mathbf{W}_l\}$, the sub-problem solution of finding the optimal $\mathbf{u}^{\text{opt}}$ can be expressed as

$$\min_{\mathbf{u}} \quad \|\mathbf{u}\|_2^2 \tag{4.8a}$$

$$\text{s.t.} \quad \mathtt{MSE}_k(u_k) \leq \gamma_k, \ \forall k \in \mathcal{K}. \tag{4.8b}$$

Noting the separable structure of the above problem with respect to $u_1, \cdots, u_K$, it can be further decomposed into $K$ parallel sub-problems given by:

$$\min_{u_k} \quad |u_k|^2 \tag{4.9a}$$

$$\begin{aligned}\text{s.t.} \quad & \sum_{q=1}^{K} \Big| \sum_{l=1}^{L} \mathbf{g}_{k,l}^H \mathbf{W}_l \mathbf{h}_{l,q} \Big|^2 |u_k|^2 - \sum_{l=1}^{L} 2\Re\left\{u_k^* \mathbf{g}_{k,l}^H \mathbf{W}_l \mathbf{h}_{l,k}\right\} \\ & \quad + \sum_{l=1}^{L} \sigma_{\mathtt{R},l}^2 \|\mathbf{g}_{k,l}^H \mathbf{W}_l\|_2^2 |u_k|^2 + \sigma_{\mathtt{D},k}^2 |u_k|^2 + 1 \leq \gamma_k. \end{aligned} \tag{4.9b}$$

Below we show that the optimal solution $u_k^{\text{opt}}$ of (4.9) in fact can be obtained as a scaled version of the MMSE filter $u_k^{\text{MMSE}}$. Observe that (4.9) is a strictly convex quadratic problem, which is also strictly feasible, i.e., Slater's constraint qualification holds [93]. Therefore, there exists a unique globally optimal solution of (4.9), which can be obtained by evaluating the Karush-Kuhn-Tucker (KKT) conditions of (4.9), which are sufficient conditions in this case. First, the Lagrangian function of (4.9) can be written as

$$\begin{aligned}\mathcal{L}(u_k, \mu_k) = |u_k|^2 + \mu_k \Big( & \sum_{q=1}^{K} \Big| \sum_{l=1}^{L} \mathbf{g}_{k,l}^H \mathbf{W}_l \mathbf{h}_{l,q} \Big|^2 |u_k|^2 - \sum_{l=1}^{L} 2\Re\left\{u_k^* \mathbf{g}_{k,l}^H \mathbf{W}_l \mathbf{h}_{l,k}\right\} \\ & + \sum_{l=1}^{L} \sigma_{\mathtt{R},l}^2 \|\mathbf{g}_{k,l}^H \mathbf{W}_l\|_2^2 |u_k|^2 + \sigma_{\mathtt{D},k}^2 |u_k|^2 + 1 - \gamma_k \Big), \end{aligned} \tag{4.10}$$

where $\mu_k \geq 0$ denotes the dual variable associated with (4.9b). The first-order KKT condition can then be formulated as

$$\begin{aligned}&\frac{\partial \mathcal{L}(u_k, \mu_k)}{\partial u_k^*} \\ &= u_k \Bigg( \mu_k \sum_{q=1}^{K} \Big| \sum_{l=1}^{L} \mathbf{g}_{k,l}^H \mathbf{W}_l \mathbf{h}_{l,q} \Big|^2 + \mu_k \sum_{l=1}^{L} \sigma_{\mathtt{R},l}^2 \|\mathbf{g}_{k,l}^H \mathbf{W}_l\|_2^2 \\ &\qquad + \mu_k \sigma_{\mathtt{D},k}^2 + 1 \Bigg) - \mu_k \sum_{l=1}^{L} \mathbf{g}_{k,l}^H \mathbf{W}_l \mathbf{h}_{l,k} = 0. \end{aligned} \tag{4.11}$$

Re-arranging the above, we obtain $u_k^{\text{opt}}$ as

$$u_k^{\text{opt}} = \frac{\mu_k \sum\limits_{l=1}^{L} \mathbf{g}_{k,l}^H \mathbf{W}_l \mathbf{h}_{l,k}}{\mu_k \left( \sum\limits_{q=1}^{K} \left| \sum\limits_{l=1}^{L} \mathbf{g}_{k,l}^H \mathbf{W}_l \mathbf{h}_{l,q} \right|^2 + \sum\limits_{l=1}^{L} \sigma_{\mathtt{R},l}^2 \|\mathbf{g}_{k,l}^H \mathbf{W}_l\|_2^2 + \sigma_{\mathtt{D},k}^2 \right) + 1}. \tag{4.12}$$

Meanwhile, the MMSE filter can be obtained by setting the partial derivative of $\mathtt{MSE}_k(u_k)$ in (2.1) to zero, yielding, $\frac{\partial \mathtt{MSE}_k(u_k)}{\partial u_k^*} = 0$, whose solution is

$$u_k^{\text{MMSE}} = \frac{\sum\limits_{l=1}^{L} \mathbf{g}_{k,l}^H \mathbf{W}_l \mathbf{h}_{l,k}}{\sum\limits_{q=1}^{K} \left| \sum\limits_{l=1}^{L} \mathbf{g}_{k,l}^H \mathbf{W}_l \mathbf{h}_{l,q} \right|^2 + \sum\limits_{l=1}^{L} \sigma_{\mathtt{R},l}^2 \|\mathbf{g}_{k,l}^H \mathbf{W}_l\|_2^2 + \sigma_{\mathtt{D},k}^2}. \tag{4.13}$$

Upon dividing (4.13) by (4.12), we obtain

$$\frac{u_k^{\text{MMSE}}}{u_k^{\text{opt}}} = \frac{\mathcal{C} + \frac{1}{\mu_k}}{\mathcal{C}}, \tag{4.14}$$

where $\mathcal{C} = \sum_{q=1}^{K} \left| \sum_{l=1}^{L} \mathbf{g}_{k,l}^H \mathbf{W}_l \mathbf{h}_{l,q} \right|^2 + \sum_{l=1}^{L} \sigma_{\mathtt{R},l}^2 \|\mathbf{g}_{k,l}^H \mathbf{W}_l\|_2^2 + \sigma_{\mathtt{D},k}^2$ is a positive constant given fixed $\{\mathbf{W}_l\}$. Upon defining $\kappa_k = 1 + \frac{1}{\mathcal{C}\mu_k} > 1$, we now recognize that $u_k^{\text{opt}}$ can be found by

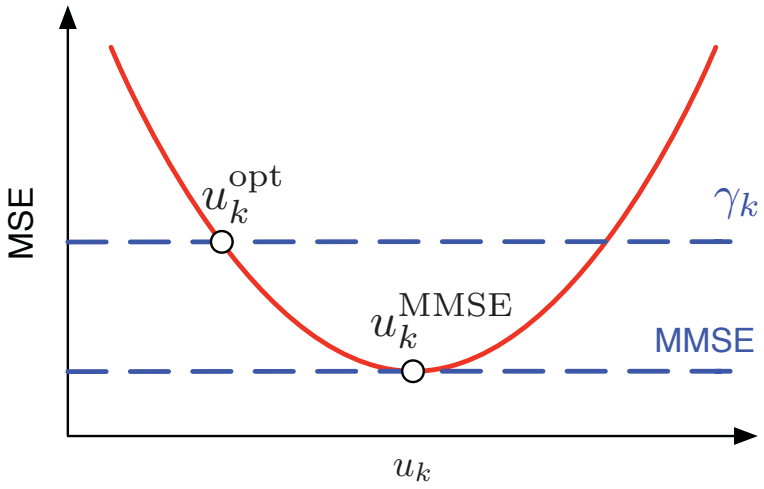



**Fig. 4.3** Illustration of how to obtain the optimal $u_k^*$ from the MMSE solution $u_k^{\text{MMSE}}$.

scaling down the MMSE filter $u_k^{\text{MMSE}}$ by $\kappa_k$ as follows:

$$u_k^{\text{opt}} = \frac{u_k^{\text{MMSE}}}{\kappa_k} = \frac{1}{\kappa_k} \frac{\sum_{l=1}^{L} \mathbf{g}_{k,l}^H \mathbf{W}_l \mathbf{h}_{l,k}}{\sum_{q=1}^{K} \left| \sum_{l=1}^{L} \mathbf{g}_{k,l}^H \mathbf{W}_l \mathbf{h}_{l,q} \right|^2 + \sum_{l=1}^{L} \sigma_{\mathrm{R},l}^2 \|\mathbf{g}_{k,l}^H \mathbf{W}_l\|_2^2 + \sigma_{\mathrm{D},k}^2}. \tag{4.15}$$

To gain more insights into the computation of $u_k^{\text{opt}}$, we illustrate the relationship between $u_k^{\text{opt}}$ and the MMSE filter $u_k^{\text{MMSE}}$ in Fig. 4.3, where for simplicity, we assume that $u_k$ is real-valued. Since $\mathtt{MSE}_k(u_k)$ is a quadratic function in $u_k$, we can first compute the global minimum, i.e., the MMSE solution $u_k^{\text{MMSE}}$, and then simply scaling it down by a factor $\kappa_k$ while still satisfying the MSE constraint (4.8b), i.e., letting $\mathtt{MSE}_k(u_k) = \gamma_k$.

### Updating $\{\mathbf{W}_l\}$

We then proceed to solve the sub-problem for $\{\mathbf{W}_l\}$ when $\mathbf{u}$ is fixed. To this end, we define the concatenated weight vectors $\mathbf{w}_l \triangleq \mathtt{vec}(\mathbf{W}_l)$ for all $l$, as well as the following matrices

and vectors, which are independent of $\{\mathbf{w}_l\}$:

$$\mathbf{Q}_{l,m}^{k} = \sum_{q=1}^{K} \left( \mathbf{h}_{l,q}^{*} u_k u_k^{*} \mathbf{h}_{l,q}^{T} \otimes \mathbf{g}_{q,l} \mathbf{g}_{q,l}^{H} \right) \tag{4.16}$$

$$\mathbf{\Psi}_l = \sigma_{\mathrm{R},l}^2 \mathbf{I}_{N_l^2} + \mathbf{I}_{N_l} \otimes \sum_{k=1}^{K} \mathbf{h}_{l,k} \mathbf{h}_{l,k}^{H} \tag{4.17}$$

$$\mathbf{\Theta}_{k,l} = \sigma_{\mathrm{R},1}^2 \mathbf{I} \otimes (\mathbf{g}_{k,1} u_k u_k^{*} \mathbf{g}_{k,1}^{H}) \tag{4.18}$$

$$\mathbf{q}_{k,l} = \mathbf{h}_{l,k}^{*} \otimes \mathbf{g}_{k,l}. \tag{4.19}$$

Then, by exploiting the useful properties of the Kronecker product, i.e., $\mathrm{Tr}\left(\mathbf{A}^H \mathbf{B} \mathbf{C} \mathbf{D}^H\right) = \mathtt{vec}\left(\mathbf{A}\right)^H \left(\mathbf{D}^T \otimes \mathbf{B}\right) \mathtt{vec}\left(\mathbf{C}\right)$, $\mathrm{Tr}\left(\mathbf{A}^H \mathbf{B} \mathbf{A}\right) = \mathtt{vec}\left(\mathbf{A}\right)^H \left(\mathbf{I} \otimes \mathbf{B}\right) \mathtt{vec}\left(\mathbf{A}\right)$ and $\mathrm{Tr}\left(\mathbf{A}^H \mathbf{B}\right) = \mathtt{vec}\left(\mathbf{B}\right)^H \mathtt{vec}\left(\mathbf{A}\right)$, we arrive at the following convex quadratically-constrained quadratic problem (QCQP):

$$\min_{\mathbf{w}} \quad \sum_{l=1}^{L} \frac{\mu}{\|\mathbf{w}_l^{(n-1)}\|_2^2 + \varepsilon} \|\mathbf{w}_l\|_2^2 P_{c,l} + \sum_{l=1}^{L} \mathbf{w}_l^H \mathbf{\Psi}_l \mathbf{w}_l \tag{4.20a}$$

$$\text{s.t.} \quad \sum_{l=1}^{L} \sum_{m=1}^{L} \mathbf{w}_l^H \mathbf{Q}_{l,m}^{k} \mathbf{w}_m - \sum_{l=1}^{L} 2\Re\left\{\mathbf{w}_l^H \mathbf{q}_{k,l}\right\} + \sum_{l=1}^{L} \mathbf{w}_l^H \mathbf{\Theta}_{k,l} \mathbf{w}_l + q_k \leq \gamma_k, \ \forall k \in \mathcal{K} \tag{4.20b}$$

$$\mathbf{w}_l^H \mathbf{\Psi}_l \mathbf{w}_l \leq P_{l,\max}, \ \forall l \in \mathcal{L}\,. \tag{4.20c}$$

The class of problems in the above form can be equivalently transformed into an SOCP using the techniques introduced in [95]. To elaborate further, we first recast (4.20b) as

$$\mathbf{w}^H \mathbf{Q}_k \mathbf{w} - 2\Re\{\mathbf{w}^H \mathbf{q}_k\} + \mathbf{w}^H \mathbf{\Theta} \mathbf{w} + q_k - \gamma_k \leq 0, \tag{4.21}$$

where $\mathbf{q}_k \triangleq [\mathbf{q}_{k,1}^T, \cdots, \mathbf{q}_{k,L}^T]^T$, $\mathbf{\Theta} \triangleq \mathtt{blkdiag}\{\mathbf{\Theta}_1, \cdots, \mathbf{\Theta}_L\}$ and $\mathbf{Q}_k$ is a block matrix with its $(l,m)^{\text{th}}$ blocked given by $\mathbf{Q}_{l,m}^{k}$. With the aid of (4.21), (4.20) can further be reformulated

as

$$\min_{\mathbf{w},\mathbf{t}_1,\mathbf{t}_2,\mathbf{t}_3,t_4} \quad \sum_{l=1}^{L} \frac{\mu P_{c,l}}{\|\mathbf{w}_l^{(n-1)}\|_2^2+\varepsilon} t_{1,l} + \sum_{l=1}^{L} t_{2,l} \tag{4.22a}$$

$$\text{s.t.} \quad t_{3,k}^2 + t_4^2 - \mathbf{q}_k \mathbf{Q}_k^{-1} \mathbf{q}_k + q_k - \gamma_k \le 0, \ \forall k \in \mathcal{K} \tag{4.22b}$$

$$t_{2,l} \le P_{l,\max}, \ \forall l \in \mathcal{L} \tag{4.22c}$$

$$\|\mathbf{w}_l\|_2^2 \le t_{1,l}, \ \forall l \in \mathcal{L} \tag{4.22d}$$

$$\|\mathbf{\Psi}_l^{1/2} \mathbf{w}_l\|_2^2 \le t_{2,l}, \ \forall l \in \mathcal{L} \tag{4.22e}$$

$$\|\mathbf{Q}^{1/2}\mathbf{w} - \mathbf{Q}_k^{-1/2}\|_2 \le t_{3,k}, \ \forall k \in \mathcal{K} \tag{4.22f}$$

$$\|\mathbf{\Theta}^{1/2}\mathbf{w}\|_2 \le t_4, \tag{4.22g}$$

where $\mathbf{t}_1 = [t_{1,1}, \cdots, t_{1,L}]^T$, $\mathbf{t}_2 = [t_{2,1}, \cdots, t_{2,L}]^T$, $\mathbf{t}_3 = [t_{3,1}, \cdots, t_{3,K}]^T$ and $t_4$ are auxiliary variables. It is observed that the objective function (4.22a) and the constraint (4.22c) are linear while (4.22f)–(4.22g) are in the form of SOCs. The remaining difficulties in solving (4.22) lie in (4.22b), (4.22d) and (4.22e), which are the so-called hyperbolic constraints [95]. To handle these constraints, we observe that for vector $\mathbf{x} \in \mathbb{C}^N$ and real scalars $y, z \ge 0$:

$$\|\mathbf{x}\|_2^2 \le yz \Longleftrightarrow \left\| \begin{bmatrix} 2\mathbf{x} \\ y - z \end{bmatrix} \right\|_2 \le y + z. \tag{4.23}$$

As a direct application of (4.23), (4.22b), (4.22d) and (4.22e) can be respectively reformulated as

$$\left\| \begin{bmatrix} 2t_{3,k} \\ 2t_4 \\ \gamma_k - q_k + \mathbf{q}_k \mathbf{Q}_k^{-1} \mathbf{q}_k - 1 \end{bmatrix} \right\|_2 \le \gamma_k - q_k + \mathbf{q}_k \mathbf{Q}_k^{-1} \mathbf{q}_k + 1 \tag{4.24}$$

$$\left\| \begin{bmatrix} 2\mathbf{w}_l \\ t_{1,l} - 1 \end{bmatrix} \right\|_2 \le t_{1,l} + 1 \tag{4.25}$$

$$\left\| \begin{bmatrix} 2\mathbf{\Psi}_l^{1/2} \mathbf{w}_l \\ t_{2,l} - 1 \end{bmatrix} \right\|_2 \le t_{2,l} + 1. \tag{4.26}$$

Substituting the above inequalities back into (4.22), the latter becomes a standard SOCP, which can then be efficiently solved by interior-point methods. To this end, one can rely

**Algorithm 4.1** Re-weighted iterative algorithm for energy minimization using BCD update

**Initialization**: $\mathbf{w}_l^{(0)}$ and $\beta_l = \frac{1}{\|\mathbf{w}_l^{(0)}\|_2^2+\epsilon}$ for all $l \in \mathcal{L}$

**repeat**

1) Update $u_k^{(n)}$ for all $k$ in parallel via (4.15) given fixed $\mathbf{W}_l = \texttt{dvec}(\mathbf{w}_l^{(n-1)})$ for all $l$
2) Update $\mathbf{w}_l^{(n)}$ by solving the SOCP (4.22) with fixed $u_k = u_k^{(n)}$ for all $k$
3) Update the weight $\beta_l = \frac{1}{\|\mathbf{w}_l^{(n)}\|_2^2+\varepsilon}$ for all $l$.

**until** *convergence*;

**Output**: $(\bar{\mathbf{w}}, \bar{\mathbf{u}})$

on state-of-the-art external software tools, see, e.g. [113].

With the aid of (4.15) and (4.22), the resultant iterative algorithm relying on the BCD and re-weighted $l_1$-norm approximation can be now summarized as seen in Algorithm 4.1, where $\texttt{dvec}(\cdot)$ denotes the matrix-vector reshaping. The computational complexity of Algorithm 4.1 is dominated by solving the SOCP (4.22). It is known that an SOCP can be solved at a worst-case complexity on the order of $\mathcal{O}(N_{\text{total}}^2 D_{\text{total}})$, where $N_{\text{total}}$ denotes the total number of optimization variables while $D_{\text{total}}$ denotes the total dimension of all the SOC constraints. To simply the analysis, let us assume that $N_l = N$ for all $l$. Hence, in (4.22), the total number of variables is $N^2L + 2L + K + 1$ and the total dimensions of (4.22b)–(4.22g) are $4K$, $L$, $(N+2)L$, $(N+2)L$, $(N+1)K$ and $N+1$, respectively. Using the above results and further assuming that $N$, $L$ and $K$ have the same order of magnitude, we can obtain the complexity of Algorithm 4.1 as $\mathcal{O}(N^5L^3 + N^5L^2K)$, i.e., the complexity of the algorithm increases on the order of $\mathcal{O}(N^5)$, $\mathcal{O}(L^3)$ and $\mathcal{O}(K)$ when considered in terms of its individual system parameters.

**Remark 4.1** (Comparison with prior works in the literature)**.** The design approach considered for our multi-user relay network is conceptually similar to that of [69] conceived for multi-cell downlink network, based on the QoS-constrained network power minimization. However, apart from the consideration of the different network topologies, the main difference between [69] and the present work is the choice of QoS metric, which inevitably leads to different solution approaches. In [69], the SINR-based constraints are imposed for the sake of guaranteeing a specific QoS level for each destination user. The SINR constraints can be equivalently written as an SOC constraint by applying a phase rotation, see, [69, eq. (10)] and therefore, the sub-problem at each iteration of the iterative algorithm is a convex SOCP. By contrast, in this work, we adopt the MSE (2.1) as our QoS metric, which is not

jointly convex in $(\mathbf{u}, \{\mathbf{W}\}_l)$. This fundamental difference subsequently leads to a different iterative algorithm from that of [69], i.e., the so-called *block coordinate re-weighted $l_1$-norm minimization*. Furthermore, in the context of the energy minimization problem considered, the convergence results of the conventional BCD algorithm [31–35, 114] are not directly applicable to the proposed algorithm. In view of this, as an additional contribution, we show below the convergence properties of the newly proposed algorithm by relying on a proposition, which has only been discovered with the advent of recent advances in the non-convex optimization theory [98]. □

### 4.2.2 Convergence Behavior

Since the original energy minimization problem of (4.6) is non-convex, it is necessary to analyze the convergence properties of Algorithm 4.1. Before proceeding further, let us take a closer look at the problem (4.6) in order to gain further insights into the proposed iterative algorithm. Similarly to the compressive sensing literature [96], a close approximation to the $l_0$-norm term of (4.6a) is the following concave function:

$$\begin{aligned}\sum_{l=1}^{L} \left\| \|\mathbf{W}_l\|_F^2 \right\|_0 P_{c,l} &= \sum_{l=1}^{L} \left\| \|\mathbf{w}_l\|_2^2 \right\|_0 P_{c,l} \\ &\approx \underbrace{\lambda_d \sum_{l=1}^{L} \log\left(1 + \|\mathbf{w}_l\|_2^2 \varepsilon^{-1}\right) P_{c,l}}_{\triangleq \mathcal{F}(\mathbf{w})} \end{aligned} \tag{4.27}$$

where $\mathbf{w} = [\mathbf{w}_1^T, \cdots, \mathbf{w}_L^T]^T$ and $\lambda_d = \frac{1}{\log(1+\varepsilon^{-1})}$. Then consider the following problem instead of (4.6):

$$\min_{(\mathbf{w},\mathbf{u})\in\Omega} \; \mathcal{F}'(\mathbf{w}, \mathbf{u}) \triangleq \mathcal{F}(\mathbf{w}) + \sum_{l=1}^{L} \mathbf{w}_l^H \mathbf{\Psi}_l \mathbf{w}_l + \mathbf{u}^H \mathbf{u}, \tag{4.28}$$

where for notational simplicity, we have $\Omega = \{(\mathbf{w}, \mathbf{u}) : (4.6\text{b}), (4.6\text{c})\}$. Since $\mathcal{F}(\cdot)$ is concave in $\|\mathbf{w}_l\|_2^2$, an efficient technique of solving (4.28) is the so-called MM algorithm [108]. The idea behind the MM algorithm is to first find a convex surrogate function, which majorizes the objective function and subsequently solves the original problem iteratively via a sequence of "convexified" sub-problems.

In (4.28), $\mathcal{F}(\cdot)$ is the only non-convex part, whose majorization function can simply be constructed from its first-order (i.e., linear) Taylor series expansion around a solution point $\mathbf{w}^{(n-1)}$ obtained from the previous iteration, i.e.,

$$\tilde{\mathcal{F}}(\mathbf{w};\mathbf{w}^{(n-1)}) = \mathcal{F}(\mathbf{w}^{(n-1)}) + \lambda_d \sum_{l=1}^{L} \frac{\|\mathbf{w}_l\|_2^2 - \|\mathbf{w}_l^{(n-1)}\|_2^2}{\|\mathbf{w}_l^{(n-1)}\|_2^2 + \varepsilon} P_{c,l} \geq \mathcal{F}(\mathbf{w}). \tag{4.29}$$

Then a modified block-type MM algorithm can be obtained for (4.28) by solving the following pair of sub-problems in a circular manner:

$$\textit{Step 1)}: \quad \mathbf{w}^{(n)} = \arg \min_{\mathbf{w}\in\Omega_{\mathrm{w}}^{(n)}} \tilde{F}(\mathbf{w};\mathbf{w}^{(n-1)}) + \sum_{l=1}^{L} \mathbf{w}_l^H \mathbf{\Psi}_l \mathbf{w}_l \tag{4.30}$$

$$\textit{Step 2)}: \quad \mathbf{u}^{(n)} = \arg \min_{\mathbf{u}\in\Omega_{\mathrm{u}}^{(n)}} \mathbf{u}^H \mathbf{u} \tag{4.31}$$

where for notational simplicity, we define the projection of $\Omega$ onto each design block at the $n^{\text{th}}$ iteration as $\Omega_{\mathrm{w}}^{(n)} = \{\mathbf{w} : (\mathbf{w}, \mathbf{u}^{(n-1)}) \in \Omega\}$ and $\Omega_{\mathrm{u}}^{(n)} = \{\mathbf{u} : (\mathbf{w}^{(n)}, \mathbf{u}) \in \Omega\}$. In Step 1), if we neglect the terms independent of $\mathbf{w}$ in $\tilde{\mathcal{F}}(\cdot;\cdot)$ and let $\lambda_d = \mu$, we observe that (4.30) becomes equivalent to the QCQP of (4.20). Furthermore, the solution to (4.31) is given by (4.15). Now we recognize that the proposed iteratively re-weighted algorithm in Algorithm 4.1 is essentially the above block-type MM algorithm, which solves the original optimization problem (4.6), however, with the aid of an approximated log-sum objective function.

Having established the above connection, we are ready to prove the convergence of Algorithm 4.1. To begin, let us first introduce the definition of the *Nash point* [98].

**Definition 4.1.** A point $(\bar{\mathbf{w}}, \bar{\mathbf{u}})$ is called a *Nash* point or *block coordinate-wise minimizer* of (4.28), if it satisfies the following Nash equilibrium conditions of (4.28):

$$\mathcal{F}'(\bar{\mathbf{w}}, \bar{\mathbf{u}}) \leq \mathcal{F}'(\mathbf{w}, \bar{\mathbf{u}}), \quad \forall \mathbf{w} \in \bar{\Omega}_{\mathrm{w}} \tag{4.32a}$$

$$\mathcal{F}'(\bar{\mathbf{w}}, \bar{\mathbf{u}}) \leq \mathcal{F}'(\bar{\mathbf{w}}, \mathbf{u}), \quad \forall \mathbf{u} \in \bar{\Omega}_{\mathrm{u}}, \tag{4.32b}$$

where $\bar{\Omega}_{\mathrm{w}} \triangleq \{\mathbf{w} : (\mathbf{w}, \bar{\mathbf{u}}) \in \Omega\}$ and similarly, $\bar{\Omega}_{\mathrm{u}} \triangleq \{\mathbf{u} : (\bar{\mathbf{w}}, \mathbf{u}) \in \Omega\}$.

**Remark 4.2** (On Nash point)**.** In general, the Nash point of a non-convex optimization problem is a generalization of a stationary point. If the feasible set $\Omega$ is separable, i.e., $\Omega$

is a Cartesian product given by $\Omega = \Omega_{\rm w} \times \Omega_{\rm u}$ where $\Omega_{\rm w}$ and $\Omega_{\rm u}$ are two convex subsets, then the above Nash equilibrium conditions become equivalent to the first-order optimality conditions [93], and subsequently $(\bar{\mathbf{w}}, \bar{\mathbf{u}})$ becomes a stationary point. In view of this, the Nash equilibrium conditions are in general weaker than the first-order optimality conditions. For problem (4.28), a stationary point must be a Nash point, but a Nash point is not necessarily a stationary point. □

With the above definition, the convergence properties of Algorithm 4.1 are formulated in the following theorem:

**Theorem 4.1.** Let $\left\{\left(\mathbf{w}^{(n)}, \mathbf{u}^{(n)}\right)\right\}$ be the solution sequence generated by Algorithm 4.1. Then any limit point of $\left\{\left(\mathbf{w}^{(n)}, \mathbf{u}^{(n)}\right)\right\}$ is a Nash point of problem (4.28), hence, satisfying the Nash equilibrium condition of (4.32).

*Proof:* The proof follows two steps. Firstly, we show that $\lim_{n\to\infty}\|\mathbf{u}^{(n)} - \mathbf{u}^{(n-1)}\|_2 = 0$ and $\lim_{n\to\infty}\|\mathbf{w}^{(n)} - \mathbf{w}^{(n-1)}\|_2 = 0$, the proof of which can be found in Appendix B.1. Then invoking [98, Theorem 2.3], it is readily shown that any limit point $(\bar{\mathbf{w}}, \bar{\mathbf{u}})$ of $\left\{\left(\mathbf{w}^{(n)}, \mathbf{u}^{(n)}\right)\right\}$ is a Nash point of (4.36). ■

Before concluding this subsection, the following remark is of interest.

**Remark 4.3** (On convergence, optimality and uniqueness of the solution)**.** In the current state-of-the-art of the BCD algorithm [31–34, 98], there are in general two classes of non-convex problems that this algorithm can efficiently solve, which are described below[2]. The first problem, denoted by $(\mathtt{P}_0)$ assumes the following form:

$$(\mathtt{P}_0): \quad \min_{\mathbf{x}} \; f(\mathbf{x}_1, \mathbf{x}_2) \quad \text{s.t. } \mathbf{x} = (\mathbf{x}_1, \mathbf{x}_2) \in \mathcal{X} \triangleq \mathcal{X}_1 \times \mathcal{X}_2,$$

where $f : \mathbb{C}^{m_1} \times \mathbb{C}^{m_2} \to \mathbb{R}$ is a continuously differentiable function, and $\mathcal{X}_i \subseteq \mathbb{C}^{m_i}, i = 1, 2$ are closed, nonempty convex subsets. Let $\{\mathbf{x}^{(n)}\}$ denote the sequence of intermediate solutions generated by the BCD algorithm. Then it has been proved in [31–34] that every limit point of $\{\mathbf{x}^{(n)}\}$ is a stationary point of $(\mathtt{P}_0)$.

Let us now proceed to a more general class of non-convex problems, which our energy

[2]To simplify the analysis, we limit our discussions on problems with two optimization variable blocks, i.e., $\mathbf{x}_1$ and $\mathbf{x}_2$. However, all the results and discussions can be extended to the case of multi-block variables.

minimization problem belongs to:

$$(\mathtt{P}_1): \quad \min_{\mathbf{x}} \ f(\mathbf{x}_1, \mathbf{x}_2) \quad \text{s.t. } \mathbf{x} \in \mathcal{X},$$

where $\mathcal{X} \subseteq \mathbb{C}^{m_1+m_2}$ is a closed and block multi-convex subset, i.e., $\mathcal{X}$ is convex with respect to each of $\mathbf{x}_1$ and $\mathbf{x}_2$, but not jointly convex in $\mathbf{x}$. In contrast to $(\mathtt{P}_0)$, now $\mathbf{x}_1$ and $\mathbf{x}_2$ are nonlinearly coupled in the constraints and therefore the feasible set $\mathcal{X}$ can no longer be expressed as a Cartesian product. The convergence behavior of the BCD algorithm for $(\mathtt{P}_1)$ has not been thoroughly investigated until very recently [98]. It has been proved that every limit point of $\{\mathbf{x}^{(n)}\}$ is a Nash point of $(\mathtt{P}_1)$, provided that $f(\cdot)$ is strongly convex with respect to each of $\mathbf{x}_1$ and $\mathbf{x}_2$ [98].

Despite the elegant convergence properties of the BCD algorithm, it is worth noting that due to the non-convex nature of both $(\mathtt{P}_0)$ and $(\mathtt{P}_1)$, there may exist many stationary or Nash points. At the time of writing, the question whether the algorithm converges to a unique stationary/Nash point remains an open issue in both the theoretical field of non-convex optimization [31–34, 53, 98] and in terms of its engineering applications [24, 35, 114]. Additionally, the local optimality of the solution obtained is not guaranteed. This is because for $(\mathtt{P}_0)$, by the definition of a stationary point, it can either be a locally optimal point or a saddle point. For $(\mathtt{P}_1)$, the current best effort is to prove the results in Theorem 4.1, where a Nash point is not necessarily a locally optimal point (see Definition 4.1). □

### 4.2.3 RRH Selection and UE Admission Control

Upon the convergence of Algorithm 4.1, when a solution $(\bar{\mathbf{w}}, \bar{\mathbf{u}})$ is obtained, the *group sparsity pattern* associated with the set of RRHs' weight vectors can be retrieved by computing the squared norm of each RRH's weight vector, yielding:

$$\boldsymbol{\mathcal{S}} = [\mathcal{S}_1, \cdots, \mathcal{S}_L] = [\|\bar{\mathbf{w}}_1\|_2^2, \cdots, \|\bar{\mathbf{w}}_L\|_2^2], \tag{4.33}$$

where $\mathcal{S}_l \triangleq \|\mathbf{w}_l\|_2^2$ is used as a sparsity indicator for RRH-$l$. In contrast to prior works [69, 70], hereby we determine the subset $\mathcal{A}$ consisting of the active RRHs in a single attempt by thresholding the sparsity indicator associated with each RRH. Specifically, we assume that when the sparsity indicator $\mathcal{S}_l$ of RRH-$l$ falls below a small threshold $\tau > 0$, RRH-$l$ is

switched off for the sake of energy efficiency, and vice versa. In this way, we obtain:

$$\mathcal{A} = \{l : \mathcal{S}_l \geq \tau, l \in \mathcal{L}\}\,. \tag{4.34}$$

Due to the reduced number of active RRHs participating in the relay-aided transmission, we have a reduced distributed diversity gain. Hence, the QoS constraints (4.6b) might be violated at some DUEs after deactivating selected RRHs. This infeasibility issue can be verified by reformulating (4.6) as a feasibility check problem. Specifically, only the subset $\mathcal{A}$ of active RRHs is now involved in (4.6) instead of all the $L$ RRHs and again, we can apply the concept of BCD update to solve the feasibility check problem. If a feasible solution is obtained, we subsequently further minimize the total transmission power of the active RRHs according to

$$\min_{\mathbf{w},\mathbf{u}} \quad \sum_{l\in\mathcal{A}} P_{t,l} \tag{4.35a}$$

$$\text{s.t.} \quad \mathtt{MSE}_k(u_k, \mathbf{w}) \leq \gamma_k,\ \forall k \in \mathcal{K} \tag{4.35b}$$

$$P_{t,l} \leq P_{l,\max},\ \forall l \in \mathcal{A} \tag{4.35c}$$

$$\mathbf{w}_l = \mathbf{0},\ \forall l \notin \mathcal{A}. \tag{4.35d}$$

Otherwise, if we fail to find a feasible solution, it becomes necessary to incorporate a mechanism which dynamically performs admission control for the end-users. A user admission control mechanism has been introduced in [115] for a multi-cell downlink scenario within a C-RAN, where the SINR is adopted as the QoS metric for each end-user. Following a similar design philosophy, hereby we propose a user admission control method for the multi-user relaying network within C-RAN.

To this end, we introduce a real-valued vector $\mathbf{z} = [z_1, \cdots, z_K]^T$, where $z_k$ is an abstract measure of the extent to which the QoS constraint at DUE-$k$ is violated. Motivated by the so-called phase-one method in [93], we can formulate the user admission control as the

**Algorithm 4.2** Iterative UE admission control algorithm

**Initialization**: $\mathcal{K} = \{1, 2, \cdots, K\}$
**while** $\mathbf{1}^T \bar{\mathbf{z}} > 0$ **do**
  1) Solve the feasibility problem (4.36) considering the QoS constraints (4.36b) for $k \in \mathcal{K}$ and obtain $\bar{\mathbf{z}}$ as its solution
  2) Set $\mathcal{K} = \mathcal{K} \backslash k$ where $k$ corresponds to the largest $\bar{z}_k$ among $\{\bar{z}_1, \cdots, \bar{z}_K\}$
**end**

following feasibility problem:

$$\min_{\mathbf{u},\mathbf{w},\mathbf{z}} \quad \mathbf{1}^T \mathbf{z} \tag{4.36a}$$

$$\text{s.t.} \quad \sum_{l\in\mathcal{A}} \sum_{m\in\mathcal{A}} \mathbf{w}_l^H \mathbf{Q}_{l,m}^k \mathbf{w}_m - \sum_{l\in\mathcal{A}}^{L} 2\Re\left\{\mathbf{w}_l^H \mathbf{q}_{k,l}\right\} + \sum_{l\in\mathcal{A}} \mathbf{w}_l^H \mathbf{\Theta}_{k,l} \mathbf{w}_l + q_k' \le z_k, \ k \in \mathcal{K} \tag{4.36b}$$

$$\mathbf{w}_l^H \mathbf{\Psi}_l \mathbf{w}_l \le P_{l,\max}, \ l \in \mathcal{A} \tag{4.36c}$$

$$\mathbf{z} \ge \mathbf{0}, \tag{4.36d}$$

where $q_k' = q_k - \gamma_k$. The above problem is always feasible and can be reformulated as a convex SOCP. The infeasibility indicator $z_k = 0$ reveals that the $k^{\text{th}}$ QoS constraint is satisfied, while $z_k > 0$ indicates the opposite. Therefore, all QoS constraints are satisfied if and only if $\mathbf{1}^T \mathbf{z} = 0$. Based on these observations, we then propose an iterative procedure to jointly check the feasibility and perform admission control, which is described by Algorithm 4.2. Specifically, at each iteration, we remove the specific DUE which has the highest infeasibility indicator, i.e., whose QoS constraint is the most "difficult" one to satisfy. The procedure is repeated, until the QoS constraints for all the remaining DUEs are satisfied. This admission control procedure can be readily incorporated into the RRH relaying design at the BBU pool, which essentially solves the inherent infeasibility issue associated with the RRH selection.

## 4.3 Simulation Results

We evaluate the performance of the joint RRH selection and AF relay optimization algorithms proposed in Section 4.2 based on computer simulations. In all simulations, we assume that the channel coefficients are generated as i.i.d. complex circular Gaussian vari-

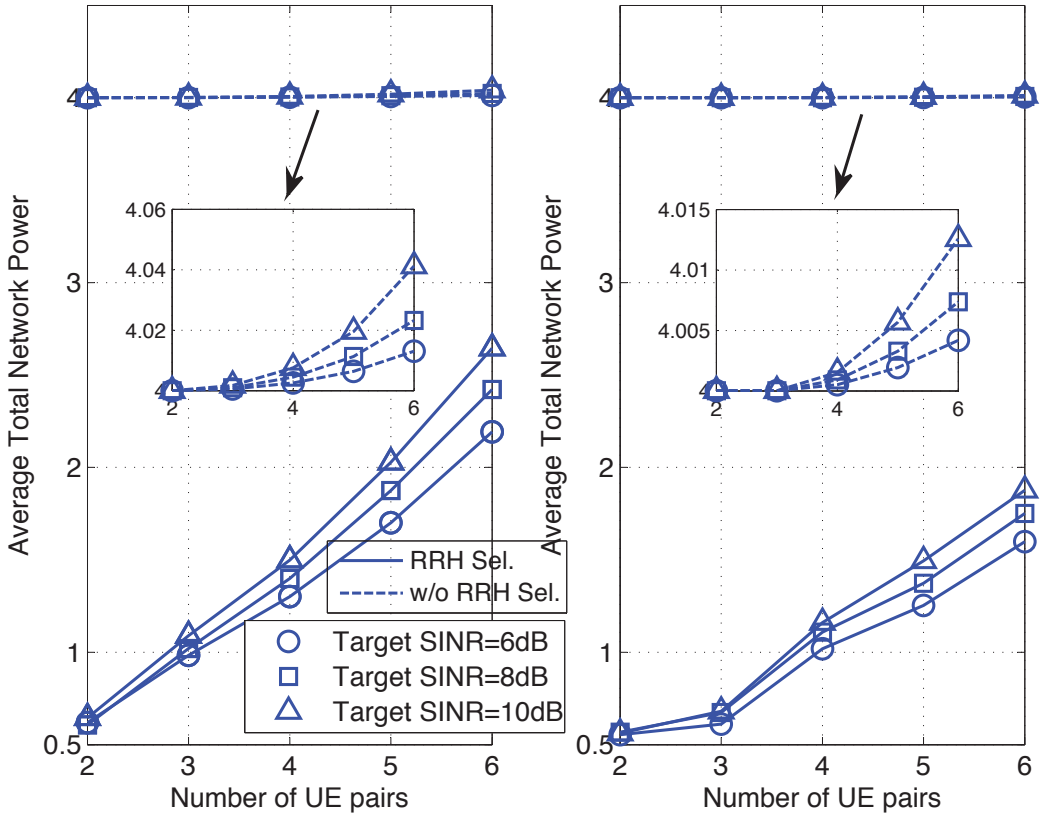



**Fig. 4.4** Comparison of power consumption of the relaying network using Algorithm 4.1 and of the one without RRH selection. Left subfigure: $N_l = 3$ $\forall l$. Right subfigure: $N_l = 4$ $\forall l$.

ables having a zero mean and a unit variance. For simplicity, all the RRHs are equipped with the same number of antennas $N_l$ while the static power consumption $P_{c,l}$ and transmit power budget $P_{t,l}$ are both set to 0.5. The noise variances at the RRHs and the DUEs are respectively set to $\sigma^2_{\mathrm{R},l} = 10^{-3}$ and $\sigma^2_{\mathrm{D},k} = 10^{-2}$. The MSE target $\gamma_k$ in (4.6) is defined in terms of the target SINR $\rho$, i.e., we have $\gamma_k = \frac{1}{\rho+1}$ for all $k \in \mathcal{K}$. Based on the observation that Algorithm 4.1 converges in about 10–15 iterations in all cases, Algorithm 4.1 is terminated after 20 iterations and $\varepsilon = 10^{-10}$ is adopted to avoid numerical instability. Each optimization problem in the form of (4.20) that is involved in Algorithms 4.1 and 4.2 is solved using the MATLAB-based interface YALMIP [116] along with the external solver MOSEK [102]. Each point or number in the subsequent figures and tables is obtained by aver-

aging the results of 200 independent realizations. The above simulation parameters remain fixed unless otherwise stated. We compare the performance of the following algorithms:

1. The proposed joint RRH selection and relay optimization, which is labeled as "RRH Selection" ("RRH Sel.");

2. The conventional collaborative relaying approach operating without RRH selection labeled as "w/o RRH Selection" ("w/o RRH Sel."), whose objective is to minimize the total transmission power of all the RRHs in the absence of sparsity-inducing $l_0$-norm [c.f. (4.6a)]. Mathematically, this problem can be written as

$$\min_{\mathbf{u},\{\mathbf{W}_l\}} \quad \sum_{l=1}^{L} P_{t,l} \tag{4.37a}$$

$$\text{s.t.} \quad \mathtt{MSE}_k(u_k, \{\mathbf{W}_l\}) \leq \gamma_k, \ \forall k \in \mathcal{K} \tag{4.37b}$$

$$P_{t,l} \leq P_{l,\max}, \ \forall l \in \mathcal{L}. \tag{4.37c}$$

We consider two simulation scenarios, namely, a generic relay network and a heterogeneous network (HetNet).

### 4.3.1 Generic Relay Network

In Fig. 4.4, we show the total relaying network's power dissipation for the different methods as a function of the number of UE pairs $K$. There are a total number of 8 RRHs in this scenario.A pair of multi-antenna settings, namely, one in which $N_l = 3$ for all $l$ and one in which $N_l = 4$ for all $l$, are considered under three different target SINRs at the DUEs. Our results reveal that in general it is inefficient to allow all the RRHs to transmit data because this requires a significant amount of static power. However, the proposed algorithm is capable of adaptively selecting the most appropriate RRHs for relaying by exploiting the knowledge of the spatial channel at a given time instance. To gain a deeper insight into this approach, we calculate the average number of RRHs that are switched off during the iterative procedure for different setup, i.e., different $K$ and $N_l$. The efficacy of the proposed algorithm becomes more evident from the corresponding results presented in Fig. 4.5, where on average 3 to 7 RRHs can be switched off for the sake of energy-efficient transmission.

Next, we examine the performance of the proposed DUE admission control mechanism

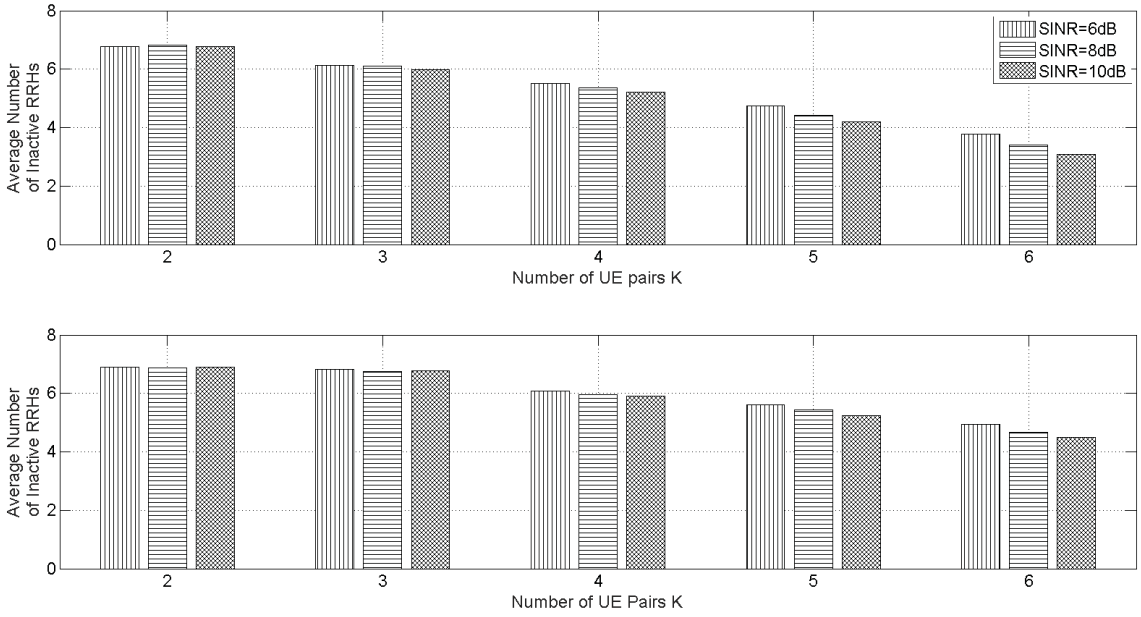



**Fig. 4.5** Average number of inactive RRHs with different number of UE pairs $K$ and target SINR $\rho$. Top subfigure: $N_l = 3$. Bottom subfigure: $N_l = 4$.

**Table 4.1** Number of non-scheduled DUEs for different $K$ and target SINR

| Number of UE pairs $K$ | 6 | | | 8 | | |
|---|---|---|---|---|---|---|
| Target SINR $\rho$ (dB) | 6 | 8 | 10 | 6 | 8 | 10 |
| Average number of excluded DUEs | 0.06 | 0.32 | 0.34 | 1.06 | 1.20 | 1.83 |

in Algorithm 4.2. Hereby we adopt a more aggressive energy-efficient design method, where only a limited number of "dominant" RRHs remain active. To achieve this, upon recovering the sparsity pattern of the RRH weight vectors, i.e., $\boldsymbol{\mathcal{S}}$ in (4.33), we compute the average of the sparsity indicators, that is, $\bar{\mathcal{S}} = \sum_{l=1}^{L} \mathcal{S}_l / L$, and then determine the set of active RRHs as $\mathcal{A} = \{l : \mathcal{S}_l \geq \eta \bar{\mathcal{S}}, l \in \mathcal{L}\}$, where $\eta$ is set to $\eta = 0.75$ in the simulations. In this way, the selected subset of active RRHs may not be capable of simultaneously maintaining an acceptable QoS at all DUEs. Then Algorithm 4.2 is invoked for iteratively removing the QoS constraints of specific DUEs from the optimization procedure. In Table 4.1, the average number of non-scheduled (excluded) DUEs is listed for different number of UE pairs $K$ and the target SINR $\rho$. It is clearly observed that the number of non-scheduled DUEs increases when more UEs requiring a higher QoS level are involved in the relay-assisted

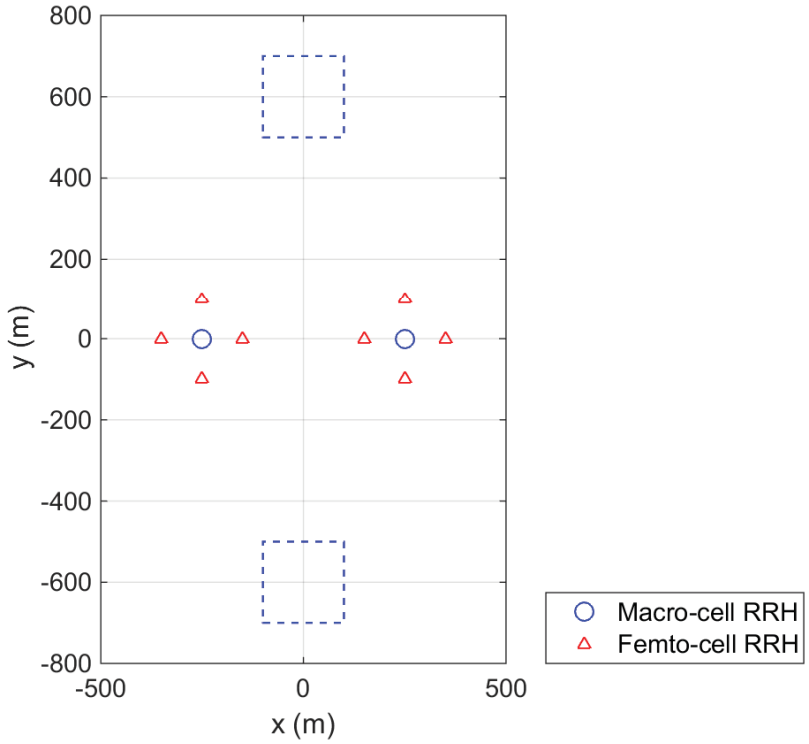



**Fig. 4.6** Layout of the HetNet considered in the simulations.

transmission.

### 4.3.2 HetNet Setup

In this subsection, we compare the performance of different methods in a HetNet setup, whose layout is sketched in Fig. 4.6. We consider a total of $L = 10$ RRHs, with two of them being macro-cell RRHs having a high power and the remaining ones are femto-cell RRHs having a relatively low power. A total of $K$ SUEs are randomly deployed in the upper dashed line rectangle, while the corresponding $K$ DUEs are randomly placed in the lower dashed line rectangle, based on the uniform distribution. The specific system parameters characterizing the HetNet are summarized in Table 4.2, where asymmetric pathloss models are considered for different types of RRHs. The performance of different methods is compared in terms of the average total network power and the number of inactive RRHs.

**Table 4.2** Simulation Parameters in the HetNet

| **Parameter** | **Value** |
|---|---|
| Pathloss between macro-cell RRH and UE at distance $d$ (km) | $128.1 + 37.6 \log_{10}(d)$ |
| Pathloss between femto-cell RRH and UE at distance $d$ (km) | $140.7 + 36.7 \log_{10}(d)$ |
| Small-scale fading $\mathbf{h}_{l,k}$ and $\mathbf{g}_{k,l}$ | Rayleigh fading |
| Noise power $\sigma_{\mathtt{R}}^2$ and $\sigma_{\mathtt{D}}^2$ | $-102$dBm |
| Static power $P_{c,l}$ | 5W for macro RRHs; 3W for femto RRHs |
| Relay power budget $P_{l,\max}$ | 5W for all RRHs |
| Antenna gain | 3dBi |

**Table 4.3** Average number of inactive (silent) RRHs for different schemes

| | | | Num. of UE Pairs $K$ | | | |
|---|---|---|---|---|---|---|
| | | Alg. | 2 | 4 | 6 | 8 |
| Relay Ant. Num. $N_l$ | 3 | RRH Sel. | 8.31 | 8 | 6.38 | 3.64 |
| | | w/o RRH Sel. | 0 | 0 | 0 | 0 |
| | 4 | RRH Sel. | 8.80 | 8.40 | 7.96 | 6.11 |
| | | w/o RRH Sel. | 0 | 0 | 0 | 0 |

In Fig. 4.7, the average network power consumption is shown as a function of the number of UE pairs $K$. Two different antenna array configurations are considered at the RRHs, namely, $N_l = 3$ and $N_l = 4$ for all $l \in \mathcal{L}$. It is observed that the proposed RRH selection achieves a significantly better energy efficiency than the conventional approach operating without RRH selection. Specifically, the performance gap is more evident, when the number of UEs is relatively small. When more UEs are involved in the transmission, the gap between the two approaches is reduced. This is because a large portion of the spatial diversity gain provided by the multiple RRHs' antenna arrays must be exploited to serve the additional end-users, hence more RRHs remain active in the relay-aided transmission. For a similar reason, the power consumption becomes lower, when additional antennas are employed at the RRHs, yielding a saving of about 35% for the case of $K = 6$ and 40% for the case of $K = 8$. The numbers of inactive RRHs in the HetNet using the two different approaches are shown in Table 4.3. For the case of $N_l = 3$, our proposed method yields a range of 3–8 inactive RRHs, while the numbers become 6–9 for the case of $N_l = 4$.

In summary, all the simulation results demonstrate the benefits of the proposed joint AF relaying optimization and RRH selection algorithm. In both generic relaying and

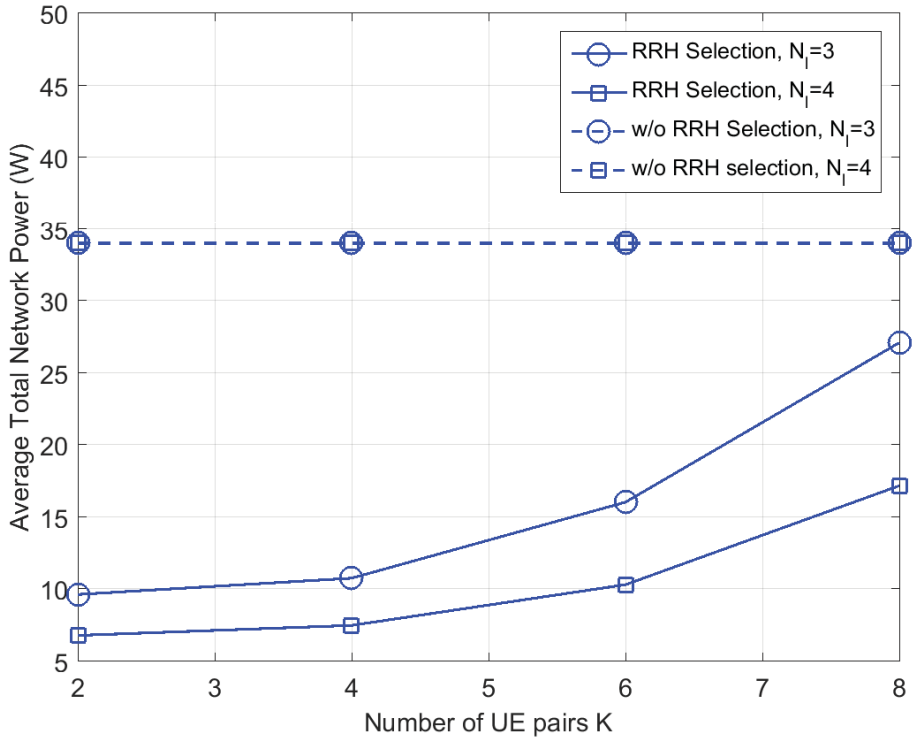



**Fig. 4.7** Average network power consumption versus the number of UE pairs in a HetNet.

heterogeneous network scenarios, the proposed method yields a significantly higher level of energy efficiency than that of its counterpart operating without RRH selection.

## 4.4 Conclusions

In this chapter, the problem of joint RRH relay selection and AF matrices design was investigated from a network energy minimization perspective for a multi-antenna multi-user relaying network within a C-RAN. Relying on the so-called re-weighted $l_1$ minimization and BCD-type methods, an iterative algorithm having a proven convergence was proposed for solving the original non-convex optimization problem. Based on the recovered group sparsity pattern associated with the RRHs' AF matrices, the set of active RRHs involved was then determined. To overcome the inherent infeasibility issue of the RRH selection, an iterative end-user admission control algorithm was proposed, which can be readily incorpo-

rated into the relaying optimization at the BBU pool. Our simulation results demonstrated the efficacy of the proposed algorithms, which significantly reduced the energy consumption of the C-RAN over that of a conventional cooperative relaying approach.

# Chapter 5

# Joint Secure Relaying and Artificial Noise Optimization

In this chapter, we consider a general wireless communication scenario, where a source (S) transmits its confidential data to a destination (D), assisted by a multi-antenna AF relay (R). Although a similar system model was studied in [87], the present thesis assumes that both phases of the two-hop transmission are overheard by a set of independent eavesdroppers (eves), as opposed to [87], where only the information leakage from the relays was considered. The power of S, the AF relaying matrix and the covariance matrix of the AN emitted by R have to be jointly optimized for protecting the message confidentiality. As an alternative to most of the prior contributions [74–77, 86–89], where the main focus has been on the maximization of the (worst-case) secrecy rate when either perfect or imperfect eavesdropper channel state information (ECSI) is available, we investigate the secrecy problem in MIMO relaying network from a practical communication performance perspective. Specifically, assuming that the ECSI errors reside in a predefined spherical region, we aim for maximizing the received SINR at D, subject to power constraints, while satisfying a set of *robust secrecy constraints* at eves. The main contributions of this chapter are listed below.

- The formulated optimization problem can be represented as a nonlinear non-convex SDP with a bilinear equality constraint due to the joint nature of the optimization

Parts of the materials in this chapter have been presented at the 2015 IEEE International Conference on Communications in London, U.K. [80], and published in IEEE Access [117].

variables. Such a class of problems are in general difficult to solve with tractable computational complexity. Towards this end, we propose a new penalized DC algorithmic framework specifically designed for the class of nonlinear non-convex SDP with bilinear equality constraints. One of the features of the proposed penalized DC algorithm is that it eliminates the need for a non-trivial feasible initialization as required by the conventional iterative algorithm [118] since finding such an initialization for a non-convex problem is in general a difficult task.

- We thoroughly investigate various aspects of the proposed penalized DC algorithm including its initialization, termination criteria, update rule of the penalty parameter and the convergence properties. Specifically, we explicitly prove that the solution sequence generated by the algorithm converges to a stationary point of the original problem.

- We further solve the secrecy constrained relaying problem by the proposed algorithm efficiently. To benchmark our solution approach, we also derive an upper bound for the secrecy constrained relaying problem by relying on the SDR technique along with an one-dimensional search algorithm. We show by numerical simulations that our proposed penalized DC algorithm is capable of achieving a performance close to the upper bound, however, at a significantly reduced complexity.

The rest of the chapter is organized as follows. Section 5.1 introduces the relay system model and formulates the secrecy-constrained robust relaying problem. In Section 5.2, we propose a new penalized DC algorithmic framework and characterize its convergence. We then invoke the proposed framework for solving the secure relaying problem in Section 5.3. In Section 5.4, a benchmarker relying on the SDR and one-dimensional exhaustive search is derived for comparison purposes. The performance of the proposed solution is quantified via numerical simulations in Section 5.5. Finally, we conclude in Section 5.6.

## 5.1 System Model and Problem Formulation

Consider a wireless network as depicted in Fig. 5.1, where source $\mathtt{S}$ communicates with destination $\mathtt{D}$, assisted by a trusted AF relay $\mathtt{R}$ operating in a half-duplex mode. The signals transmitted during the $\mathtt{S} \rightarrow \mathtt{R}$ and $\mathtt{R} \rightarrow \mathtt{D}$ hops are overheard by $K$ independent

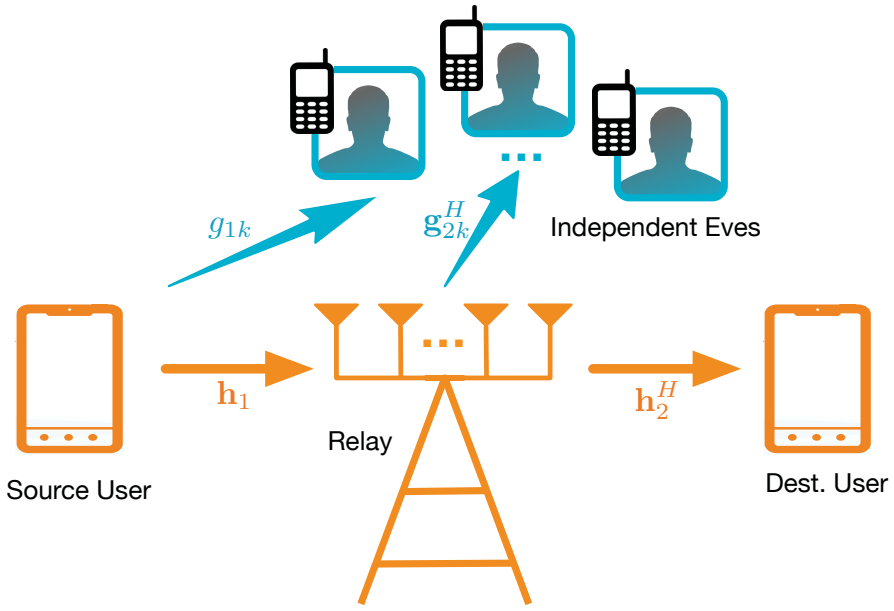



**Fig. 5.1** MIMO relay network in the presence of multiple single-antenna eves.

eves, $\mathtt{E}_k$ for $k \in \mathcal{K} \triangleq \{1, 2, \cdots, K\}$. We assume that S, D and $\mathtt{E}_k$, $\forall k \in \mathcal{K}$ are single-antenna UEs having limited signal processing capabilities and low power budgets. By contrast, R is equipped with $N_{\mathtt{R}} \geq 2$ antennas. It is assumed that no direct link is available between S–D due to the severe pathloss.

A narrowband flat-fading channel model is considered, where we denote the S–R channel by $\mathbf{h}_1 \in \mathbb{C}^{N_{\mathtt{R}} \times 1}$ and the Hermitian transpose of the R–D channel by $\mathbf{h}_2 \in \mathbb{C}^{N_{\mathtt{R}} \times 1}$. Let $s$ denote the S information symbol, modeled as a zero-mean Gaussian random variable with a power of $\sigma_{\mathtt{S}}^2 \leq P_{\mathtt{S}}$, where $P_{\mathtt{S}}$ denotes the S power budget. During the first transmission slot, the signal received at R is given by

$$\mathbf{z} = \mathbf{h}_1 s + \mathbf{n}_{\mathtt{R}}, \tag{5.1}$$

where $\mathbf{n}_{\mathtt{R}}$ is a zero-mean additive noise vector with covariance of $\sigma_{\mathtt{R}}^2 \mathbf{I}_{N_{\mathtt{R}}}$. Then R applies a linear AF transformation matrix $\mathbf{W} \in \mathbb{C}^{N_{\mathtt{R}} \times N_{\mathtt{R}}}$ to the received signal, and superimposes an AN vector onto the linearly processed signal. Hence, the signal to be forwarded to D is given by

$$\mathbf{r} = \mathbf{W}\mathbf{z} + \mathbf{v} = \mathbf{W}\mathbf{h}_1 s + \mathbf{W}\mathbf{n}_{\mathtt{R}} + \mathbf{v}, \tag{5.2}$$

where $\mathbf{v}$ denotes the AN vector with zero mean and covariance of $\mathbb{E}\{\mathbf{v}\mathbf{v}^H\} = \mathbf{\Psi} \succeq \mathbf{0}$ to be

optimized. The relay R has the power constraint of $\sigma_{\mathtt{S}}^2\|\mathbf{W}\mathbf{h}_1\|^2 + \sigma_{\mathtt{R}}^2\|\mathbf{W}\|_F^2 + \mathrm{Tr}(\mathbf{\Psi}) \leq P_{\mathtt{R}}$, where $P_{\mathtt{R}}$ denotes its power budget. During the second transmission slot, D receives the following signal:

$$y_{\mathtt{D}} = \mathbf{h}_2^H\mathbf{W}\mathbf{h}_1 s + \mathbf{h}_2^H\mathbf{W}\mathbf{n}_{\mathtt{R}} + \mathbf{h}_2^H\mathbf{v} + n_{\mathtt{D}}, \tag{5.3}$$

where $n_{\mathtt{D}}$ is an additive noise with zero mean and a variance of $\sigma_{\mathtt{D}}^2$.

We adopt, as a metric of transmission reliability, the received SINR at D given by

$$\mathtt{SINR}_{\mathtt{D}} = \frac{\sigma_{\mathtt{S}}^2|\mathbf{h}_2^H\mathbf{W}\mathbf{h}_1|^2}{\sigma_{\mathtt{R}}^2\|\mathbf{h}_2^H\mathbf{W}\|^2 + \mathbf{h}_2^H\mathbf{\Psi}\mathbf{h}_2 + \sigma_{\mathtt{D}}^2}. \tag{5.4}$$

During the transmission, each $\mathtt{E}_k$ is potentially capable of overhearing the signals transmitted both from S and R. Let $g_{1k}$ and $\mathbf{g}_{2k} \in \mathbb{C}^{N_{\mathtt{R}}\times 1}$, respectively, denote the S–$\mathtt{E}_k$ channel and the Hermitian transpose of the R–$\mathtt{E}_k$ channel. Then the signals observed by $\mathtt{E}_k$ from S and R, respectively, are given by

$$y_{\mathbb{E},k}^{\mathtt{S}} = g_{1k}s + n_{\mathbb{E},1k} \tag{5.5}$$

$$y_{\mathbb{E},k}^{\mathtt{R}} = \mathbf{g}_{2k}^H\mathbf{W}\mathbf{h}_1 s + \mathbf{g}_{2k}^H\mathbf{W}\mathbf{n}_{\mathtt{R}} + \mathbf{g}_{2k}^H\mathbf{v} + n_{\mathbb{E},2k}, \tag{5.6}$$

where $n_{\mathbb{E},1k}$ and $n_{\mathbb{E},2k}$ are additive noise terms with zero mean and a variance of $\sigma_{\mathbb{E},k}^2$. In our work, it is reasonable to assume that $\mathtt{E}_k$, for $k \in \mathcal{K}$, relies on selection combining (SC) of $y_{\mathbb{E},k}^{\mathtt{S}}$ and $y_{\mathbb{E},k}^{\mathtt{R}}$ for the sake of simpler exposition (However, our work can be extended to the case of maximum ratio combining (MRC), see Remark 5.4 for more justifications.). On this basis, the mutual information leakage to each $\mathtt{E}_k$ can therefore be expressed as

$$\mathtt{C}_{\mathbb{E},k}(\sigma_{\mathtt{S}}, \mathbf{W}, \mathbf{\Psi}) = \frac{1}{2}\max\left\{\log_2\left(1 + \frac{\sigma_{\mathtt{S}}^2|\mathbf{g}_{1k}|^2}{\sigma_{\mathbb{E},k}^2}\right), \log_2\left(1 + \frac{\sigma_{\mathtt{S}}^2|\mathbf{g}_{2k}^H\mathbf{W}\mathbf{h}_1|^2}{\sigma_{\mathtt{R}}^2\|\mathbf{g}_{2k}^H\mathbf{W}\|^2 + \mathbf{g}_{2k}^H\mathbf{\Psi}\mathbf{g}_{2k} + \sigma_{\mathbb{E},k}^2}\right)\right\}, \tag{5.7}$$

where the coefficient $\frac{1}{2}$ is due to the fact that the relay-assisted transmission requires a pair of orthogonal time slots in half-duplex mode.

In practice, due to the lack of explicit cooperation between the legitimate UEs and eves, only imperfect estimates of the ECSI may be available at the legitimate UEs. Like most of the prior contributions in the robust transceiver design literature, we model the unknown

ECSI by taking into account the error terms $\Delta g_{1k}$ and $\Delta \mathbf{g}_{2k}$, yielding:

$$g_{1k} = \hat{g}_{1k} + \Delta g_{1k},\ \mathbf{g}_{2k} = \hat{\mathbf{g}}_{2k} + \Delta \mathbf{g}_{2k}, \tag{5.8}$$

where $\hat{g}_{1k}$ and $\hat{\mathbf{g}}_{2k}$ denote the imperfect ECSI estimates, while again, $\Delta g_{1k}$ and $\Delta \mathbf{g}_{2k}$ represent the corresponding uncertainties. Hereby we assume that the ECSI errors lie in some predefined bounded sets, yielding:

$$\mathcal{G}_{1k} \triangleq \left\{ \Delta g_{1k} : |\Delta g_{1k}|^2 \leq \varepsilon_{1k} \right\} \tag{5.9}$$

$$\mathcal{G}_{2k} \triangleq \left\{ \Delta \mathbf{g}_{2k} : \|\Delta \mathbf{g}_{2k}\|^2 \leq \varepsilon_{2k} \right\}, \tag{5.10}$$

where $\varepsilon_{ik},\ i = 1, 2$ denotes the radius of the uncertainty region. The above bounded error model has been extensively used in robust MIMO transceiver optimization literature to capture the effects of channel estimation errors or quantization errors due to the finite-rate feedback, see, e.g., [59] for more details. The above error model is also applicable in some secure communication scenarios. A notable example is the D2D discovery and communication defined in 3GPP LTE Rel. 12 [2]. Each UE (including the potential `eve`s) periodically broadcasts its own beacon signals and listens to others using a subset of resources reserved for D2D operations. In this way, each UE is able to discover the presence of other UEs (including potential `eve`s in its proximity) and subsequently infers an imprecise ECSI estimate based on the channel reciprocity. In this case, the bounded error model can be invoked to quantify the channel estimation errors.

In a practical communication system, `S` can operate at a fixed data rate of $R_d$ with specific modulation and coding scheme (MCS), i.e., during a specific scheduling period in LTE. The objective of the secure relaying design is to jointly optimize $\sigma_{\mathtt{S}}$, $\mathbf{W}$ and $\mathbf{\Psi}$, subject to the power constraints, in order to maximize the received SINR at the legitimate end-user `D`, while satisfying a set of *robust secrecy constraints* at the `eve`s. Mathematically,

this problem can be formulated as

$$\max_{\sigma_{\mathtt{S}},\mathbf{W},\mathbf{\Psi}} \quad \mathtt{SINR}_{\mathtt{D}} \tag{5.11a}$$

$$\text{s.t.} \quad \mathtt{C}_{\mathbb{E},k}(\sigma_{\mathtt{S}},\mathbf{W},\mathbf{\Psi};\Delta g_{1k},\Delta\mathbf{g}_{2k}) \leq \kappa R_d,$$

$$\forall \Delta g_{1k} \in \mathcal{G}_{1k}, \Delta\mathbf{g}_{2k} \in \mathcal{G}_{2k}, k \in \mathcal{K} \tag{5.11b}$$

$$\sigma_{\mathtt{S}}^2\|\mathbf{W}\mathbf{h}_1\|^2 + \sigma_{\mathtt{R}}^2\|\mathbf{W}\|_F^2 + \mathrm{Tr}(\mathbf{\Psi}) \leq P_{\mathtt{R}} \tag{5.11c}$$

$$\sigma_{\mathtt{S}}^2 \leq P_{\mathtt{S}},\ \mathbf{\Psi} \succeq \mathbf{0}. \tag{5.11d}$$

In the above formulation, (5.11b) denotes the so-called *robust secrecy constraints*, which aims to guarantee the secrecy for all possible realizations of the ECSI errors $\Delta g_{1k}$ and $\Delta\mathbf{g}_{2k}$ within uncertainty regions as defined in (5.9) and (5.10), respectively. The parameter $\kappa$ is used to introduce more flexibility in controlling the security level of the communication. Before leaving this section, two important remarks are presented:

**Remark 5.1** (On the assumption of eves' receive combining)**.** It is worth pointing out that in contrast to prior contributions, hereby we assume information leakage during both the two-hop relay-assisted transmission. This more general assumption grants the eves the opportunities of enhancing their quality of reception via diversity combining. Two popular diversity combing schemes are available, namely, SC and MRC. The implement of MRC requires an accurate estimate of the phases of the received signals during the two stages of relay-assisted transmission. When channel estimation errors are in general invoked, the performance of MRC would significantly deteriorate. Additionally, to coherently combine the signals from the two-stage transmission, eves' clocks need to be perfectly synchronized to that of the legitimate network, which is quite challenging if the eves are not part of the legitimate network. It is observed in [119] that the MRC with two branches only yields marginal performance gain over the SC, however, at the expense of higher complexity. Hence, to bypass the aforementioned requirements, it is reasonable to assume that eves adopt the SC, also for the sake of lower hardware complexity. However, to better appreciate the generality of our proposed algorithm, in Remark 5.4 of Section IV, we will elaborate on how the proposed algorithm can be applied to solve the secure relaying problem when the MRC is employed by eves. □

**Remark 5.2** (On the problem formulation)**.** In the literature, another popular approach for improving the transmission secrecy is to maximize the secrecy capacity of the relay-

assisted network from the perspective of information theory. The latter in general relies on the underlying assumption that there exists a capacity-achieving coding scheme based on non-constructive random coding theorem. Such design approach is therefore useful as a benchmark from system design viewpoint. In practical communication systems whereby specific MCSs are used, e.g., 3GPP LTE-Advanced, it is better to consider a physical layer design approach, which can be readily incorporated into on-going standards. The proposed design approach well suits several use cases in LTE-Advanced such as the D2D broadcast scenarios. Specifically, by enforcing the mutual information leakage $\mathtt{C}_{\mathbb{E},k}$ to fall below the data rate of the legitimate UE, i.e., $\mathtt{C}_{\mathbb{E},k} < \kappa R_d$, eves are impossible to perfectly decode the confidential messages from the legitimate UEs. □

## 5.2 Theory: Penalized DC Algorithmic Framework

In this section, we propose a new penalized DC algorithmic framework, which aims to solve a class of nonlinear non-convex SDPs. Following some preliminary, we first present the framework, which can be considered as an evolutionary variant of the conventional DC framework [120]. However, the results of convergence analysis in the literature of conventional DC algorithm are not directly applicable to the proposed framework. Hence as a further contribution, we explicitly state the convergence properties of this new framework.

### 5.2.1 Preliminary

We first provide some definitions which will be used throughout the subsequent derivations of the algorithm.

**Definition 5.1** (Positive Semi-Definite (PSD)-Convex Mapping)**.** A matrix-valued mapping $\boldsymbol{\mathcal{F}}(\cdot) : \mathbb{C}^n \to \mathbb{H}^p$ is called *PSD*-convex on a convex subset $\Omega \subseteq \mathbb{C}^n$, if for all $\mathbf{x}, \mathbf{y} \in \Omega$ and $\theta$ with $0 \leq \theta \leq 1$, we have

$$\boldsymbol{\mathcal{F}}\left(\theta\mathbf{x} + (1-\theta)\mathbf{y}\right) \preceq \theta\boldsymbol{\mathcal{F}}(\mathbf{x}) + (1-\theta)\boldsymbol{\mathcal{F}}(\mathbf{y}). \tag{5.12}$$

The PSD-convex mapping is a generalization of a convex function by noting that any convex function with $f(\cdot) : \mathbb{C}^n \to \mathbb{R}$ is PSD-convex in conjunction with $p = 1$. The derivative of a matrix-valued mapping $\boldsymbol{\mathcal{F}}(\cdot)$ at a point $\mathbf{x}$ is defined as a linear mapping $\boldsymbol{\mathcal{DF}} : \mathbb{C}^n \to \mathbb{C}^{p\times p}$ given by

**Definition 5.2** (Directional Derivative of Matrix-Valued Mapping [121])**.** The directional derivative of a matrix-valued mapping $\boldsymbol{\mathcal{F}}$ at $\mathbf{x}$ is a linear mapping $\boldsymbol{\mathcal{DF}} : \mathbb{C}^n \to \mathbb{C}^{p\times p}$, which is defined by

$$\boldsymbol{\mathcal{DF}}(\mathbf{x})\mathbf{h} = \sum_{i=1}^{n} h_i \frac{\partial \boldsymbol{\mathcal{F}}}{\partial x_i}(\mathbf{x}),\ \forall \mathbf{h} \in \mathbb{C}^n. \tag{5.13}$$

For a given convex subset $\Omega \subseteq \mathbb{C}^n$, the matrix-valued mapping $\boldsymbol{\mathcal{F}}(\cdot)$ is said to be differentiable on $\Omega$ if its directional derivative $\boldsymbol{\mathcal{DF}}$ exits at every $\mathbf{x} \in \Omega$. For ease of discussion, we assume that all the functions and matrix-valued mappings are twice differentiable on their corresponding domains throughout the chapter.

The first-order condition for a PSD-convex mapping is given in the following proposition:

**Proposition 5.1** (First-Order Condition)**.** A mapping $\boldsymbol{\mathcal{F}}$ is PSD-convex if and only if for all $\mathbf{x}, \mathbf{y} \in \mathbb{C}^n$, the following inequality holds

$$\boldsymbol{\mathcal{F}}(\mathbf{y}) \succeq \boldsymbol{\mathcal{F}}(\mathbf{x}) + \boldsymbol{\mathcal{DF}}(\mathbf{x})(\mathbf{y} - \mathbf{x}). \tag{5.14}$$

Now we can proceed to the definition of a PSD DC mapping.

**Definition 5.3** (PSD DC Mapping)**.** A matrix-valued mapping $\boldsymbol{\mathcal{H}}(\cdot)$ is called a PSD DC mapping if $\boldsymbol{\mathcal{H}}$ can be represented as a difference of two PSD-convex mappings, i.e.,

$$\boldsymbol{\mathcal{H}}(\mathbf{x}) = \boldsymbol{\mathcal{F}}(\mathbf{x}) - \boldsymbol{\mathcal{G}}(\mathbf{x}). \tag{5.15}$$

Note that the concept of the PSD DC mapping generalizes the conventional scalar-valued DC function, i.e., $h(\mathbf{x}) = f(\mathbf{x}) - g(\mathbf{x})$.

### 5.2.2 Optimization of a PSD DC Program with Bilinear Matrix Equality Constraint

To simplify the exposition, in this subsection let us use matrix $\mathbf{X}$ as an optimization variable instead of using vector $\mathbf{x}$. The reason is that in the problem formulation of our interest, there exists a bilinear matrix equality constraint, as will seen below. However, it should be pointed out that any matrix variable $\mathbf{X} \in \mathbb{C}^{m\times n}$ can be equivalently expressed in the vector form, i.e., $\mathbf{x} \in \mathbb{C}^{mn\times 1}$ via $\mathbf{x} = \mathtt{vec}(\mathbf{X})$. Since the vectorization is a linear operation, the aforementioned PSD-convexity is preserved under linear operation.

We are interested in solving the following problem:

$$\min_{\mathbf{X}} \quad f_0(\mathbf{X}) - g_0(\mathbf{X}) \tag{5.16a}$$

$$\text{s.t.} \quad \boldsymbol{\mathcal{F}}_i(\mathbf{X}) - \boldsymbol{\mathcal{G}}_i(\mathbf{X}) \preceq \mathbf{0},\ i = 1, 2, \cdots, I-1 \tag{5.16b}$$

$$\mathbf{X}_2 = \mathbf{X}_0\mathbf{X}_1 \tag{5.16c}$$

$$\mathbf{X} \in \Omega, \tag{5.16d}$$

where the optimization variable $\mathbf{X}$ is defined as a tuple, $\mathbf{X} = (\mathbf{X}_0, \mathbf{X}_1, \mathbf{X}_2, \cdots, \mathbf{X}_{N-1})$ with $\Omega$ is a non-empty, closed convex set, $f_0(\cdot)$, $g_0(\cdot)$ are convex functions on $\Omega$, and $\boldsymbol{\mathcal{F}}_i(\cdot)$, $\boldsymbol{\mathcal{G}}_i(\cdot)$ are PSD-convex mappings on $\Omega$. For the ease of presentation, we use (5.16c) to represent that some of the optimization variables are nonlinearly coupled in the bilinear form. However, it can be conveniently extended to the case of $\mathbf{X}_i = \mathbf{X}_j\mathbf{X}_k$ for $i, j, k \in \{0, 1, \cdots, N-1\}$.

Next, we rely on the following lemma to show that (5.16) can be equivalently rewritten as a PSD DC program.

**Lemma 5.1** (Lemma 1 of [67])**.** Given $\mathbf{X}_0$, $\mathbf{X}_1$ and $\mathbf{X}_2$ of appropriate dimensions, which satisfy the following relation:

$$\mathbf{X}_2 = \mathbf{X}_0\mathbf{X}_1, \tag{5.17}$$

then the above matrix equality is equivalent to the following two constraints:

$$\begin{bmatrix} \mathbf{Y}_1 & \mathbf{X}_2 & \mathbf{X}_0 \\ \mathbf{X}_2^H & \mathbf{Y}_2 & \mathbf{X}_1^H \\ \mathbf{X}_0^H & \mathbf{X}_1 & \mathbf{I} \end{bmatrix} \succeq \mathbf{0} \tag{5.18}$$

$$\mathrm{Tr}\left(\mathbf{Y}_1\right) - \mathrm{Tr}\left(\mathbf{X}_0\mathbf{X}_0^H\right) \leq 0 \tag{5.19}$$

where $\mathbf{Y}_1$ and $\mathbf{Y}_2$ are auxiliary matrix variables with appropriate dimensions. □

It is observed that (5.18) is an LMI constraint and (5.19) is a DC constraint. Therefore, we can conveniently embed (5.18) into the convex subset $\Omega$ and its convexity remains unaffected. Additionally, since the DC function in (5.19) is a special case of the PSD DC mapping with $p = 1$, we can incorporate (5.19) into (5.16b), and re-express (5.16) as a

standard PSD DC program, which is defined as follows:

**Definition 5.4** (PSD DC Program)**.** A PSD DC program assumes the form of

$$\min_{\mathbf{x}} \quad \varphi_(\mathbf{x}) \triangleq f_0(\mathbf{x}) - g_0(\mathbf{x}) \tag{5.20a}$$

$$\text{s.t.} \quad \boldsymbol{\mathcal{F}}_i(\mathbf{x}) - \boldsymbol{\mathcal{G}}_i(\mathbf{x}) \preceq \mathbf{0},\ i \in \mathcal{I} \triangleq \{1, 2, \cdots, I\} \tag{5.20b}$$

$$\mathbf{x} \in \Omega, \tag{5.20c}$$

where $\mathbf{x}$ is a tuple, which collectively denotes all the optimization variables and auxiliary variables with appropriate linear transformation, i.e., $\mathbf{x} \triangleq (\mathtt{vec}(\mathbf{X}), \mathtt{vec}(\mathbf{Y}_1), \mathtt{vec}(\mathbf{Y}_2))$.

The above PSD DC program represents a generalization of the conventional DC program [118], where the DC inequality constraint, e.g., $f_i(\mathbf{x}) - g_i(\mathbf{x}) \leq 0$ is now extended to the *generalized inequality* $\preceq$ on the PSD cone. If the convex subset $\Omega$ is a polyhedral (which is true for most MIMO-aided transceiver optimization problems), the formulation in (5.20) can properly represent several classes of optimization problems:

- If at least one of $f_0$, $g_0$, $\boldsymbol{\mathcal{F}}_i$ and $\boldsymbol{\mathcal{G}}_i$ for $i \in \mathcal{I}$ is nonlinear, then (5.20) is a nonlinear SDP;
- If $g_0$ and $\boldsymbol{\mathcal{G}}_i$ for $i \in \mathcal{I}$ are linear, then (5.20) subsequently becomes a convex nonlinear SDP
- If at least one of $g_0$ and $\boldsymbol{\mathcal{G}}_i$ for $i \in \mathcal{I}$ are nonlinear, (5.20) represents a general nonlinear non-convex SDP.

### 5.2.3 Issues with the Conventional DC Algorithm

Since (5.20) can be considered as a direct extension of a conventional DC program involving only scalar-valued functions, a natural question arises as to whether the conventional DC algorithm developed in [118] is applicable to solving (5.20). Following the line of [118], an iterative algorithm can be developed for (5.20), where the key ingredient is to find a local linear approximation of the non-convex parts of the objective function (5.20a) and the PSD DC constraints (5.20b), i.e., $-g_0(\cdot)$ and $-\boldsymbol{\mathcal{G}}_i(\cdot)$, around the solution $\mathbf{x}^{(n-1)}$ obtained in the previous iteration, such that the resultant sub-problem becomes a convex SDP. The original non-convex problem can then be iteratively solved by a sequence of these

"convexified" SDPs. Assuming that $\mathbf{x}^{(n)}$ is a solution obtained at the $n^{\text{th}}$ iteration, the linearized sub-problem is then given by

$$\min_{\mathbf{x}} \quad f_0(\mathbf{x}) - g_0(\mathbf{x}^{(n)}) - \nabla g_0^T(\mathbf{x}^{(n)})(\mathbf{x}-\mathbf{x}^{(n)}) \tag{5.21a}$$

$$\text{s.t.} \quad \boldsymbol{\mathcal{F}}_i(\mathbf{x}) - \boldsymbol{\mathcal{G}}_i(\mathbf{x}^{(n)}) - \boldsymbol{\mathcal{D}}\boldsymbol{\mathcal{G}}_i(\mathbf{x}^{(n)})(\mathbf{x}-\mathbf{x}^{(n)}) \preceq \mathbf{0},\ i \in \mathcal{I} \tag{5.21b}$$

$$\mathbf{x} \in \Omega. \tag{5.21c}$$

Since $f_0$ and $\boldsymbol{\mathcal{F}}_i$ are convex function and mapping in $\mathbf{x}$ and the remaining terms are linear in $\mathbf{x}$, the above problem is a convex (nonlinear) SDP. The iterative algorithm therefore generates a sequence of intermediate solutions $\{\mathbf{x}^{(n)}\}_{n=0}^{\infty}$. Before proceeding to analyze the feasibility of $\{\mathbf{x}^{(n)}\}$, we first define the feasible set of the original PSD DC program in (5.20) as

$$\mathcal{D} \triangleq \{\mathbf{x} \in \Omega : \boldsymbol{\mathcal{F}}_i(\mathbf{x}) - \boldsymbol{\mathcal{G}}_i(\mathbf{x}) \preceq \mathbf{0},\ i \in \mathcal{I}\}, \tag{5.22}$$

and the relative interior of $\mathcal{D}$ as

$$\text{ri}(\mathcal{D}) \triangleq \{\mathbf{x} \in \text{ri}(\Omega) : \boldsymbol{\mathcal{F}}_i(\mathbf{x}) - \boldsymbol{\mathcal{G}}_i(\mathbf{x}) \prec \mathbf{0},\ i \in \mathcal{I}\}. \tag{5.23}$$

In order to guarantee that the obtained solution sequence $\{\mathbf{x}^{(n)}\}$ lies in the feasible set $\mathcal{D}$, a strictly feasible initialization, i.e., $\mathbf{x}^{(0)} \in \text{ri}(\Omega)$ is required by the conventional DC algorithm[1]. Hence, the following requirements are necessary for the conventional DC algorithm:

**Requirement 5.1.** A strictly feasible initialization $\mathbf{x}^{(0)} \in \text{ri}(\mathcal{D})$ is required by the conventional DC algorithm.

Subsequently, it is straightforward to have

**Requirement 5.2.** The relative interior of the feasible set is nonempty, i.e., $\text{ri}(\mathcal{D}) \neq \emptyset$.

We now explain the practical difficulties in satisfying the above requirements. As mentioned earlier, since $\mathcal{D}$ is a non-convex set, finding a strictly feasible initialization within a

[1]This is due to the fact that the first-order Taylor series expansion of the concave function $-\boldsymbol{\mathcal{G}}_i(\cdot)$ is its upper bound, such that we have $\text{F}_i(\mathbf{x}) - \boldsymbol{\mathcal{G}}_i(\mathbf{x}) \preceq \text{F}_i(\mathbf{x}) - \boldsymbol{\mathcal{G}}_i(\mathbf{x}^{(n)}) - \boldsymbol{\mathcal{D}}\boldsymbol{\mathcal{G}}_i(\mathbf{x}^{(n)})(\mathbf{x}-\mathbf{x}^{(n)}) \preceq \mathbf{0}$ for all $\mathbf{x} \in \Omega$.

non-convex set corresponds to the following non-convex feasibility search problem

$$\text{Find } \mathbf{x} \quad \text{s.t.}\, \mathbf{x} \in \mathcal{D}, \tag{5.24}$$

which in principle is not a simple task. Otherwise, if the algorithm starts with an infeasible point, then it can lead to further infeasibility issues during the successive iterations.

Additionally, the following claim also prevents the direct application of the conventional DC algorithm to (5.20).

**Claim 5.1.** The relative interior of the feasible set of (5.20) is empty, i.e., $\text{ri}(\mathcal{D}) = \emptyset$.

*Proof:* We show by contradiction. Recall that a strictly feasible solution to (5.20) has to satisfy

$$\text{Tr}(\mathbf{Y}_1) - \text{Tr}(\mathbf{X}_0\mathbf{X}_0^H) < 0. \tag{5.25}$$

By applying the Schur complement to (5.18), we have

$$\begin{aligned} &\begin{bmatrix} \mathbf{Y}_1 & \mathbf{X}_2 \\ \mathbf{X}_2^H & \mathbf{Y}_2 \end{bmatrix} - \begin{bmatrix} \mathbf{X}_0 \\ \mathbf{X}_1^H \end{bmatrix} \begin{bmatrix} \mathbf{X}_0^H & \mathbf{X}_1 \end{bmatrix} \succeq \mathbf{0} \\ &\quad \Longleftrightarrow \begin{bmatrix} \mathbf{Y}_1 & \mathbf{X}_2 \\ \mathbf{X}_2^H & \mathbf{Y}_2 \end{bmatrix} - \begin{bmatrix} \mathbf{X}_0\mathbf{X}_0^H & \mathbf{X}_0\mathbf{X}_1 \\ \mathbf{X}_1^H\mathbf{X}_0^H & \mathbf{X}_1^H\mathbf{X}_1 \end{bmatrix} \succeq \mathbf{0} \\ &\quad \Longrightarrow \mathbf{Y}_1 \succeq \mathbf{X}_0\mathbf{X}_0^H, \end{aligned} \tag{5.26}$$

which obviously contradicts (5.25). Therefore, we must have

$$\text{Tr}(\mathbf{Y}_1) - \text{Tr}(\mathbf{X}_0\mathbf{X}_0^H) = 0, \tag{5.27}$$

which implies that $\text{ri}(\mathcal{D}) = \emptyset$. ■

Based on the above analysis, it is known that both requirements of the conventional DC algorithm cannot be satisfied. Motivated by the latter, we shall propose a new approach where the concept of penalized DC algorithm is developed for the considered PSD DC program. The proposed penalized DC algorithm, which can be considered as an evolutionary variant of the conventional DC algorithm, can solve a wider range of PSD DC programs. In particular, it eliminates the requirements of a non-trivial initialization and of a feasible

set with non-empty relative interior.

### 5.2.4 Penalized PSD DC Algorithmic Framework

Instead of solving (5.21), hereby we introduce a set of matrix auxiliary variables $\{\mathbf{S}_i\}_{i=1}^I$ and penalize (5.21a) with a linear regularization term, i.e.,

$$\begin{aligned}
\min_{\mathbf{x},\mathbf{S}} \quad & \hat{\varphi}^{(n)}(\mathbf{x},\mathbf{S};\mathbf{x}^{(n)}) \\
& \triangleq f_0(\mathbf{x}) - g_0(\mathbf{x}^{(n)}) - \nabla g_0^T(\mathbf{x}^{(n)})(\mathbf{x}-\mathbf{x}^{(n)}) + \tau^{(n)} \sum_{i=1}^{I} \mathrm{Tr}(\mathbf{S}_i) && \text{(5.28a)} \\
\text{s.t.} \quad & \boldsymbol{\mathcal{F}}_i(\mathbf{x}) - \boldsymbol{\mathcal{G}}_i(\mathbf{x}^{(n)}) - \boldsymbol{\mathcal{D}}\boldsymbol{\mathcal{G}}_i(\mathbf{x}^{(n)})(\mathbf{x}-\mathbf{x}^{(n)}) \preceq \mathbf{S}_i && \text{(5.28b)} \\
& \mathbf{S}_i \succeq \mathbf{0},\ i \in \mathcal{I} && \text{(5.28c)} \\
& \mathbf{x} \in \Omega, && \text{(5.28d)}
\end{aligned}$$

where $\tau^{(n)} \geq 0$ denotes the weight associated with the penalty term at the $n^{\text{th}}$ iteration and $\mathbf{S}$ collectively denotes $\mathbf{S} \triangleq (\mathbf{S}_1, \cdots, \mathbf{S}_I)$. The auxiliary variable $\mathbf{S}_i \in \mathbb{H}^{p_i}$ can be viewed as an abstract measure of the extent to which the $i^{\text{th}}$ constraint in (5.21b) is violated. Specifically, $\mathrm{Tr}(\mathbf{S}_i) = 0$ reveals that the $i^{\text{th}}$ constraint is satisfied while $\mathrm{Tr}(\mathbf{S}_i) > 0$ indicates the opposite. Therefore, a feasible solution $\mathbf{x} \in \Omega$ is found if

$$\sum_{i=1}^{I} \mathrm{Tr}(\mathbf{S}_i) = 0. \tag{5.29}$$

With the introduction of the penalized sub-problem (5.28), we now develop an iterative procedure for solving the PSD DC program (5.20). The rationale of the proposed penalized DC algorithm is that it starts with an arbitrary point within the convex subset $\Omega$, i.e., $\mathbf{x}^{(0)} \in \Omega$, as opposed to $\mathbf{x}^{(0)} \in \mathrm{ri}(\mathcal{D})$ (hence possibly infeasible), and a small penalty $\tau$ such that it facilitates a fast descent of the objective function at the beginning while the constraints are temporarily allowed to be violated, i.e, $\mathbf{S}_i \succ \mathbf{0}$. As iterations evolve, the value of $\tau$ gradually increases according to some designed rule in order to enforce the solution to be closer to and finally lie in the feasible region $\mathcal{D}$.

The penalized DC algorithm, which iteratively solves a sequence of sub-problems (5.28) with a specifically designed updating rule of $\tau$ is then described as Algorithm 1. Below we

**Algorithm 5.1** Penalized DC Algorithm

Initialization: An initial point $\mathbf{x}^{(0)} \in \Omega$, $\tau^{(0)} > 0$, $\delta_1 > 0$ and $\delta_2 > 0$. Set $n = 0$.

**repeat**

1. *Convexify*: Compute the first-order approximates

$$g_0(\mathbf{x}) \approx g_0\left(\mathbf{x}^{(n)}\right) + \nabla g_0^T\left(\mathbf{x}^{(n)}\right)\left(\mathbf{x} - \mathbf{x}^{(n)}\right)$$
$$\boldsymbol{\mathcal{G}}_i(\mathbf{x}) \approx \boldsymbol{\mathcal{G}}_i\left(\mathbf{x}^{(n)}\right) + \boldsymbol{\mathcal{DG}}_i\left(\mathbf{x}^{(n)}\right)\left(\mathbf{x} - \mathbf{x}^{(n)}\right)$$

2. *Solve*: Compute $\mathbf{x}^{(n+1)}$ by solving (5.28)
3. *Update* $\tau$: Obtain the dual variable $\boldsymbol{\Phi}_i^{(n+1)}$ associated with (5.28b) and set

$$\tau^{(n+1)} = \begin{cases} \tau^{(n)} & \text{if } \tau^{(n)} \geq r^{(n)} \\ \tau^{(n)} + \delta_2 & \text{if } \tau^{(n)} < r^{(n)} \end{cases} \tag{5.30}$$

where

$$r^{(n)} \triangleq \min\left\{ \|\mathbf{x}^{(n+1)} - \mathbf{x}^n\|^{-1}, \lambda_{\max}\Big[\sum_{i=1}^{I} \boldsymbol{\Phi}_i^{(n+1)}\Big] + \delta_1 \right\}$$

4. *Update iteration*: $n \leftarrow n + 1$

**until** Termination criterion is satisfied or a maximum number of iterations are reached

Output: The optimized $\mathbf{x}^*$.

discuss a few important implementation aspects of Algorithm 1.

*1) Initialization:* Instead of finding an initialization within the relative interior of a non-convex feasible set [c.f. (5.23)], i.e., $\mathbf{x}^{(0)} \in \mathrm{ri}(\mathcal{D})$, Algorithm 1 can now be initialized with a point $\mathbf{x}^{(0)} \in \Omega$, which corresponds to a more computationally efficient convex feasible search problem. For implementation, one may rely on the general-purpose optimization solvers to find $\mathbf{x}^{(0)}$. More importantly, in many practical problems, $\mathbf{x}^{(0)}$ can be easily found by exploiting the specific structure of the convex subset $\Omega$ in that problem, (see the considered secure relaying design problem in Section 5.3).

*2) Termination Criterion:* In practical implementation, Algorithm 1 needs to be terminated within a maximum of number iterations. Thus, a reasonable termination criterion is that the successive difference in the solution becomes small, i.e., $\|\mathbf{x}^{(n+1)} - \mathbf{x}^{(n)}\| \leq \delta$ and $\mathbf{x}^{(n)}$ is (nearly) feasible, i.e., $\sum_{i=1}^{I} \mathrm{Tr}(\mathbf{S}_i) \approx 0$. If the criterion cannot be satisfied within a maximum number of iterations, we claim that the algorithm fails to find a feasible solution given a limited time frame.

*3) On the Updating Rule* (5.30)*:* The updating rule of $\tau$ is motivated by the theory of

exact penalty function methods for nonlinear optimization problem [122, 123]. The theory suggests that if the penalty $\tau$ is larger than all the dual variables $\{\mathbf{\Phi}_i\}$ associated with (5.28b) (in our case, it is in the form of PSD ordering), i.e., $\tau\mathbf{I} \succeq \mathbf{\Phi}_i$ for all $i$, then (5.28) and (5.21) become equivalent. Also from the definition of $r^{(n)}$ below (5.30), we see that the unboundness of $\{\tau^{(n)}\}$ leads to the unboundness of $\{\mathbf{\Phi}_i^{(n)}\}$ and $\|\mathbf{x}^{(n+1)} - \mathbf{x}^n\| \to 0$. This key property will be exploited later in proving the convergence of Algorithm 5.1.

*4) Solving the Convex Sub-Problem* (5.28)*:* As mentioned earlier, (5.28) is a general nonlinear convex SDP, which can be solved by a general interior-point method. To our best knowledge, the external solvers supporting a general nonlinear SDP are still limited, i.e., some widely-used solvers such as `SeDuMi` and `MOSEK` do not support nonlinear SDPs at the current stage while `PENLAB` is the only public nonlinear SDP solver. However, many MIMO transceiver optimization problems exhibit some common structures. Specifically:

1. The convex subset $\Omega$ can be represented by a finite number of LMIs, i.e.,

$$\Omega \triangleq \{\mathbf{x} : \boldsymbol{\mathcal{A}}_l(\mathbf{x}) + \mathbf{C}_l \succeq \mathbf{0}, l = 1, \cdots, L\}, \tag{5.31}$$

   where $\boldsymbol{\mathcal{A}}_l(\mathbf{x})$ is a linear mapping of $\mathbf{x}$.

2. The mappings $\boldsymbol{\mathcal{F}}_i(\mathbf{x})$ for $i \in \mathcal{I}$ are so-called Schur PSD-convex mappings, which assumes the form of

$$\boldsymbol{\mathcal{F}}_i(\mathbf{x}) \triangleq \boldsymbol{\mathcal{S}}_i(\mathbf{x})\boldsymbol{\mathcal{R}}_i^{-1}(\mathbf{x})\boldsymbol{\mathcal{S}}_i^H(\mathbf{x}) - \boldsymbol{\mathcal{Q}}_i(\mathbf{x}), \tag{5.32}$$

   where $\boldsymbol{\mathcal{R}}_i = \boldsymbol{\mathcal{R}}_i^H$, $\boldsymbol{\mathcal{Q}}_i(\mathbf{x}) = \boldsymbol{\mathcal{Q}}_i^H(\mathbf{x})$ and $\boldsymbol{\mathcal{S}}_i(\mathbf{x})\boldsymbol{\mathcal{S}}_i^H(\mathbf{x})$ are linear mappings of $\mathbf{x}$ and $\boldsymbol{\mathcal{R}}_i \succ \mathbf{0}$;

3. The function $f_0(\mathbf{x})$ in the objective function is quadratic in $\mathbf{x}$:

$$f_0(\mathbf{x}) = \mathbf{x}^H \mathbf{B}\, \mathbf{x} + 2\,\mathrm{Re}\{\mathbf{b}^H\, \mathbf{x}\} + c. \tag{5.33}$$

Below we show that the sub-problem (5.20) with the above structure can be equivalently transformed into a standard SDP, which can be efficiently solved by state-of-the-art optimization tools. The transformation simply invokes the Schur complement and the introduction of auxiliary variables. In this case, one can transform (5.28) into a standard

SDP:

$$\min_{\substack{\mathbf{x},\mathbf{S},\\ t,\{\mathbf{T}_i\}}} \quad t - \nabla g_0^T(\mathbf{x}^{(n)})\,\mathbf{x} + \tau^{(n)} \sum_{i=1}^{I} \mathrm{Tr}(\mathbf{S}_i) \tag{5.34a}$$

$$\text{s.t.} \quad \mathbf{T}_i - \boldsymbol{\mathcal{G}}_i(\mathbf{x}^{(n)}) - \boldsymbol{\mathcal{D}}\boldsymbol{\mathcal{G}}_i(\mathbf{x}^{(n)})(\mathbf{x} - \mathbf{x}^{(n)}) - \mathbf{S}_i \preceq \mathbf{0} \tag{5.34b}$$

$$\boldsymbol{\mathcal{A}}_l(\mathbf{x}) + \mathbf{C}_l \succeq \mathbf{0},\ l = 1, \cdots, L \tag{5.34c}$$

$$\begin{bmatrix} \boldsymbol{\mathcal{Q}}_i(\mathbf{x}) + \mathbf{T}_i & \boldsymbol{\mathcal{S}}_i(\mathbf{x}) \\ \boldsymbol{\mathcal{S}}_i^H(\mathbf{x}) & \boldsymbol{\mathcal{R}}_i \end{bmatrix} \succeq \mathbf{0},\ i \in \mathcal{I} \tag{5.34d}$$

$$\begin{bmatrix} \mathbf{x} \\ 1 \end{bmatrix}^H \begin{bmatrix} \mathbf{B} & \mathbf{b} \\ \mathbf{b}^H & c - t \end{bmatrix} \begin{bmatrix} \mathbf{x} \\ 1 \end{bmatrix} \preceq \mathbf{0}, \tag{5.34e}$$

where $t$ and $\{\mathbf{T}_i\}$ are auxiliary variables.

### 5.2.5 Convergence Analysis of the Penalized DC Algorithm

Since Algorithm 5.1 is designed to start with a possibly infeasible initialization, the iterative procedure may admit an infeasible final solution to the original PSD DC program (5.20). Therefore, two important aspects regarding the convergence of Algorithm 5.1 need to be examined:

1. whether the solution generated by Algorithm 1 is feasible to the PSD DC program (5.20).

2. whether the convergence properties of conventional DC algorithm still hold for the penalized DC algorithm.

In this subsection, the convergence properties of Algorithm 1 are analytically established. Let $\bar{\mathbf{x}}$ be a point within the convex subset $\Omega$, i.e., $\bar{\mathbf{x}} \in \Omega$. The PSD DC constraint (5.20b) at $\bar{\mathbf{x}}$ is called *inactive* if the strict inequality holds, that is, $\boldsymbol{\mathcal{F}}_i(\bar{\mathbf{x}}) - \boldsymbol{\mathcal{G}}_i(\bar{\mathbf{x}}) \prec \mathbf{0}$. Otherwise, the PSD DC constraint is called *active*, i.e., $\boldsymbol{\mathcal{F}}_i(\bar{\mathbf{x}}) - \boldsymbol{\mathcal{G}}_i(\bar{\mathbf{x}}) \not\prec \mathbf{0}$. Let us denote the set of active constraints at $\bar{\mathbf{x}}$ by

$$\mathcal{U}(\bar{\mathbf{x}}) \triangleq \left\{ i \in \mathcal{I} \middle| \boldsymbol{\mathcal{F}}_i(\bar{\mathbf{x}}) - \boldsymbol{\mathcal{G}}_i(\bar{\mathbf{x}}) \not\prec \mathbf{0} \right\}. \tag{5.35}$$

We call a vector $\mathbf{h} \in \texttt{cone}(\Omega - \bar{\mathbf{x}})$ a *feasible direction* to (5.20) at $\bar{\mathbf{x}}$ if we have

$$(\boldsymbol{\mathcal{D}}\boldsymbol{\mathcal{F}}_i(\bar{\mathbf{x}}) - \boldsymbol{\mathcal{D}}\boldsymbol{\mathcal{G}}_i(\bar{\mathbf{x}}))\,\mathbf{h} \prec \mathbf{0},\ \forall i \in \mathcal{U}(\bar{\mathbf{x}}). \tag{5.36}$$

We now make our first assumption, which is called the *extended Mangasarian-Fromovitz constraint qualification* (MFCQ) [124]:

**Assumption 5.1.** For any $\bar{\mathbf{x}} \in \Omega$, there exists a feasible direction $\mathbf{h} \in \texttt{cone}(\Omega - \bar{\mathbf{x}})$ to (5.20).

The extended MFCQ is a quite common constraint qualification in nonlinear optimization theory such that it guarantees the KKT necessary conditions to hold at a local point[2]. A geometric interpretation of the extended MFCQ can be described as follows. The gradients of the active inequality constraints (recall that $\boldsymbol{\mathcal{F}}_i(\bar{\mathbf{x}}) - \boldsymbol{\mathcal{G}}_i(\bar{\mathbf{x}}) \not\prec \mathbf{0}$) at $\bar{\mathbf{x}}$ form a pointed cone, and there exists a feasible direction in this cone that is tangent to the surface formed by active inequality constraints.

In addition, we also make the following common assumptions:

**Assumption 5.2.** $\Omega$ is bounded and the objective function $\varphi(\mathbf{x}) = f_0(\mathbf{x}) - g_0(\mathbf{x})$ is bounded from below on $\Omega$.

Assumption 5.2 is a mild assumption from practical perspective. Specifically, $\Omega$ is bounded due to the power constraints imposed in the design problem, while the objective function is usually a performance metric such as the SINR or MSE, which is lower-bounded by zero.

Before formally stating the convergence theorem, we first present the following lemma, which shows that $\Delta\,\mathbf{x}^{(n)} \triangleq \mathbf{x}^{(n+1)} - \mathbf{x}^{(n)}$ is a descent direction of the PSD DC program (5.20). The latter is a key property in proving the convergence of Algorithm 5.1.

**Lemma 5.2.** Let us denote the penalized objective function by $\hat{\varphi}^{(n)}(\mathbf{x}, \mathbf{S}) \triangleq f_0(\mathbf{x}) - g_0(\mathbf{x}) + \sum_{i=1}^{I} \tau^{(n)} \operatorname{Tr}(\mathbf{S}_i)$. Suppose that $\{\mathbf{x}^{(n)},\ n = 0, 1, \cdots\}$ is a sequence of solutions generated by Algorithm 5.1. Then we have:

[2]A similar example in convex optimization theory is that the Slater condition guarantees that the sufficient KKT conditions hold at some points for a convex problem.

1. The following inequality holds for $n \geq 0$:

$$\hat{\varphi}^{(n)}(\mathbf{x}^{(n)}, \mathbf{S}^{(n)}) - \hat{\varphi}^{(n)}(\mathbf{x}^{(n+1)}, \mathbf{S}^{(n+1)}) \geq \frac{\rho_f + \rho_g}{2} \|\mathbf{x}^{(n+1)} - \mathbf{x}^{(n)}\|^2, \tag{5.37}$$

   where $\rho_f$ and $\rho_g$ denote the convexity parameters of $f_0$ and $g_0$, respectively, i.e., $\rho_f, \rho_g > 0$ if $f_0$ and $g_0$ are strongly convex function and $\rho_f, \rho_g = 0$ otherwise.

2. If either $f_0$ or $g_0$ is strongly convex, i.e., $\rho_f + \rho_g > 0$, then $\Delta\,\mathbf{x}^{(n)}$ is a sufficient descent direction of (5.20) for all $n \geq 0$.

*Proof:* Please see Appendix C.1. ■

Subsequently, we assume that

**Assumption 5.3.** Either $f_0(\cdot)$ or $g_0(\cdot)$ is strongly convex.

The above assumption is needed to ensure $\Delta\,\mathbf{x}^{(n)}$ is a sufficient descent direction of (5.20) for all $n \geq 0$. To justify this assumption, let us consider a DC function $f(\mathbf{x}) = f_1(\mathbf{x}) - f_2(\mathbf{x})$, then it is trivial to observe that $f(\mathbf{x}) = (f_1(\mathbf{x}) + \frac{\rho}{2}\|\mathbf{x}\|^2) - (f_2(\mathbf{x}) + \frac{\rho}{2}\|\mathbf{x}\|^2)$ for any given $\rho > 0$. Therefore, without loss of generality, we can always find a DC decomposition $f_1, f_2$ where both $f_1$ and $f_2$ are strongly convex.

The following theorem states the convergence properties of Algorithm 5.1:

**Theorem 5.1.** Let $\left\{\mathbf{x}^{(n)}\right\}$ be the solution sequence generated by Algorithm 1. Suppose (5.20) is feasible and Assumptions 5.1–5.3 hold for (5.20), then one of the following scenarios applies:

*1)* Algorithm 5.1 terminates after a finite number of $\breve{n}$ iterations and $\mathbf{x}^{(\breve{n})}$ is a stationary point of (5.20);

*2)* Algorithm 5.1 generates an infinite sequence every limit point of $\{\mathbf{x}^{(n)}\}$ is a stationary point of (5.20)[3].

*Proof:* Please see Appendix C.2. ■

Based on the above theorem, we can further state that the sequence of the objective function $\left\{\varphi(\mathbf{x}^{(n)})\right\}$ of (5.20) obtained by Algorithm 5.1 is also convergent.

[3]In order to perform asymptotic analysis on the algorithm, here we assume that the algorithm will only terminate if the criteria $\|\mathbf{x}^{(n+1)} - \mathbf{x}^{(n)}\| = 0$ and $\sum_{i=1}^{I} \mathrm{Tr}(\mathbf{S}_i) = 0$ are satisfied.

## 5.3 Application: Secure MIMO AF Relaying Optimization

In this section, we apply the proposed penalized DC algorithm in the previous section to the secure MIMO AF relaying optimization problem (5.11). We first show that the latter can be reformulated as a PSD DC program (5.20) by exploiting the so-called $\mathcal{S}$-procedure and by performing changes of variables. Subsequently, the penalized DC algorithm is adapted to solve the transformed optimization problem.

### 5.3.1 Transformation of the Secure Relaying Problem into a PSD DC Program

The robust secure relaying optimization (5.11) can be equivalently written as the following after substituting (5.4) and (5.7) into (5.11),

$$\max_{\sigma_{\mathtt{S}},\mathbf{W},\mathbf{\Psi}} \quad \frac{\sigma_{\mathtt{S}}^2|\mathbf{h}_2^H\mathbf{W}\mathbf{h}_1|^2}{\sigma_{\mathtt{R}}^2\|\mathbf{h}_2^H\mathbf{W}\|^2+\mathbf{h}_2^H\mathbf{\Psi}\mathbf{h}_2+\sigma_{\mathtt{D}}^2} \tag{5.38a}$$

$$\text{s.t.} \quad \max_{\Delta g_{1k}\in\mathcal{G}_{1k}} \frac{\sigma_{\mathtt{S}}^2|g_{1k}|^2}{\sigma_{\mathbb{E},k}^2} \le \gamma,\ k\in\mathcal{K} \tag{5.38b}$$

$$\max_{\Delta\mathbf{g}_{2k}\in\mathcal{G}_{2k}} \frac{\sigma_{\mathtt{S}}^2|\mathbf{g}_{2k}^H\mathbf{W}\mathbf{h}_1|^2}{\sigma_{\mathtt{R}}^2\|\mathbf{g}_{2k}^H\mathbf{W}\|^2+\mathbf{g}_{2k}^H\mathbf{\Psi}\mathbf{g}_{2k}+\sigma_{\mathbb{E},k}^2} \le \gamma,\ k\in\mathcal{K} \tag{5.38c}$$

$$\sigma_{\mathtt{S}}^2\|\mathbf{W}\mathbf{h}_1\|^2+\sigma_{\mathtt{R}}^2\|\mathbf{W}\|_F^2+\mathrm{Tr}(\mathbf{\Psi}) \le P_{\mathtt{R}} \tag{5.38d}$$

$$\sigma_{\mathtt{S}}^2 \le P_{\mathtt{S}},\ \mathbf{\Psi}\succeq\mathbf{0}, \tag{5.38e}$$

where $\gamma=2^{2\kappa R_d}-1$. Constraint (5.38b) can be equivalently rewritten as the following by exploiting the Cauchy-Schwarz inequality:

$$\sigma_{\mathtt{S}} \le \min_{k\in\mathcal{K}}\left\{\frac{\gamma\sigma_{\mathbb{E},k}^2}{\left||\hat{g}_{1k}|+\sqrt{\varepsilon_{1k}}\right|^2}\right\}. \tag{5.39}$$

Then to tackle the infiniteness associated with (5.38c), after some manipulations, we can rewrite (5.38c) as

$$\begin{aligned}&\Delta\mathbf{g}_{2k}^H\mathbf{\Theta}(\mathbf{W},\mathbf{\Psi})\Delta\mathbf{g}_{2k}+2\,\mathrm{Re}\left\{\hat{\mathbf{g}}_{2k}^H\mathbf{\Theta}(\mathbf{W},\mathbf{\Psi})\Delta\mathbf{g}_{2k}\right\}\\ &\quad+\hat{\mathbf{g}}_{2k}^H\mathbf{\Theta}(\mathbf{W},\mathbf{\Psi})\hat{\mathbf{g}}_{2k}-\gamma\sigma_{\mathbb{E},k}^2\le 0,\ \forall\Delta\mathbf{g}_{2k}\in\mathcal{G}_{2k}\end{aligned} \tag{5.40}$$

where we have defined $\mathbf{\Theta}(\mathbf{W},\mathbf{\Psi}) \triangleq \mathbf{W}\left(\sigma_{\mathtt{S}}^2\mathbf{h}_1\mathbf{h}_1^H - \gamma\sigma_{\mathtt{R}}^2\mathbf{I}_{N_{\mathtt{R}}}\right)\mathbf{W}^H - \gamma\mathbf{\Psi}$. As a popular technique of tackling the infiniteness in the robust optimization theory, we invoke the so-called $\mathcal{S}$-Procedure [93] for equivalently recasting (5.40) as

$$\mathbf{P}_k^H\mathbf{\Theta}(\mathbf{W},\mathbf{\Psi})\mathbf{P}_k - \mathbf{\Lambda}_k(\rho_k) \preceq \mathbf{0}, \tag{5.41}$$

where we have $\mathbf{\Lambda}_k(\rho_k) = \mathtt{blkdiag}(\rho_k\mathbf{I}_{N_{\mathtt{R}}}, \gamma\sigma_{\mathbb{E},k}^2 - \varepsilon_{2k}\rho_k)$ and $\mathbf{P}_k = [\mathbf{I}_{N_{\mathtt{R}}}, \hat{\mathbf{g}}_{2k}]$ with $\mathtt{blkdiag}(\cdot,\cdot)$ denoting the construction of a block diagonal matrix from the input arguments.

To further transform (5.38) into a PSD DC program in the form of (5.20), let us introduce an auxiliary variable $t$. Plugging (5.40) and (5.41) back into (5.38), we obtain

$$\begin{aligned}
\max_{\sigma_{\mathtt{S}},\mathbf{W},\mathbf{\Psi}} \quad & \frac{\sigma_{\mathtt{S}}^2|\mathbf{h}_2^H\mathbf{W}\mathbf{h}_1|^2}{t} && \text{(5.42a)}\\
\text{s.t.} \quad & \sigma_{\mathtt{R}}^2\|\mathbf{h}_2^H\mathbf{W}\|^2 + \mathbf{h}_2^H\mathbf{\Psi}\mathbf{h}_2 + \sigma_{\mathtt{D}}^2 \le t && \text{(5.42b)}\\
& \sigma_{\mathtt{S}} \le \min_{k\in\mathcal{K}}\left\{\frac{\gamma\sigma_{\mathbb{E},k}^2}{\left||\hat{g}_{1k}| + \sqrt{\varepsilon_{1k}}\right|^2}\right\} && \text{(5.42c)}\\
& \mathbf{P}_k^H\mathbf{\Theta}(\mathbf{W},\mathbf{\Psi})\mathbf{P}_k - \mathbf{\Lambda}_k(\rho_k) \preceq \mathbf{0},\ k\in\mathcal{K} && \text{(5.42d)}\\
& \sigma_{\mathtt{S}}^2\|\mathbf{W}\mathbf{h}_1\|^2 + \sigma_{\mathtt{R}}^2\|\mathbf{W}\|_F^2 + \mathrm{Tr}(\mathbf{\Psi}) \le P_{\mathtt{R}} && \text{(5.42e)}\\
& \sigma_{\mathtt{S}} \le \sqrt{P_{\mathtt{S}}},\ \mathbf{\Psi}\succeq\mathbf{0}. && \text{(5.42f)}
\end{aligned}$$

Observe that in the above formulation, the source power $\sigma_{\mathtt{S}}$ and relay AF matrix $\mathbf{W}$ are nonlinearly coupled in the objective (5.42a) and constraints (5.42d) and (5.42e). To transform (5.42) into a more convenient form, we introduce a new optimization variable $\mathbf{U}$, which is related to $\sigma_{\mathtt{S}}$ and $\mathbf{W}$ via the following bilinear matrix equality:

$$\mathbf{U} = \sigma_{\mathtt{S}}\mathbf{W}. \tag{5.43}$$

With the aid of (5.43), (5.42) can then be expressed as a PSD DC program with bilinear

matrix equality constraint as defined in (5.16), i.e.,

$$\max_{\mathbf{x}} \quad \frac{|\mathbf{h}_2^H \mathbf{U} \mathbf{h}_1|^2}{t} \tag{5.44a}$$

$$\text{s.t.} \quad \mathbf{P}_k^H \boldsymbol{\Theta}(\mathbf{W}, \boldsymbol{\Psi}) \mathbf{P}_k - \boldsymbol{\Lambda}_k(\rho_k) \preceq \mathbf{0}, \ k \in \mathcal{K} \tag{5.44b}$$

$$\mathbf{U} = \sigma_{\mathrm{S}} \mathbf{W} \tag{5.44c}$$

$$\mathbf{x} \in \Omega, \tag{5.44d}$$

where $\mathbf{x}$ collectively denotes all the optimization variables (including both the original and auxiliary variables), i.e.,

$$\mathbf{x} \triangleq [\sigma_{\mathrm{S}}, \mathtt{vec}(\mathbf{U})^T, \mathtt{vec}(\mathbf{W})^T, \mathtt{vec}(\boldsymbol{\Psi})^T, \boldsymbol{\rho}^T, t]^T, \tag{5.45}$$

and $\Omega$ is a compact convex subset defined as

$$\Omega \triangleq \{\mathbf{x} : (5.42\text{b}), (5.42\text{c}), (5.42\text{e}), (5.42\text{f})\}, \tag{5.46}$$

which can easily be represented as a finite number of LMIs by exploiting the techniques introduced in [94].

To tackle the bilinear matrix equality constraint (5.44c), we follow the procedure proposed in the previous section, i.e., exploit the results in Lemma 5.1, and conveniently convert (5.44c) into

$$\begin{bmatrix} \mathbf{Y}_1 & \mathbf{U} & \sigma_{\mathrm{S}} \mathbf{I}_{N_{\mathrm{R}}} \\ \mathbf{U}^H & \mathbf{Y}_2 & \mathbf{W}^H \\ \sigma_{\mathrm{S}} \mathbf{I}_{N_{\mathrm{R}}} & \mathbf{W} & \mathbf{I}_{N_{\mathrm{R}}} \end{bmatrix} \succeq \mathbf{0} \tag{5.47}$$

$$\mathrm{Tr}(\mathbf{Y}_1) - \mathrm{Tr}(\sigma_{\mathrm{S}}^2 \mathbf{I}_{N_{\mathrm{R}}}) \leq 0, \tag{5.48}$$

where $\mathbf{Y}_1$ and $\mathbf{Y}_2$ are auxiliary matrix variables with appropriate dimensions, (5.47) is an LMI, and (5.48) is a DC constraint (special case of PSD DC constraint with dimension one). Therefore, we can now embed the LMI (5.47) into $\Omega$ while preserving its convex structure. Additionally, the collection of optimization variables represented by $\mathbf{x}$ is augmented with the new auxiliary variables $\mathbf{Y}_1$ and $\mathbf{Y}_2$.

Finally, note that the matrix inequality constraint (5.44b) can be expressed as a PSD

DC constraint as follows:

$$\underbrace{\mathbf{P}_k^H\mathbf{U}\mathbf{h}_1\mathbf{h}_1^H\mathbf{U}^H\mathbf{P}_k - \mathbf{\Lambda}_k(\rho_k) - \gamma\mathbf{P}_k^H\mathbf{\Psi}\mathbf{P}_k}_{\boldsymbol{\mathcal{F}}_k(\cdot)} - \underbrace{\gamma\sigma_{\mathtt{R}}^2\mathbf{P}_k^H\mathbf{W}\mathbf{W}^H\mathbf{P}_k}_{\boldsymbol{\mathcal{G}}_k(\cdot)} \preceq \mathbf{0},\ k\in\mathcal{K}, \tag{5.49}$$

where both $\boldsymbol{\mathcal{F}}_k(\cdot)$ and $\boldsymbol{\mathcal{G}}_k(\cdot)$ are PSD-convex mappings.

Based on the above derivations, we arrive at the following PSD DC program:

$$\begin{aligned}
\min_{\mathbf{x}} \quad & -\frac{|\mathbf{h}_2^H\mathbf{U}\mathbf{h}_1|^2}{t} && \text{(5.50a)}\\
\text{s.t.} \quad & \mathbf{P}_k^H\mathbf{U}\mathbf{h}_1\mathbf{h}_1^H\mathbf{U}^H\mathbf{P}_k - \mathbf{\Lambda}_k(\rho_k) - \gamma\mathbf{P}_k^H\mathbf{\Psi}\mathbf{P}_k \\
& \qquad - \gamma\sigma_{\mathtt{R}}^2\mathbf{P}_k^H\mathbf{W}\mathbf{W}^H\mathbf{P}_k \preceq \mathbf{0},\ k\in\mathcal{K} && \text{(5.50b)}\\
& \mathrm{Tr}(\mathbf{Y}_1) - \mathrm{Tr}(\sigma_{\mathtt{S}}^2\mathbf{I}_{N_{\mathtt{R}}}) \le 0 && \text{(5.50c)}\\
& \mathbf{x}\in\Omega. && \text{(5.50d)}
\end{aligned}$$

### 5.3.2 Penalized PSD DC Algorithm for Secure Relaying Design

For simplicity, let us denote

$$g_0(\mathbf{U},t) = -\frac{|\mathbf{h}_2^H\mathbf{U}\mathbf{h}_1|^2}{t} \tag{5.51}$$

$$g_1(\sigma_{\mathtt{S}}) = \mathrm{Tr}(\sigma_{\mathtt{S}}^2\mathbf{I}_{N_{\mathtt{R}}}) = N_{\mathtt{R}}\sigma_{\mathtt{S}}^2. \tag{5.52}$$

The algorithm designed for (5.50) is described as Algorithm 5.2, where $\hat{g}_0(\cdot;\cdot)$, $\hat{g}_1(\cdot;\cdot)$ and $\hat{\boldsymbol{\mathcal{G}}}_k(\cdot;\cdot)$ denote the first-order approximations of their corresponding functions/mapping around a solution from the previous iteration:

$$\begin{aligned}
\hat{g}_0(\mathbf{U},t;\mathbf{U}^{(n)},t^{(n)}) = {} & \frac{\left|\mathbf{h}_2^H\mathbf{U}^{(n)}\mathbf{h}_1\right|^2}{t^{(n)}} - \frac{\left|\mathbf{h}_2^H\mathbf{U}^{(n)}\mathbf{h}_1\right|^2}{(t^{(n)})^2}\left(t-t^{(n)}\right) \\
& + \frac{1}{t^{(n)}}2\,\mathrm{Re}\left\{\mathbf{h}_2^H\mathbf{U}^{(n)}\mathbf{h}_1\mathbf{h}_1^H\left(\mathbf{U}-\mathbf{U}^{(n)}\right)^H\mathbf{h}_2\right\} && \text{(5.53)}\\
\hat{g}_1(\sigma_{\mathtt{S}};\sigma_{\mathtt{S}}^{(n)}) = {} & N_{\mathtt{R}}(\sigma_{\mathtt{S}}^{(n)})^2 - 2N_{\mathtt{R}}\sigma_{\mathtt{S}}^{(n)}\left(\sigma_{\mathtt{S}}+\sigma_{\mathtt{S}}^{(n)}\right) && \text{(5.54)}\\
\hat{\boldsymbol{\mathcal{G}}}_k(\mathbf{W};\mathbf{W}^{(n)}) = {} & \gamma\sigma_{\mathtt{R}}^2\mathbf{P}_k^H\mathbf{W}^{(n)}\left(\mathbf{W}^{(n)}\right)^H\mathbf{P}_k + 2\gamma N_{\mathtt{R}}\,\mathrm{Re}\left(\mathbf{P}_k^H\mathbf{W}^{(n)}\left(\mathbf{W}-\mathbf{W}^{(n)}\right)^H\mathbf{P}_k\right).
\end{aligned}$$

We now briefly analyze the theoretical complexity of solving each sub-problem in (5.55).

**Algorithm 5.2** Penalized PSD DC Algorithm for Secure Relaying Design

Initialization: An initial point $\mathbf{x}^{(0)} \in \Omega$, $\tau^{(0)} > 0$, $\delta_1 > 0$ and $\delta_2 > 0$. Set $n = 0$.
**repeat**
Compute $\mathbf{x}^{(n+1)}$ by solving the convex sub-problem:

$$\min_{\mathbf{x}} \quad -\hat{g}_0(\mathbf{U}, t; \mathbf{U}^{(n)}, t^{(n)}) + \tau^{(n)}(s + \sum_{k=1}^{K} \mathrm{Tr}(\mathbf{S}_k)) \tag{5.55a}$$

$$\text{s.t.} \quad \mathrm{F}_k - \hat{\boldsymbol{\mathcal{G}}}_k(\mathbf{W}, \mathbf{W}^{(n)}) \preceq \mathbf{S}_k, \ k \in \mathcal{K} \tag{5.55b}$$

$$\mathrm{Tr}(\mathbf{Y}_1) - \hat{g}_1(\sigma_{\mathtt{S}}, \sigma_{\mathtt{S}}^{(n)}) \leq s \tag{5.55c}$$

$$\mathbf{x} \in \Omega. \tag{5.55d}$$

Update $\tau$ via (5.30);
Update iteration: $n \leftarrow n + 1$
**until** Termination criterion is satisfied or a maximum number of iterations are reached

Since (5.55) is a standard SDP, its complexity mainly depends on the number of optimization variables and the number of semidefinite cone constraints. It is not difficult to verify (5.55) involves on the order of $\mathcal{O}(N_{\mathtt{R}}^2 + N_{\mathtt{R}} + K + 1)$ optimization variables and $K$ semidefinite cone constraints of dimension $(N_{\mathtt{R}} + 1)^2$ Therefore, as analyzed in [94], (5.55) can be solved at a worst case complexity, which is on the order of $\mathcal{O}((N_{\mathtt{R}}^2 + N_{\mathtt{R}} + K + 1)^2 (N_{\mathtt{R}} + 1)^2$.

Before leaving this section, the following remarks are of interests:

**Remark 5.3** (On the initialization of Algorithm 5.2)**.** Since $\Omega$ defined in (5.46) is a compact convex subset, we are able to efficiently exploit its bounded structure, and conveniently select a feasible initialization, e.g.,

$$\sigma_{\mathtt{S}}^{(0)} = \min\left\{\sqrt{P_{\mathtt{S}}}, \min_{k\in\mathcal{K}}\left\{\frac{\gamma\sigma_{\mathbb{E},k}^2}{\left||\hat{g}_{1k}| + \sqrt{\varepsilon_{1k}}\right|^2}\right\}\right\} - \epsilon$$

$$\mathbf{W}^{(0)} = \left(\frac{P_{\mathtt{R}}}{(\sigma_{\mathtt{S}}^{(0)})^2\|\mathbf{h}_1\|^2 + \sigma_{\mathtt{R}}^2 N_{\mathtt{R}}}\right)^{\frac{1}{2}} \mathbf{I}_{N_{\mathtt{R}}}$$

$$\mathbf{U}^{(0)} = \sigma_{\mathtt{S}}^{(0)}\mathbf{W}^{(0)}.$$

where $\epsilon$ is a small positive number. □

**Remark 5.4** (Extension to the case of MRC)**.** We show that the proposed penalized DC algorithm is also applicable to the case where `eve`s adopt a more complicated receive

MRC scheme for decoding the messages from the legitimate UEs. With MRC, the mutual information leakage to the eves can now be given by

$$\mathtt{C}_{\mathbb{E},k}(\sigma_{\mathtt{S}}, \mathbf{W}, \mathbf{\Psi}) = \frac{1}{2}\log_2\left(1 + \frac{\sigma_{\mathtt{S}}^2|\mathbf{g}_{1k}|^2}{\sigma_{\mathbb{E},k}^2} + \frac{\sigma_{\mathtt{S}}^2|\mathbf{g}_{2k}^H\mathbf{W}\mathbf{h}_1|^2}{\sigma_{\mathtt{R}}^2\|\mathbf{g}_{2k}^H\mathbf{W}\|^2 + \mathbf{g}_{2k}^H\mathbf{\Psi}\mathbf{g}_{2k} + \sigma_{\mathbb{E},k}^2}\right). \tag{5.56}$$

We adopt a rate-splitting approach, i.e., we introduce a pair of weights $(\gamma_1, \gamma_2)$ with $\gamma_1+\gamma_2 = \gamma$ as defined below (5.38), and the robust secrecy constraint can be subsequently formulated as

$$\begin{cases} \max\limits_{\Delta g_{1k}\in\mathcal{G}_{1k}} \frac{\sigma_{\mathtt{S}}^2|g_{1k}|^2}{\sigma_{\mathbb{E},k}^2} \le \gamma_1, \ k \in \mathcal{K} \\ \max\limits_{\Delta \mathbf{g}_{2k}\in\mathcal{G}_{2k}} \frac{\sigma_{\mathtt{S}}^2|\mathbf{g}_{2k}^H\mathbf{W}\mathbf{h}_1|^2}{\sigma_{\mathtt{R}}^2\|\mathbf{g}_{2k}^H\mathbf{W}\|^2+\mathbf{g}_{2k}^H\mathbf{\Psi}\mathbf{g}_{2k}+\sigma_{\mathbb{E},k}^2} \le \gamma_2, \ k \in \mathcal{K} \end{cases}. \tag{5.57}$$

It can be observed that in the case of MRC, the robust secrecy constraints have a similar form to that of the SC by prefixing a pair of weights $(\beta, 2R_d-\beta)$. Therefore, given different values of $\beta$, we can obtain a set of solutions $(\sigma_{\mathtt{S}}, \mathbf{W}, \mathbf{\Psi})$ using Algorithm 5.2. Within such set of solution, the best solution can be achieved by the one that attains the maximum $\mathtt{SINR}_{\mathtt{D}}$ in the objective function. □

## 5.4 Benchmarker: SDR-Based Exhaustive Search Method

To benchmark the proposed penalized DC algorithm, in this section we derive an SDR-based approach that yields an upper-bound for the robust secrecy problem (5.11), however, at the expense of higher computational complexity.

It is in general challenging to jointly optimize the tuple of $(\sigma_{\mathtt{S}}, \mathbf{W}, \mathbf{\Psi})$ due to its non-convex nature and therefore, we can consider a sub-problem of (5.11) solving for the optimal pair $(\mathbf{W}, \mathbf{\Psi})$, while temporarily fixing the value of $\sigma_{\mathtt{S}}$. Substituting the expression of $\mathtt{C}_{\mathbb{E},k}$ in (5.7) into (5.11b) and neglecting the terms independent of $(\mathbf{W}, \mathbf{\Psi})$, we arrive at the

following sub-problem:

$$\tau(\sigma_{\mathtt{S}}) \triangleq \max_{\mathbf{W},\mathbf{\Psi}\succeq\mathbf{0}} \quad \frac{\sigma_{\mathtt{S}}^2|\mathbf{h}_2^H\mathbf{W}\mathbf{h}_1|^2}{\sigma_{\mathtt{R}}^2\|\mathbf{h}_2^H\mathbf{W}\|^2+\mathbf{h}_2^H\mathbf{\Psi}\mathbf{h}_2+\sigma_{\mathtt{D}}^2} \tag{5.58a}$$

$$\text{s.t.} \quad \log_2\left(1+\frac{\sigma_{\mathtt{S}}^2|\mathbf{g}_{2k}^H\mathbf{W}\mathbf{h}_1|^2}{\sigma_{\mathtt{R}}^2\|\mathbf{g}_{2k}^H\mathbf{W}\|^2+\mathbf{g}_{2k}^H\mathbf{\Psi}\mathbf{g}_{2k}+\sigma_{\mathbb{E},k}^2}\right) \leq \kappa R_d,\ \forall \Delta\mathbf{g}_{2k}\in\mathcal{G}_{2k}, k\in\mathcal{K} \tag{5.58b}$$

$$\sigma_{\mathtt{S}}^2\|\mathbf{W}\mathbf{h}_1\|^2+\sigma_{\mathtt{R}}^2\|\mathbf{W}\|_F^2+\mathrm{Tr}(\mathbf{\Psi})\leq P_{\mathtt{R}}, \tag{5.58c}$$

where $\tau(\sigma_{\mathtt{S}})$ denotes its objective value, which depends on the value of $\sigma_{\mathtt{S}}$. With the aid of (5.58), the original problem (5.11) can equivalently be expressed as

$$\max_{\sigma_{\mathtt{S}}} \quad \tau(\sigma_{\mathtt{S}}) \quad \text{s.t.} \quad 0\leq\sigma_{\mathtt{S}}\leq\bar{\sigma}_{\mathtt{S}}, \tag{5.59}$$

where $\sigma_{\mathtt{S}}$ is lower bounded by zero, while its upper bound $\bar{\sigma}_{\mathtt{S}}$ is given by [c.f. (5.39)]

$$\bar{\sigma}_{\mathtt{S}}=\min\left\{\sqrt{P_{\mathtt{S}}},\min_{k\in\mathcal{K}}\left\{\frac{\gamma\sigma_{\mathbb{E},k}^2}{\left||\hat{g}_{1k}|+\sqrt{\varepsilon_{1k}}\right|^2}\right\}\right\}. \tag{5.60}$$

The reformulated problem in (5.59) leads to a simpler single-variable optimization problem defined over the interval $[0,\ \bar{\sigma}_{\mathtt{S}}]$. Assuming that $\tau(\sigma_{\mathtt{S}})$ can be evaluated at any feasible $\sigma_{\mathtt{S}}$, a one-dimensional exhaustive search procedure can be invoked for finding the global optimum of (5.59)[4]. Let us now focus our attention on computing $\tau(\sigma_{\mathtt{S}})$ for a given feasible $\sigma_{\mathtt{S}}$, which however requires solving the non-convex sub-problem (5.58). The solution to (5.58) will be addressed in the following.

Recall that the infiniteness of the constraint in (5.58b) can be tackled by the $\mathcal{S}$-procedure [c.f. (5.40), (5.41)], which leads to the following equivalent reformulation:

$$\mathbf{\Lambda}_k(\rho_k)-\mathbf{P}_k^H\mathbf{\Theta}(\mathbf{W},\mathbf{\Psi})\mathbf{P}_k\succeq\mathbf{0}. \tag{5.61}$$

Replacing (5.58b) by (5.61), the sub-problem in $(\mathbf{W},\mathbf{\Psi})$ of (5.58) can now be rewritten in

[4]Another widely-used search approach in optimization is the line search, which leads to significantly lower complexity than the exhaustive search. Note that the line search strategy requires the computation of a descent direction and a step size. In the case of (5.58), the optimized $\mathbf{W}$ and $\mathbf{\Psi}$ will implicitly depend on $\sigma_{\mathtt{S}}$ such that a descent direction for the objective function $\tau(\sigma_{\mathtt{S}})$ (e.g., the gradient) will be difficult to find. Therefore, we focus on the exhaustive search in this section.

a finite form:

$$\max_{\mathbf{W},\boldsymbol{\Psi}} \quad \frac{\sigma_{\mathtt{S}}^2|\mathbf{h}_2^H\mathbf{W}\mathbf{h}_1|^2}{\sigma_{\mathtt{R}}^2\|\mathbf{h}_2^H\mathbf{W}\|^2+\mathbf{h}_2^H\boldsymbol{\Psi}\mathbf{h}_2+\sigma_{\mathtt{D}}^2} \tag{5.62a}$$

$$\text{s.t.} \quad \sigma_{\mathtt{S}}^2\|\mathbf{W}\mathbf{h}_1\|^2+\sigma_{\mathtt{R}}^2\|\mathbf{W}\|_F^2+\mathrm{Tr}(\boldsymbol{\Psi})\leq P_{\mathtt{R}} \tag{5.62b}$$

$$\boldsymbol{\Lambda}_k(\rho_k)-\mathbf{P}_k^H\boldsymbol{\Theta}(\mathbf{W},\boldsymbol{\Psi})\mathbf{P}_k\succeq\mathbf{0},\ k\in\mathcal{K} \tag{5.62c}$$

$$\boldsymbol{\Psi}\succeq\mathbf{0}. \tag{5.62d}$$

The above transformed formulation is still non-convex and to proceed, we have to transform it into an appropriate formulation, where the SDR is applicable. Let us define $\mathbf{w}=\mathtt{vec}(\mathbf{W})$ and $\mathbf{X}=\mathbf{w}\mathbf{w}^H$. Interestingly, after some tedious matrix manipulations, which have been relegated to Appendix C.5, we are now able to rewrite (5.62) in a form, which only involves the linear terms of $\mathbf{X}$ and $\boldsymbol{\Psi}$. The results are summarized in the following proposition:

**Proposition 5.2.** Define

$$\mathbf{Q}_0=\sigma_{\mathtt{S}}^2(\mathbf{h}_1^*\mathbf{h}_1^T)\otimes(\mathbf{h}_2\mathbf{h}_2^H) \tag{5.63}$$

$$\mathbf{Q}_1=\sigma_{\mathtt{R}}^2\mathbf{I}_{N_{\mathtt{R}}}\otimes(\mathbf{h}_2\mathbf{h}_2^H) \tag{5.64}$$

$$\mathbf{Q}_2=\sigma_{\mathtt{S}}^2(\mathbf{h}_1^*\mathbf{h}_1^T)\otimes\mathbf{I}_{N_{\mathtt{R}}}+\sigma_{\mathtt{R}}^2\mathbf{I}_{N_{\mathtt{R}}^2} \tag{5.65}$$

$$\mathbf{Q}_3(\mathbf{X},\boldsymbol{\Psi})=\sigma_{\mathtt{S}}^2\mathbf{H}_1\mathbf{X}\mathbf{H}_1^H-\gamma\sigma_{\mathtt{R}}^2\sum_{l=1}^{N_{\mathtt{R}}}\mathbf{E}_l\mathbf{X}\mathbf{E}_l^H-\gamma\boldsymbol{\Psi}, \tag{5.66}$$

where $\mathbf{Q}_3(\cdot)$ is a linear mapping of $\mathbf{X}$ and $\boldsymbol{\Psi}$ with $\mathbf{H}_1=\mathbf{h}_1^T\otimes\mathbf{I}_{N_{\mathtt{R}}}$ and $\mathbf{E}_l$ is defined as $\mathbf{E}_l=\left[\mathbf{0}_{N_{\mathtt{R}}\times(l-1)N_{\mathtt{R}}},\mathbf{I}_{N_{\mathtt{R}}},\mathbf{0}_{N_{\mathtt{R}}\times(N_{\mathtt{R}}-l)N_{\mathtt{R}}}\right]$. Then problem (5.62) can equivalently be rewritten in the following form:

$$\max_{\mathbf{X},\boldsymbol{\Psi},\boldsymbol{\rho}} \quad \frac{\mathrm{Tr}(\mathbf{Q}_0\mathbf{X})}{\mathrm{Tr}(\mathbf{Q}_1\mathbf{X})+\mathrm{Tr}(\mathbf{h}_2\mathbf{h}_2^H\boldsymbol{\Psi})+\sigma_{\mathtt{D}}^2} \tag{5.67a}$$

$$\text{s.t.} \quad \mathrm{Tr}(\mathbf{Q}_2\mathbf{X})+\mathrm{Tr}(\boldsymbol{\Psi})\leq P_{\mathtt{R}} \tag{5.67b}$$

$$\boldsymbol{\Lambda}_k(\rho_k)-\mathbf{P}_k^H\mathbf{Q}_3(\mathbf{X},\boldsymbol{\Psi})\mathbf{P}_k\succeq\mathbf{0},\ k\in\mathcal{K} \tag{5.67c}$$

$$\mathbf{X}\succeq\mathbf{0},\ \boldsymbol{\Psi}\succeq\mathbf{0},\ \ \mathrm{Rank}(\mathbf{X})=1. \tag{5.67d}$$

Upon neglecting the non-convex rank-one constraint in (5.67d), (5.67) is relaxed to a so-called fractional SDP, which can further be transformed into a standard SDP via the

Charnes-Cooper transformation [125]. Specifically, by introducing an auxiliary variable $s > 0$, and defining $\overline{\mathbf{X}} = s\mathbf{X}$, $\overline{\mathbf{\Psi}} = s\mathbf{\Psi}$ and $\overline{\boldsymbol{\rho}} = s\boldsymbol{\rho}$, (5.67) is conveniently recast as

$$\max_{\overline{\mathbf{X}},\overline{\mathbf{\Psi}},\overline{\boldsymbol{\rho}},s>0} \quad \mathrm{Tr}(\mathbf{Q}_0\overline{\mathbf{X}}) \tag{5.68a}$$

$$\text{s.t.} \quad \mathrm{Tr}(\mathbf{Q}_1\overline{\mathbf{X}}) + \mathrm{Tr}(\mathbf{h}_2\mathbf{h}_2^H\overline{\mathbf{\Psi}}) + s\sigma_{\mathtt{D}}^2 \leq 1 \tag{5.68b}$$

$$\mathrm{Tr}(\mathbf{Q}_2\overline{\mathbf{X}}) + \mathrm{Tr}(\overline{\mathbf{\Psi}}) \leq sP_{\mathtt{R}} \tag{5.68c}$$

$$\mathbf{\Lambda}_k(\overline{\rho}_k) - \mathbf{P}_k^H\mathbf{Q}_3(\overline{\mathbf{X}},\overline{\mathbf{\Psi}})\mathbf{P}_k \succeq \mathbf{0},\ k \in \mathcal{K} \tag{5.68d}$$

$$\overline{\mathbf{X}} \succeq \mathbf{0},\ \overline{\mathbf{\Psi}} \succeq \mathbf{0}. \tag{5.68e}$$

Interestingly, (5.68) now becomes a convex SDP, which is efficiently solvable by generic optimization tools such as `SeDuMi` [101] and `MOSEK` [102] relying on interior-point methods [103]. We remark that (5.68) and the rank-relaxed version of (5.67) are equivalent in the sense that the optimal solution $\mathbf{X}^*$ to (5.67) after rank-one relaxation can be retrieved by the optimal solution $(\overline{\mathbf{X}}^*, s^*)$ to (5.68), i.e., $\mathbf{X}^* = \frac{\overline{\mathbf{X}}^*}{s^*}$, and the resultant objective values of the two problems are equivalent.

After obtaining the rank-relaxed solution $\mathbf{X}^*$, a natural question arises as to how good a solution is $\mathbf{X}^*$, i.e., does it satisfy the rank-one optimality condition of (5.67)? Answering these questions directly from the formulation of (5.68) is still an open problem in the literature[5]. To overcome this difficulty, we follow an approach similar to [87]. Specifically, denoting the objective value of (5.68) by $\tau_{\text{relax}}^*(\sigma_{\mathtt{S}})$, we consider the following power minimization problem:

$$\min_{\mathbf{X},\mathbf{\Psi},\boldsymbol{\rho}} \quad \mathrm{Tr}(\mathbf{Q}_2\mathbf{X}) \tag{5.69a}$$

$$\text{s.t.} \quad \frac{\mathrm{Tr}(\mathbf{Q}_0\mathbf{X})}{\mathrm{Tr}(\mathbf{Q}_1\mathbf{X}) + \mathrm{Tr}(\mathbf{h}_2\mathbf{h}_2^H\mathbf{\Psi}) + \sigma_{\mathtt{D}}^2} \geq \tau_{\text{relax}}^*(\sigma_{\mathtt{S}}) \tag{5.69b}$$

$$\mathrm{Tr}(\mathbf{Q}_2\mathbf{X}) + \mathrm{Tr}(\mathbf{\Psi}) \leq P_{\mathtt{R}} \tag{5.69c}$$

$$\mathbf{\Lambda}_k(\rho_k) - \mathbf{P}_k^H\mathbf{Q}_3(\mathbf{X},\mathbf{\Psi})\mathbf{P}_k \succeq \mathbf{0},\ k \in \mathcal{K} \tag{5.69d}$$

$$\mathbf{X} \succeq \mathbf{0},\ \mathbf{\Psi} \succeq \mathbf{0}. \tag{5.69e}$$

[5]Note the methods developed in [126, 127] can be used to obtain an optimal rank-one solution to an SDP similar to (5.68), however, with only a limited number of constraints. In our case, since the number of `eves` can be arbitrarily large, we proceed with an alternative method by solving an inverse problem to (5.68).

Observe that (5.69) is also a standard SDP and therefore it is readily solvable by existing optimization tools. Furthermore, its specific structure allows us to obtain the following useful results, based on which we are able to retrieve an optimal rank-one solution of (5.67).

**Proposition 5.3.** Let us denote the optimal solution of (5.69) by $(\mathbf{X}^o, \mathbf{\Psi}^o, \boldsymbol{\rho}^o)$. Assuming suitable constraint qualification of (5.69), $(\mathbf{X}^o, \mathbf{\Psi}^o, \boldsymbol{\rho}^o)$ is also an optimal solution of (5.67), i.e., $\mathbf{X}^o$ must be of rank one.

*Proof:* Please see appendix C.6. ■

In summary, obtaining an optimal solution of (5.59) now consists of two steps: *1)* solve the rank-relaxed SDP (5.68) and obtain the largest $\tau_{\text{relax}}(\sigma_{\mathtt{S}})$ by exhaustive search over $\sigma_{\mathtt{S}}$; *2)* solve the power minimization problem (5.69) based on $\tau_{\text{relax}}(\sigma_{\mathtt{S}})$. Since the rank-one optimality condition of $\mathbf{X}^o$ is guaranteed, the optimal AF matrix $\mathbf{W}^o$ can be retrieved by the rank-one decomposition of $\mathbf{X}^o$, i.e., $\mathbf{X}^o = \mathbf{x}^o(\mathbf{x}^o)^H$ and subsequently converting $\mathbf{x}^o$ to $\mathbf{W}^o$ via the vector-matrix reshaping.

We should point out that solving (5.59) requires performing an exhaustive search for $\sigma_{\mathtt{S}}$ over $[0,\ \bar{\sigma}_{\mathtt{S}}]$. In each step, we have to solve the SDP (5.68), which involves on the order of $\mathcal{O}(N_{\mathtt{R}}^4 + N_{\mathtt{R}}^2 + 1)$ optimization variables and $K$ semidefinite cone constraints of dimension $(N_{\mathtt{R}} + 1)^2$. Therefore, it can be solved at a *worst-case* complexity, which is on the order of $\mathcal{O}\left(K(N_{\mathtt{R}}^4 + N_{\mathtt{R}}^2 + 1)^2(N_{\mathtt{R}} + 1)^2\right)$ [94]. As compared to the complexity of the proposed penalized DC algorithm (see, e.g., analysis below Algorithm 5.2), The associated computational cost escalates significantly faster as the size of the relay antenna array and the number of eves increase, which may become computationally prohibitive in practical problems.

## 5.5 Numerical Examples

The efficacy of the proposed solutions to the robust secure relaying problem is verified by a few numerical examples. In all simulations, all the coefficients of the legitimate channels $\mathbf{h}_1$ and $\mathbf{h}_2$, and the estimated eves' channels $\{\hat{g}_{1,k}\}$ and $\{\hat{\mathbf{g}}_{2,k}\}$ are generated following i.i.d. complex circular Gaussian distribution with zero-mean and unit-variance. Equal radii are assumed for all $\Delta g_{1,k}$ and for all $\Delta\mathbf{g}_{2,k}$, i.e., $\varepsilon_{1,k} = \varepsilon_1$ and $\varepsilon_{2,k} = \varepsilon_2$ for all $k$. The power budget of S is normalized to one and we set higher power budget for R with

$P_{\mathtt{R}} = 2$. It is also assumed that an antenna array of size $N_{\mathtt{R}} = 3$ is employed by $\mathtt{R}$. The noise variances are $\sigma_{\mathtt{R}}^2 = 0.05$, $\sigma_{\mathtt{D}}^2 = 0.05$ and $\sigma_{\mathbb{E},k}^2 = 0.01\ \forall k$. The above parameters are fixed unless otherwise explicitly stated. In all figures, we denote the proposed penalized DC algorithm in Section 5.3 by "Proposed P-DCA" and the derived benchmarker in Section 5.4 by "SDR+Search".

### Convergence

We first study the convergence behavior of Algorithm 1. We simulate 200 channel realizations and among which, two classes of behaviors are observed. A representative case for each class is then plotted in the left and right parts of Fig. 5.2. In each case, the top sub-figure shows the convergence of the achieved SINR at $\mathtt{D}$ while the bottom sub-figure plots the evolution of the FI. The first case shows a behavior similar to conventional DC algorithm. The second example shows a more interesting behavior where the algorithm begins with an infeasible point and in the first few iterations, the algorithm targets finding a region (still infeasible) with larger objective function. As the penalty terms gradually play more important roles, more emphasis will be on finding a feasible point near the above located region. Therefore, the value of objective function drops since the feasibility has to be enforced now. Finally, the SINR remains approximately the same because a stationary point is achieved. The convergence behavior is consistent with the discussions and proof in Section 5.3.

### Secrecy

To evaluate the secrecy of relaying transmission achieved by the proposed solutions, i.e., how consistently the robust secrecy constraints (5.11b) can be satisfied, we follow a probabilistic approach similar to [30, Section VI–B]. In this example, the coefficients of $\Delta\hat{g}_{1k}$ and $\Delta\hat{\mathbf{g}}_{2k}$ are generated by i.i.d. zero-mean complex circular Gaussian distribution with variance $\sigma_h^2 = 0.05$. The radii of uncertainty regions in (5.9) and (5.10) are then determined by $\varepsilon_1 = \sigma_h^2 \times \mathtt{gammaincinv}(\mathrm{Pr}, 0.5)$ and $\varepsilon_2 = \sigma_h^2 \times \mathtt{gammaincinv}(\mathrm{Pr}, 0.5N_{\mathtt{R}}^2)$ where $\mathtt{gammaincinv}(\cdot)$ is the inverse of incomplete gamma function defined in `MATLAB` and Pr is a predefined bounding probability, say, Pr = 95%, c.f. [30, (61)]. The empirical CDFs of mutual information leakage at both $\mathtt{eve}$s are shown in Fig 5.3. Both the proposed solutions ensures that the mutual information leakage never exceeds the data rate of legitimate UEs while the non-

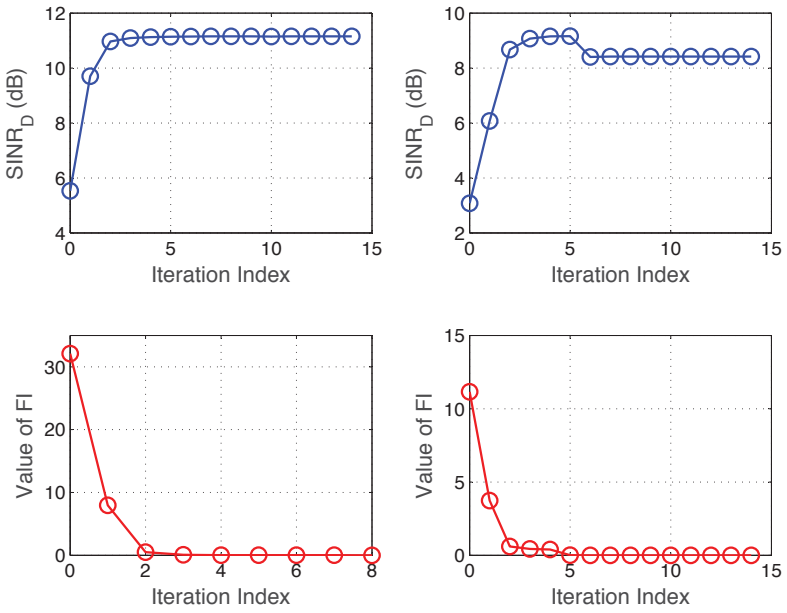



**Fig. 5.2** Convergence behavior of Algorithm 1. Left set of sub-figures: The first case. Right set of sub-figures: The second case.

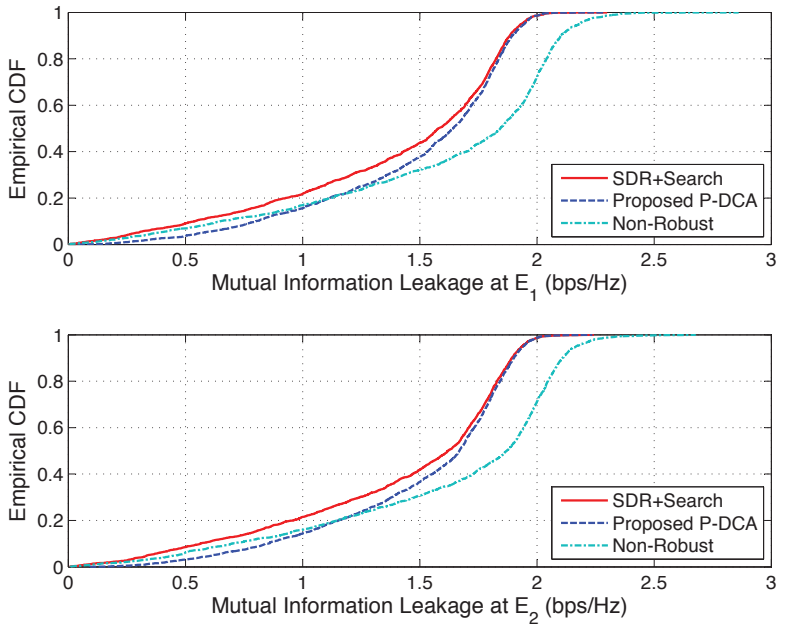



**Fig. 5.3** Empirical CDFs of mutual information leakage at eves. The legitimate S is transmitting at $R_d = 2$ bps/Hz.

robust design leads to a frequent violation of the secrecy constraints, namely for more than 20% of the realizations. Considering the practical MCS with finite coding block length, a proper selection of $\kappa$ would lead to sufficiently high block error rate at eves. Although the proposed method can prevent the eves from perfectly decoding the information signals, we need to point out the use of the secrecy constraints does not guarantee perfect secrecy from the information theoretical perspective. However, we can view our design as a means to cause additional confusion to eves.

### Reliability

Having verified the secrecy of the proposed solutions, we now compare the transmission reliability in terms of the achieved SINR at D. In Fig. 5.4, $\mathtt{SINR}_{\mathtt{D}}$ for a set of 50 independent experiments are plotted. The curve labeled "Nullspace Beamforming" refers to the method where R first nullifies eves' reception by first projecting its received signal onto the null space of $[\hat{\mathbf{g}}_{2,1}, \cdots, \hat{\mathbf{g}}_{2,K}]$ and then performs AF relaying. Therefore, the method is only applicable when $N_{\mathtt{R}} > K$. Two cases $K = 2$ and $K = 4$ are considered. In both cases, we observe that the performance of the proposed penalized DC algorithm is very close to the SDR-based benchmarker. In the case of $K = 2$, the proposed solution significantly outperforms the nullspace beamforming method.

We then study how different system configurations impact the achieved SINR by different approaches. In the left sub-figure of Fig. 5.5, the achieved SINR of the proposed solutions and the nullspace beamforming is plotted as a function of the number of antenna elements employed at R. Two sizes of uncertainty regions are considered with $\varepsilon_1 = \varepsilon_2 = 0.1$ and $\varepsilon_1 = \varepsilon_2 = 0.2$. In both scenarios, the achieved SINR monotonically increases as $N_{\mathtt{R}}$ increases due to the higher diversity one can exploit from the antenna array. Again, both the proposed solutions consistently exhibit better performance than the nullspace beamforming. Notice also when more channel uncertainties are now present ($\varepsilon_1 = \varepsilon_2 = 0.2$), the legitimate UEs are confined to relatively low transmission power to satisfy the robust secrecy constraints, leading to lower received SINR at D. In the right sub-figure of Fig. 5.5, the impact of different number of eves on the achieved SINR is assessed. The SINR monotonically decreases when there are more eves around and therefore, the legitimate UEs have to lower their transmission power to prevent the information leakage more carefully. For completeness, in Fig. 5.6, we also investigate how robustly the proposed solutions can

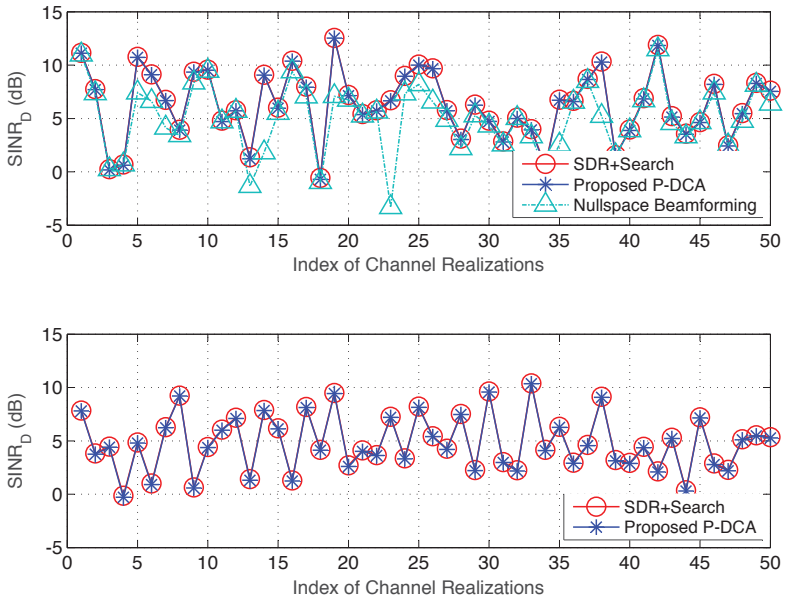



**Fig. 5.4** Achieved SINR at D. Top sub-figure: $K = 2$ eves. Bottom sub-figure: $K = 4$ eves.

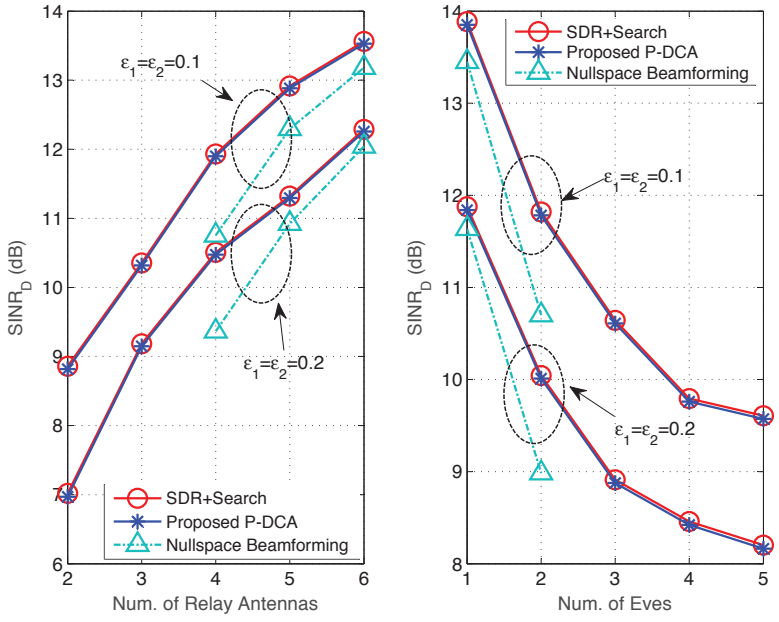



**Fig. 5.5** Achieved SINR at D. Left sub-figure: $\mathtt{SINR}_{\mathtt{D}}$ versus $N_{\mathtt{R}}$. $K = 3$ eves are considered. Right sub-figure: $\mathtt{SINR}_{\mathtt{D}}$ versus $K$.

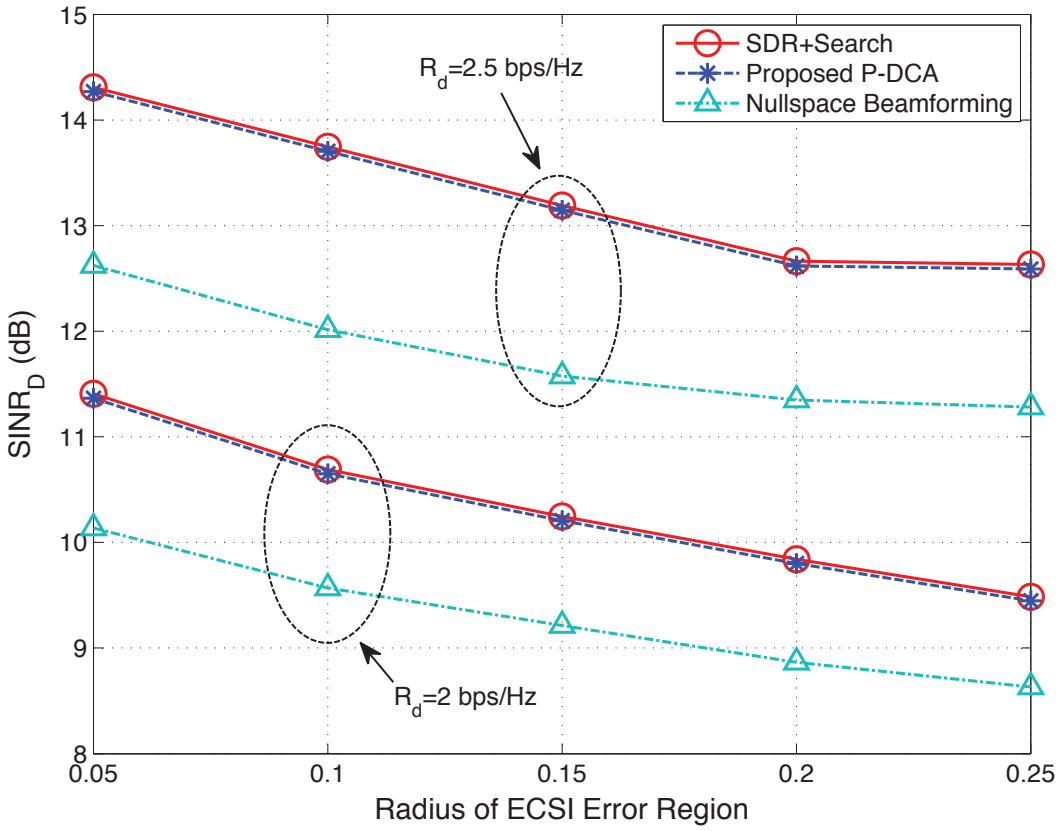



**Fig. 5.6** Achieved SINR at D as a function of size of uncertainty region. Two data rates of legitimate UEs are considered, namely, $R_d = 2$bps/Hz and $R_d = 2.5$bps/Hz.

behave against the ECSI errors by varying the sizes of the channel uncertainty regions. Again, the results are as expected and showing the superiority of our proposed solutions.

### Computational Complexity

Last but not least, we need to justify the lower complexity of the proposed penalized DC algorithm as compared to the SDR-based benchmarker proposed in Section 5.4. The averaged solver time over 100 independent realizations is shown in Table 5.1 for different values of $N_{\mathrm{R}}$ and $K$. It is observed that the solver time for the SDR approach scales very fast with increases in $N_{\mathrm{R}}$ and $K$, which is consistent with the worst-case complexity analysis in Section 5.4. In the meantime, the solver time of the proposed penalized DC algorithm

**Table 5.1** Average solver time (in seconds) for different algorithms

| Num. of eves $K$ | | Alg. | Num. of Relay Ant. $N_{\mathtt{R}}$ = 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|---|
| | 2 | SDR | 1.31 | 4.44 | 22.44 | 111.26 | 553.11 |
| | | P-DCA | 0.69 | 1.42 | 2.58 | 4.73 | 8.38 |
| | 3 | SDR | 1.59 | 5.95 | 28.61 | 141.54 | 680.90 |
| | | P-DCA | 0.88 | 1.80 | 3.45 | 6.52 | 11.31 |
| | 4 | SDR | 1.83 | 7.14 | 33.29 | 165.29 | 798.69 |
| | | P-DCA | 1.05 | 2.24 | 4.60 | 8.01 | 14.09 |
| | 5 | SDR | 2.19 | 8.38 | 41.23 | 203.20 | 924.32 |
| | | P-DCA | 1.20 | 2.66 | 5.43 | 9.95 | 17.12 |

increases more slowly.

## 5.6 Conclusions

In this chapter, robust design of secure MIMO relaying in the presence of multiple eves was studied. We jointly optimized the power of S, the AF matrix and covariance of AN at R to maximize the received SINR at D while imposing a set of mutual information leakage-based secrecy constraints. With only imperfect ECSI, the resultant problem has been shown to be non-convex and challenging. A computationally efficient sub-optimal solution relying on the new penalized DC algorithmic framework was developed. This algorithm is capable of finding a stationary solution to a general non-convex SDP representable by a PSD DC program. The latter can be efficiently solved by the penalized DC algorithm without finding a non-trivial feasible initialization. To benchmark the proposed scheme, an SDR-based approach was also proposed, which yields an upper bound of the secure MIMO relaying problem, however, with significantly higher complexity. We compared the performance of the proposed algorithm and the benchmarking schemes using a few numerical examples. It shows that the proposed solutions yield a significantly better performance than the non-robust and null-space beamforming methods. In addition, the penalized DC algorithm often reaches performance close to the SDR-based approach.

# Chapter 6

# Concluding Remarks

## 6.1 Summary

Wireless MIMO relaying can improve the QoS of destination users in multi-user interference-limited networks. To fully leverage the performance benefits provided by multiple antennas, MIMO transceivers employed by the sources, relays and destinations need to be carefully designed. This thesis focused on the joint linear MIMO AF transceiver design for multi-user relaying networks by taking into consideration some crucial practical constraints such as imperfect CSI, network energy consumption and transmission secrecy. Specifically, the main contributions of the thesis are summarized as follows.

In Chapter 3, we considered the problem of MIMO AF relaying for multi-user networks, where each source transmits multiple substreams to its corresponding destination with the assistance of multiple relays. Assuming realistic imperfect CSI of all the source-relay and relay-destination links, we proposed a robust optimization framework for the joint design of the source TPCs, relay AF transceiver matrices and destination receive filters. Specifically, two well-known CSI error models were considered, namely the *statistical* and the *norm-bounded* error models. We considered the first design problem of minimizing the maximum per-stream MSE subject to the source and relay power constraints (*min-max problem*). The statistically robust and worst-case robust versions of this problem, which respectively take into account the statistical and norm-bounded CSI errors, were formulated. Both the resultant optimization problems are non-convex and therefore, numerical algorithms were proposed by resorting to the iterative BCD-type update along with optimization techniques such as SDP and SOCP. We then considered the second design problem of

minimizing the maximum per-relay power subject to the QoS constraints for each substream and the source power constraints (*QoS problem*). The proposed algorithms for the min-max problem can correspondingly be adapted to solve the QoS problem. However, finding a proper initialization for the QoS problem is a difficult task and therefore, an efficient initial feasibility search algorithm was proposed based on the relationship between the feasibility check and the min-max problems. Numerical results showed that the proposed algorithms converge in around 30 iterations. The robust transceiver design approach can achieve an improved robustness against different types of CSI errors compared to non-robust approaches, which treat the estimated CSI as the true CSI and ignore the effects of CSI errors.

In Chapter 4, we addressed the design of a multi-user relaying sub-network within a C-RAN from an energy-efficient perspective. In the relaying sub-network, multiple source-destination pairs communicate with the assistance of multiple RRHs serving as relays, which are connected to a centralized processing node, i.e., the BBU pool. Exploiting the flexible centralized processing structure of C-RAN, where RRHs can be adaptively activated/deactivated, we formulated the problem as a QoS-based network energy minimization problem via joint RRH selection and relaying matrix optimization. Since the resultant optimization problem is non-convex and mathematically challenging, we proposed an iterative solution based on the concept of the re-weighted $l_1$-norm along with a BCD-type algorithm. Analysis showed that the proposed algorithm converges to a so-called Nash point as opposed to a local optimal point. Based on the outputs of the algorithm, the subset of active RRHs was then determined in a single attempt by thresholding a group sparsity pattern associated with the set of all RRH relaying matrices. To circumvent a potentially undesirable condition, where the selected subset of RRHs fails to simultaneously satisfy all the destination users' QoS levels, we conceived a UE admission control mechanism for overcoming the associated *infeasibility* problem. The proposed energy-efficient relaying scheme was compared with the conventional relaying scheme under a generic relaying subnetwok and a HetNet. In both scenarios, it has been showed that the proposed joint RRH selection and optimization approach leads to a small number of active RRHs. For example, to support 6 source-destination UE pairs, on average 6 out of a total of 10 RRHs can be switched off, yielding a saving of about 35% of the network energy consumption.

In Chapter 5, we designed a secure relay transmission strategy in a network, where a pair of legitimate UEs communicate with the aid of an MIMO relay in the presence of multiple

eavesdroppers. The relay adopts AF scheme to forward the confidential information from the source to the legitimate destination while emitting AN to degrade the reception quality of the eavesdroppers. Assuming that the ECSI errors reside in a predefined spherical region, we jointly optimized the power of the source UE, the relay AF matrix and the covariance of the AN transmitted by the relay, in order to maximize the received SINR at the destination, while imposing a set of *robust secrecy constraints* at the eavesdroppers. The formulated optimization problem can be represented as a nonlinear non-convex SDP with a bilinear equality constraint. Such a class of problems are in general difficult to solve with tractable computational complexity. To this end, we proposed a new penalized DC algorithm, which is specifically designed for solving this class of non-convex SDPs. One of the features of the proposed algorithm is that it eliminates the need for a non-trivial feasible initialization as required by the conventional DC algorithm [118]. We explicitly proved that the solution sequence generated by the algorithm converges to a stationary point of the original non-convex SDP. We further showed how this penalized DC framework can be invoked for solving the robust secure relaying problem. To benchmark the proposed algorithm, we subsequently proposed an SDR-based exhaustive search approach, which yields an upper bound of the secure relaying performance, however, with significantly higher complexity. Simulation results showed that the proposed solution can ensure the secrecy of the relay-aided transmission and significantly improving the robustness towards the ECSI errors compared to the non-robust counterparts. It was also demonstrated that the proposed penalized DC algorithm yields a performance close to the upper bound at a significantly reduced complexity.

## 6.2 Potential Future Works

In this section, we discuss some potential future works, which are closely related to the contributions presented in this thesis.

In Chapter 3, the proposed joint transceiver optimization algorithms for multi-user MIMO relaying networks take into account two types of CSI error models, namely, statistical (Gaussian) and norm-bounded models. As mentioned in Chapter 1, one main source that causes CSI errors during the channel reporting stage is the quantization error. For example, CSI can be quantized using channel direction information (CDI) and channel magnitude information (CMI). To report the CDI, a pre-selected codebook can be

used, which can be constructed from random vector quantization (RVQ) [128–130], discrete Fourier transform (DFT) [131] or Grassmannian manifold [132, 133]. For instance, when RVQ is used, the statistical properties of the CSI quantization errors have been analyzed in [129, 130]. It may be possible to develop a new statistically robust transceiver optimization approach for the case of RVQ. The idea behind this new approach would be similar to the one proposed in Section 3.2, e.g., the objective would be to minimize the expected MSE under the CSI quantization errors subject to power constraints. However, instead of assuming that the CSI errors are Gaussian distributed [c.f. Section 3.2], new statistical properties of the CSI errors due to the RVQ quantization [129, 130] need to be exploited, which may lead to different problem formulations and solutions.

In Chapter 4, the joint relay selection and transceiver optimization in a C-RAN was developed under the assumption of perfect CSI for all links between UEs and RRHs. However, as mentioned in Sections 2.4 and 3.1, the CSI collected at the BBU pool in a C-RAN is subject to various sources of imperfections such as the estimation errors at the receiver side, the quantization errors due to the finite-rate feedback and the feedback delay. Therefore, it is desirable to take into the consideration the effects of CSI error in the design of energy efficient relaying. For example, assuming that the statistical properties of the CSI errors are known [c.f. Section 3.2], a robust version of the energy efficient relaying problem in 4.6 may be expressed as

$$\min_{\mathbf{u},\{\mathbf{W}_l\}} \quad \sum_{l=1}^{L} \left\| \|\mathbf{W}_l\|_F^2 \right\|_0 P_{c,l} + \mathbb{E}\left\{ \sum_{l=1}^{L} P_{t,l}\left( \{\Delta \mathbf{h}_{l,k}\}_{k=1}^{K} \right) \right\} + \|\mathbf{u}\|_2^2 \tag{6.1a}$$

$$\text{s.t.} \quad \mathbb{E}\left\{ \mathtt{MSE}_k\left( u_k, \{\mathbf{W}_l\}; \{\Delta \mathbf{g}_{k,l}\}_{l=1}^{L}, \{\Delta \mathbf{h}_{l,k}\} \right) \right\} \leq \gamma_k, \ \forall k \in \mathcal{K} \tag{6.1b}$$

$$\mathbb{E}\left\{ P_{t,l}\left( \{\Delta \mathbf{h}_{l,k}\}_{k=1}^{K} \right) \right\} \leq P_{l,\max}, \ \forall l \in \mathcal{L}, \tag{6.1c}$$

where $\Delta \mathbf{h}_{l,k}$ and $\Delta \mathbf{g}_{k,l}$ denote the CSI errors and the expectations are computed with respect to the CSI errors. A robust version of the block coordinate re-weighted $l_1$-norm minimization algorithm [c.f. Algorithm 4.1] may be developed with the aid of convex approximation techniques depending on the forms of the computed expectations. Another possible extension of the works in Chapter 4 may be the consideration of other CSI error models, e.g., norm-bounded and RVQ quantization, in the design of energy efficient relaying.

In Chapter 5, the secure relaying strategies were designed for a network model consisting of one source, one relay, one destination and multiple eavesdroppers. It may be

desirable to also consider multi-user networks, which may bring additional challenges. Secure relaying with improved performance needs to be designed to tackle the additional information leakage due to the concurrent transmission from multiple users. Furthermore, as we discussed in Chapter 5, the proposed penalized DC algorithm is a general scheme, which can solve a large class of non-convex SDPs. In addition to the robust secure relaying problem formulated in Section 5.3, it may be useful to exploit the possibility of applying the penalized DC algorithm to solve various optimization problems appearing in other applications/disciplines.

# Appendix A

# Appendices of Chapter 3

## A.1 Transformation of the QCLP into a standard SOCP

By exploiting the separable structure of (3.37) and the properties of quadratic terms, the problem can be recast as

$$\min_{\substack{t,\{\mathbf{f}_k\},\\ \{\boldsymbol{\lambda}^{k,l}\},\{\boldsymbol{\theta}^m\}}} \quad t \tag{A.1a}$$

$$\text{s.t.} \quad \left\| \left(\mathbf{A}_{1,q}^{k,l}\right)^{1/2} \mathbf{f}_q \right\| \le \lambda_q^{k,l}, \quad \forall q,k \in \mathcal{K}, q \ne k, l \in \mathcal{D}_k \tag{A.1b}$$

$$\left\| \left(\mathbf{A}_{1,k}^{k,l}\right)^{1/2} \mathbf{f}_k - \left(\mathbf{A}_{1,k}^{k,l}\right)^{-1/2} \mathbf{a}_2^{k,l} \right\| \le \lambda_k^{k,l}, \quad \forall k \in \mathcal{K}, l \in \mathcal{D}_k \tag{A.1c}$$

$$\left\|\boldsymbol{\lambda}^{k,l}\right\|^2 - (\mathbf{a}_2^{k,l})^H (\mathbf{A}_{1,k}^{k,l})^{-1} \mathbf{a}_2^{k,l} + a_3^{k,l} \le \frac{t}{\kappa_{k,l}}, \quad \forall k \in \mathcal{K}, l \in \mathcal{D}_k \tag{A.1d}$$

$$\left\|(\mathbf{A}_{4,k}^m)^{1/2} \mathbf{f}_k\right\| \le \theta_k^m, \ \forall k \in \mathcal{K}, m \in \mathcal{M} \tag{A.1e}$$

$$\|\boldsymbol{\theta}^m\| \le \sqrt{\eta_{\mathtt{R},m}}, \ \forall m \in \mathcal{M} \tag{A.1f}$$

$$\|\mathbf{f}_k\| \le \sqrt{P_{\mathtt{S},k}^{\max}}, \ \forall k \in \mathcal{K}, \tag{A.1g}$$

where $\boldsymbol{\lambda}^{k,l} = [\lambda_1^{k,l}, \cdots, \lambda_K^{k,l}]^T$, $\boldsymbol{\theta}^m = [\theta_1^m, \cdots, \theta_K^m]^T$ and $t$ are auxiliary variables. The main difficulty in solving this problem is with (A.1d), which is a so-called *hyperbolic constraint* [95], while the remaining constraints are already in the form of SOC.

To tackle (A.1d), we observe that for any $\mathbf{x}$ and $y, z \geq 0$, the following equation holds

$$\|\mathbf{x}\|^2 \leq yz \iff \left\| \begin{bmatrix} 2\mathbf{x} \\ y - z \end{bmatrix} \right\| \leq y + z. \tag{A.2}$$

We can apply (A.2) to transform (A.1d) into

$$\left\| \begin{bmatrix} 2\boldsymbol{\lambda}^{k,l} \\ \frac{t}{\kappa_{k,l}} + \left(\mathbf{a}_2^{k,l}\right)^H \left(\mathbf{A}_{1,k}^{k,l}\right)^{-1} \mathbf{a}_2^{k,l} - a_3^{k,l} - 1 \end{bmatrix} \right\| \leq \frac{t}{\kappa_{k,l}} + \left(\mathbf{a}_2^{k,l}\right)^H \left(\mathbf{A}_{1,k}^{k,l}\right)^{-1} \mathbf{a}_2^{k,l} - a_3^{k,l} + 1. \tag{A.3}$$

Therefore, substituting (A.1d) by (A.3), we can see that (A.1) is in the form of a standard SOCP.

## A.2 Proof of Proposition 3.1

First, we define $\boldsymbol{\mathcal{T}}_k \triangleq [\boldsymbol{\mathcal{T}}_{k,1}, \cdots, \boldsymbol{\mathcal{T}}_{k,K}]$ and $\boldsymbol{\mathcal{G}}_k \triangleq [\sigma_{\mathtt{R},1}\boldsymbol{\mathcal{G}}_{k,1}, \cdots, \sigma_{\mathtt{R},M}\boldsymbol{\mathcal{G}}_{k,M}]$. We exploit the fact that for any vectors $\{\mathbf{a}_k\}_{k=1}^N$, the following identity holds:

$$\sum_{k=1}^{N} \|\mathbf{a}_k\|^2 = \left\| \left[\mathbf{a}_1^T, \cdots, \mathbf{a}_N^T\right] \right\|^2. \tag{A.4}$$

The per-stream MSE (3.13) can subsequently be expressed as

$$\begin{aligned} \varepsilon_{k,l} = & \left\| \mathbf{u}_{k,l}^H \boldsymbol{\mathcal{T}}_k + \sum_{m=1}^{M} \mathbf{u}_{k,l}^H \Delta\mathbf{G}_{k,m} \left[\boldsymbol{\mathcal{W}}_{m,1}\mathbf{F}_1, \cdots, \boldsymbol{\mathcal{W}}_{m,K}\mathbf{F}_K\right] \right. \\ & + \sum_{q=1}^{K}\sum_{m=1}^{M} \left[\mathbf{0}_{1\times\sum_{t=1}^{q} d_t}, \mathbf{u}_{k,l}^H \boldsymbol{\mathcal{G}}_{k,m}\, \Delta\mathbf{H}_{m,q}\mathbf{F}_q, \mathbf{0}_{1\times\sum_{q+1}^{K} d_t}\right] \Big\|^2 \\ & + \left\| \sum_{m=1}^{M} \left[\mathbf{0}_{1\times\sum_{p=1}^{m-1} N_{\mathtt{R},p}}, \mathbf{u}_{k,l}^H \Delta\mathbf{G}_{k,m}\mathbf{W}_m, \mathbf{0}_{1\times\sum_{p=m+1}^{M} N_{\mathtt{R},p}}\right] \mathbf{u}_{k,l}^H \boldsymbol{\mathcal{G}}_k \right\|^2 + \sigma_{\mathtt{D},k}^2 \|\mathbf{u}_{k,l}\|^2 \end{aligned} \tag{A.5}$$

Upon applying the identity $\mathtt{vec}^T(\mathbf{ABC}) = \mathtt{vec}(\mathbf{B})^T\left(\mathbf{C}\otimes\mathbf{A}^T\right)$ to (A.5), we arrive at:

$$\varepsilon_{k,l} = \left\|\mathbf{u}_{k,l}^{\mathbf{H}}\boldsymbol{\mathcal{T}}_k - \overline{\mathbf{e}}_{k,l}^T + \sum_{m=1}^{M}\mathbf{g}_{k,m}^T\mathbf{C}_{1,m}^{k,l} + \sum_{m,q}\mathbf{h}_{m,q}^T\mathbf{D}_{m,q}^{k,l}\right\|^2 + \left\|\mathbf{u}_{k,l}^{\mathbf{H}}\boldsymbol{\mathcal{G}}_k + \sum_{m=1}^{M}\mathbf{g}_{k,m}^T\mathbf{C}_{2,m}^{k,l}\right\|^2 + \left\|\sigma_{\mathtt{D},k}\mathbf{u}_{k,l}^{\mathbf{H}}\right\|^2, \tag{A.6}$$

where $\mathbf{h}_{m,k} \triangleq \mathtt{vec}(\Delta\mathbf{H}_{m,k})$ and $\mathbf{g}_{k,m} \triangleq \mathtt{vec}(\Delta\mathbf{G}_{k,m})$ denote the vectorized CSI errors, $\overline{\mathbf{e}}_{k,l} \triangleq \left[\mathbf{0}_{1\times\sum_{t=1}^{k-1}d_t}, \mathbf{e}_{k,l}^T, \mathbf{0}_{1\times\sum_{t=k+1}^{K}d_t}\right]^T$, and the following matrices have also been introduced:

$$\mathbf{C}_{1,m}^{k,l} \triangleq \left[(\boldsymbol{\mathcal{W}}_{m,1}\mathbf{F}_1)\otimes\mathbf{u}_{k,l}^*, \cdots, (\boldsymbol{\mathcal{W}}_{m,K}\mathbf{F}_K)\otimes\mathbf{u}_{k,l}^*\right] \tag{A.7}$$

$$\mathbf{C}_{2,m}^{k,l} \triangleq \left[\mathbf{0}_{N_{\mathtt{D},k}N_{\mathtt{R},m}\times\sum_{p=1}^{m-1}N_{\mathtt{R},p}}, \mathbf{W}_m\otimes\mathbf{u}_{k,l}^*, \mathbf{0}_{N_{\mathtt{D},k}N_{\mathtt{R},m}\times\sum_{p=m+1}^{M}N_{\mathtt{R},p}}\right] \tag{A.8}$$

$$\mathbf{D}_{m,q}^{k,l} \triangleq \left[\mathbf{0}_{N_{\mathtt{S},q}N_{\mathtt{R},m}\times\sum_{t=1}^{q-1}d_t}, \mathbf{F}_q\otimes\left(\boldsymbol{\mathcal{G}}_{k,m}^T\mathbf{u}_{k,l}^*\right), \mathbf{0}_{N_{\mathtt{S},q}N_{\mathtt{R},m}\times\sum_{t=q+1}^{K}d_t}\right]. \tag{A.9}$$

Again, by exploiting the property in (A.4), we can write (A.6) in a more compact form as follows:

$$\varepsilon_{k,l} = \Bigg\| \underbrace{\left[\mathbf{u}_{k,l}^{\mathbf{H}}\boldsymbol{\mathcal{T}}_k - \overline{\mathbf{e}}_{k,l}, \mathbf{u}_{k,l}^{\mathbf{H}}\overline{\boldsymbol{\mathcal{G}}}_k, \sigma_{\mathtt{D},k}\mathbf{u}_{k,l}^{\mathbf{H}}\right]}_{\boldsymbol{\theta}_{k,l}} + \sum_{m=1}^{M}\mathbf{g}_{k,m}^T\underbrace{\left[\mathbf{C}_{1,m}^{k,l}, \mathbf{C}_{2,m}^{k,l}, \mathbf{0}_{N_{\mathtt{D},k}N_{\mathtt{R},m}\times N_{\mathtt{D},k}}\right]}_{\boldsymbol{\Theta}_m^{k,l}} + \sum_{m=1}^{M}\sum_{q=1}^{K}\mathbf{h}_{m,q}^T\underbrace{\left[\mathbf{D}_{m,q}^{k,l}, \mathbf{0}_{N_{\mathtt{R},m}N_{\mathtt{S},q}\times N_{\mathtt{R}}+N_{\mathtt{D},k}}\right]}_{\boldsymbol{\Phi}_{m,q}^{k,l}}\Bigg\|^2. \tag{A.10}$$

Substituting (A.10) into (3.46b), we can express (3.46b) as

$$\left(\boldsymbol{\theta}_{k,l} + \sum_{m=1}^{M}\mathbf{g}_{k,m}^T\boldsymbol{\Theta}_m^{k,l} + \sum_{m=1}^{M}\sum_{q=1}^{K}\mathbf{h}_{m,q}^T\boldsymbol{\Phi}_{m,q}^{k,l}\right)^H\left(\boldsymbol{\theta}_{k,l} + \sum_{m=1}^{M}\mathbf{g}_{k,m}^T\boldsymbol{\Theta}_m^{k,l} + \sum_{m=1}^{M}\sum_{q=1}^{K}\mathbf{h}_{m,q}^T\boldsymbol{\Phi}_{m,q}^{k,l}\right) \leq \frac{t}{\kappa_{k,l}}, \tag{A.11}$$

where the uncertain blocks $\mathbf{h}_{m,k}$ and $\mathbf{g}_{k,m}$ should satisfy $\|\mathbf{h}_{m,k}\|_S = \|\mathbf{h}_{m,k}\| \leq \xi_{m,k}$ and $\|\mathbf{g}_{k,m}\|_S = \|\mathbf{g}_{k,m}\| \leq \eta_{k,m}$, respectively. Through a direct application of Lemma 3.1, (A.11) can readily be recast as (3.49) where the nonnegativity of $\boldsymbol{\tau}_{k,l}^{\mathrm{g}}$ and $\boldsymbol{\tau}_{k,l}^{\mathrm{h}}$ has been implicitly included in the positive semidefinite nature of $\mathbf{Q}_{k,l}$.

# Appendix B

# Appendices of Chapter 4

## B.1 Proof of the First Step in Theorem 4.1

We have to show that the sequence $\left\{(\mathbf{w}^{(n)}, \mathbf{u}^{(n)})\right\}$ is a square summable sequence such that $\lim_{n\to\infty}\|\mathbf{w}^{(n)} - \mathbf{w}^{(n-1)}\|_2^2 = 0$ and $\lim_{n\to\infty}\|\mathbf{u}^{(n)} - \mathbf{u}^{(n-1)}\|_2^2 = 0$. Based on the majorization relation in (4.29), we have

$$\begin{aligned}
&\mathcal{F}(\mathbf{w}^{(n)}) + \sum_{l=1}^{L} \mathbf{w}_l^{(n)H} \mathbf{\Psi}_l \mathbf{w}_l^{(n)} + \mathbf{u}^{(n)H}\mathbf{u}^{(n)} \\
&\quad \leq \tilde{\mathcal{F}}(\mathbf{w}^{(n)}; \mathbf{w}^{(n-1)}) + \sum_{l=1}^{L} \mathbf{w}_l^{(n)H} \mathbf{\Psi}_l \mathbf{w}_l^{(n)} + \mathbf{u}^{(n)H}\mathbf{u}^{(n)} \\
&\quad \leq \mathcal{F}(\mathbf{w}^{(n-1)}) + \sum_{l=1}^{L} \mathbf{w}_l^{(n-1)H} \mathbf{\Psi}_l \mathbf{w}_l^{(n-1)} + \mathbf{u}^{(n-1)H}\mathbf{u}^{(n-1)} \\
&\qquad - \frac{\tau(\mathbf{w})}{2}\|\mathbf{w}^{(n)} - \mathbf{w}^{(n-1)}\|_2^2 - \frac{\tau(\mathbf{u})}{2}\|\mathbf{u}^{(n)} - \mathbf{u}^{(n-1)}\|_2^2, \qquad \text{(B.1)}
\end{aligned}$$

where the first inequality is due to the majorization relation in (4.29) while the second inequality results from $\mathcal{F}(\mathbf{w}^{(n-1)}) = \mathcal{F}(\mathbf{w}^{(n-1)}; \mathbf{w}^{(n-1)})$ and the strong convexity of $\sum_{l=1}^{L} \mathbf{w}_l^H \mathbf{\Psi}_l \mathbf{w}_l$ and $\mathbf{u}^H\mathbf{u}$ with parameters $\tau(\mathbf{w}) > 0$ and $\tau(\mathbf{u}) > 0$, respectively. Then

summing the above inequality over $n$ from 0 to $\bar{n}$, we arrive at

$$\begin{aligned}
&\mathcal{F}(\mathbf{w}^{(0)}) + \sum_{l=1}^{L} \mathbf{w}_l^{(0)^H} \mathbf{\Psi}_l \mathbf{w}_l^{(0)} + \mathbf{u}^{(0)^H} \mathbf{u}^{(0)} \\
&\quad - \left( \mathcal{F}(\mathbf{w}^{(\bar{n})}) + \sum_{l=1}^{L} \mathbf{w}_l^{(\bar{n})^H} \mathbf{\Psi}_l \mathbf{w}_l^{(\bar{n})} + \mathbf{u}^{(\bar{n})^H} \mathbf{u}^{(\bar{n})} \right) \\
&\geq \sum_{n=0}^{\bar{n}} \left( \frac{\tau(\mathbf{w})}{2} \|\mathbf{w}^{(n)} - \mathbf{w}^{(n-1)}\|_2^2 + \frac{\tau(\mathbf{u})}{2} \|\mathbf{u}^{(n)} - \mathbf{u}^{(n-1)}\|_2^2 \right). \qquad \text{(B.2)}
\end{aligned}$$

It is not difficult to observe that the objective function in (4.28) is monotonically decreasing after each iteration and lower bounded at least by zero, so that the left-hand side of the above equation is a finite positive number. Upon letting $\bar{n} \to \infty$, we therefore conclude that $\lim_{n\to\infty} \|\mathbf{w}^{(n)} - \mathbf{w}^{(n-1)}\|_2^2 = 0$ and $\lim_{n\to\infty} \|\mathbf{u}^{(n)} - \mathbf{u}^{(n-1)}\|_2^2 = 0$.

# Appendix C

# Appendices of Chapter 5

## C.1 Proof of Lemma 5.2

From Step 2 of Algorithm 5.1, we obtain $(\mathbf{x}^{(n+1)}, \mathbf{S}^{(n+1)}$ is an optimal solution of the convex sub-problem (5.28) and $\mathbf{\Phi}_i^{(n+1)} \succeq \mathbf{0}, \mathbf{Z}_i^{(n+1)} \succeq \mathbf{0}$ for $i \in \mathcal{I}$ are the corresponding Lagrange multipliers. Since (5.28) is convex and strictly feasible, i.e., the Slater's constraint qualification holds, the optimal primal-dual pair must satisfy the following sufficient generalized KKT conditions:

$$\mathbf{0} \in \nabla f_0(\mathbf{x}^{(n+1)}) - \nabla g_0(\mathbf{x}^{(n)}) + \sum_{i=1}^{I} \left( \left( \boldsymbol{\mathcal{D}}\boldsymbol{\mathcal{F}}_i(\mathbf{x}^{(n+1)}) - \boldsymbol{\mathcal{D}}\boldsymbol{\mathcal{G}}_i(\mathbf{x}^{(n)}) \right) * \mathbf{\Phi}_i^{(n+1)} \right) + \mathcal{N}(\Omega, \mathbf{x}^{(n+1)}) \tag{C.1a}$$

$$\tau^{(n)}\mathbf{I} - \mathbf{\Phi}_i^{(n+1)} - \mathbf{Z}_i^{(n+1)} = \mathbf{0}, \ i \in \mathcal{I} \tag{C.1b}$$

$$\boldsymbol{\mathcal{F}}_i(\mathbf{x}^{(n+1)}) - \boldsymbol{\mathcal{G}}_i(\mathbf{x}^{(n)}) - \boldsymbol{\mathcal{D}}\boldsymbol{\mathcal{G}}(\mathbf{x}^{(n)})(\mathbf{x}^{(n+1)} - \mathbf{x}^{(n)}) \preceq \mathbf{S}_i^{(n+1)}, \ i \in \mathcal{I} \tag{C.1c}$$

$$\mathrm{Tr}\left( \mathbf{\Phi}_i^{(n+1)} \left( \boldsymbol{\mathcal{F}}_i(\mathbf{x}^{(n+1)}) - \boldsymbol{\mathcal{G}}_i(\mathbf{x}^{(n)}) - \boldsymbol{\mathcal{D}}\boldsymbol{\mathcal{G}}(\mathbf{x}^{(n)})(\mathbf{x}^{(n+1)} - \mathbf{x}^{(n)}) - \mathbf{S}_i^{(n+1)} \right) \right) = 0, \ i \in \mathcal{I} \tag{C.1d}$$

$$\mathbf{x}^{(n+1)} \in \Omega, \ \mathbf{S}_i^{(n+1)} \succeq \mathbf{0}, \ \mathbf{\Phi}_i^{(n+1)} \succeq \mathbf{0}, \ \mathbf{Z}_i^{(n+1)} \succeq \mathbf{0}, \ \mathrm{Tr}(\mathbf{S}_i^{(n+1)}\mathbf{Z}_i^{(n+1)}) = 0, \ i \in \mathcal{I}, \tag{C.1e}$$

where $\mathbf{A}*$ denotes the adjoint operator of $\mathbf{A} = \sum_{i=1}^{n} x_i \mathbf{A}_i$ with $\mathbf{A}_i \in \mathbb{H}^p$ for $i = 1, \cdots, n$, i.e., $\mathbf{A} * \mathbf{Z} = [\mathrm{Tr}(\mathbf{A}_1\mathbf{Z}), \cdots, \mathrm{Tr}(\mathbf{A}_n\mathbf{Z})]^T$ for any $\mathbf{Z} \in \mathbb{H}^p$. $\mathcal{N}(\Omega, \mathbf{x})$ denotes the normal cone of $\Omega$ at $\mathbf{x}$ defined as:

$$\mathcal{N}(\Omega, \mathbf{x}) \triangleq \{\mathbf{w} \in \mathbb{C}^n | \mathbf{w}^H(\mathbf{x} - \mathbf{y}) \geq 0, \forall \mathbf{y} \in \Omega\}$$

To simply the notation, let us define $\hat{\varphi}^{(n)}(\mathbf{x}, \mathbf{S}) = \varphi(\mathbf{x}) + + \sum_{i=1}^{I} \tau^{(n)} \mathrm{Tr}(\mathbf{S}_i) = f_0(\mathbf{x}) - g_0(\mathbf{x}) + \sum_{i=1}^{I} \tau^{(n)} \mathrm{Tr}(\mathbf{S}_i)$.

First multiplying the both sides of (C.1a) by $(\mathbf{x}^{(n)} - \mathbf{x}^{(n+1)})^T$ and re-arranging the consequence, we obtain

$$\begin{aligned} &\left(\nabla f_0^T(\mathbf{x}^{(n+1)}) - \nabla g_0^T(\mathbf{x}^{(n)})\right)(\mathbf{x}^{(n)} - \mathbf{x}^{(n+1)}) \\ &\quad + \sum_{i=1}^{I} \left[\left(\boldsymbol{\mathcal{D}}\boldsymbol{\mathcal{F}}_i(\mathbf{x}^{(n+1)}) - \boldsymbol{\mathcal{D}}\boldsymbol{\mathcal{G}}_i(\mathbf{x}^{(n)})\right) * \boldsymbol{\Phi}_i\right]^T (\mathbf{x}^{(n)} - \mathbf{x}^{(n+1)}) \geq 0 \end{aligned} \tag{C.2}$$

By the assumption of convexity of $f_0(\cdot)$ and $g_0(\cdot)$, we have

$$f_0(\mathbf{x}^{(n)}) \geq f_0(\mathbf{x}^{(n+1)}) + \nabla f_0^T(\mathbf{x}^{(n+1)})(\mathbf{x}^{(n)} - \mathbf{x}^{(n+1)}) + \frac{\rho_f}{2}\|\mathbf{x}^{(n+1)} - \mathbf{x}^{(n)}\|^2 \tag{C.3}$$

$$g_0(\mathbf{x}^{(n+1)}) \geq g_0(\mathbf{x}^{(n)}) + \nabla g_0^T(\mathbf{x}^{(n)})(\mathbf{x}^{(n+1)} - \mathbf{x}^{(n)}) + \frac{\rho_g}{2}\|\mathbf{x}^{(n+1)} - \mathbf{x}^{(n)}\|^2, \tag{C.4}$$

where we recall that $\rho_f \geq 0$ and $\rho_g \geq 0$ are the convexity parameters.

Combining (C.3) and (C.4) and rearranging the consequence, we further obtain

$$\left(\nabla f_0^T(\mathbf{x}^{(n+1)}) - \nabla g_0^T(\mathbf{x}^{(n)})\right)\left(\mathbf{x}^{(n)} - \mathbf{x}^{(n+1)}\right) \leq \varphi(\mathbf{x}^{(n)}) - \varphi(\mathbf{x}^{(n+1)}) - \frac{\rho_f + \rho_g}{2}\|\mathbf{x}^{(n+1)} - \mathbf{x}^{(n)}\|^2. \tag{C.5}$$

By the PSD-convexity of $\boldsymbol{\mathcal{F}}_i(\cdot)$, we obtain

$$\boldsymbol{\mathcal{F}}_i(\mathbf{x}^{(n)}) \succeq \boldsymbol{\mathcal{F}}_i(\mathbf{x}^{(n+1)}) + \boldsymbol{\mathcal{D}}\boldsymbol{\mathcal{F}}_i(\mathbf{x}^{(n+1)})(\mathbf{x}^{(n)} - \mathbf{x}^{(n+1)}), \tag{C.6}$$

which further lead to

$$\begin{aligned}&\left[\boldsymbol{\mathcal{D}}\boldsymbol{\mathcal{F}}_i(\mathbf{x}^{(n+1)}) - \boldsymbol{\mathcal{D}}\boldsymbol{\mathcal{G}}_i(\mathbf{x}^{(n)})\right](\mathbf{x}^{(n)} - \mathbf{x}^{(n+1)})\\ &\quad\preceq \boldsymbol{\mathcal{F}}_i(\mathbf{x}^{(n)}) - \boldsymbol{\mathcal{G}}_i(\mathbf{x}^{(n)}) - \left[\boldsymbol{\mathcal{F}}_i(\mathbf{x}^{(n+1)}) - \boldsymbol{\mathcal{G}}_i(\mathbf{x}^{(n)}) - \boldsymbol{\mathcal{D}}\boldsymbol{\mathcal{G}}_i(\mathbf{x}^{(n)})(\mathbf{x}^{(n+1)} - \mathbf{x}^{(n)})\right]. \qquad \text{(C.7)}\end{aligned}$$

For simplicity, let us denote the second term on the right hand side of (C.7) by $\mathbf{A}$. Multiplying the both sides of (C.7) by $\boldsymbol{\Phi}_i^{(n+1)} \succeq \mathbf{0}$ leads to

$$\begin{aligned}&\mathrm{Tr}\left(\boldsymbol{\Phi}_i^{(n+1)}\left[\boldsymbol{\mathcal{D}}\boldsymbol{\mathcal{F}}_i(\mathbf{x}^{(n+1)}) - \boldsymbol{\mathcal{D}}\boldsymbol{\mathcal{G}}_i(\mathbf{x}^{(n)})\right](\mathbf{x}^{(n)} - \mathbf{x}^{(n+1)})\right)\\ &\quad\preceq \mathrm{Tr}\left(\boldsymbol{\Phi}_i^{(n+1)}\left[\boldsymbol{\mathcal{F}}_i(\mathbf{x}^{(n)}) - \boldsymbol{\mathcal{G}}_i(\mathbf{x}^{(n)})\right]\right) + \mathrm{Tr}\left(\boldsymbol{\Phi}_i^{(n+1)}\mathbf{A}\right) \qquad \text{(C.8)}\end{aligned}$$

Noting that

$$\begin{aligned}&\mathrm{Tr}\left(\boldsymbol{\Phi}_i^{(n+1)}\left[\boldsymbol{\mathcal{D}}\boldsymbol{\mathcal{F}}_i(\mathbf{x}^{(n+1)}) - \boldsymbol{\mathcal{D}}\boldsymbol{\mathcal{G}}_i(\mathbf{x}^{(n)})\right](\mathbf{x}^{(n)} - \mathbf{x}^{(n+1)})\right)\\ &\quad= \left[\left(\boldsymbol{\mathcal{D}}\boldsymbol{\mathcal{F}}_i(\mathbf{x}^{(n+1)}) - \boldsymbol{\mathcal{D}}\boldsymbol{\mathcal{G}}_i(\mathbf{x}^{(n)})\right) * \boldsymbol{\Phi}_i\right]^T(\mathbf{x}^{(n)} - \mathbf{x}^{(n+1)}) \qquad \text{(C.9)}\end{aligned}$$

Substituting the results of (C.8), (C.9) and (C.1d) into (C.7), we have

$$\begin{aligned}&\left[\left(\boldsymbol{\mathcal{D}}\boldsymbol{\mathcal{F}}_i(\mathbf{x}^{(n+1)}) - \boldsymbol{\mathcal{D}}\boldsymbol{\mathcal{G}}_i(\mathbf{x}^{(n)})\right) * \boldsymbol{\Phi}_i\right]^T(\mathbf{x}^{(n)} - \mathbf{x}^{(n+1)})\\ &\quad\le \mathrm{Tr}\left(\boldsymbol{\Phi}_i^{(n+1)}\left[\boldsymbol{\mathcal{F}}_i(\mathbf{x}^{(n)}) - \boldsymbol{\mathcal{G}}_i(\mathbf{x}^{(n)})\right]\right) - \mathrm{Tr}\left(\boldsymbol{\Phi}_i^{(n+1)}\mathbf{S}_i^{(n+1)}\right). \qquad \text{(C.10)}\end{aligned}$$

Observing that $\boldsymbol{\mathcal{F}}_i(\mathbf{x}^{(n)}) - \boldsymbol{\mathcal{G}}_i(\mathbf{x}^{(n)}) \preceq \mathbf{S}_i^{(n)}$ and $\tau^{(n)}\mathbf{I} \succeq \boldsymbol{\Phi}_i^{(n+1)}$ [c.f., (C.1b)], (C.10) can further be derived as

$$\begin{aligned}&\left[\left(\boldsymbol{\mathcal{D}}\boldsymbol{\mathcal{F}}_i(\mathbf{x}^{(n+1)}) - \boldsymbol{\mathcal{D}}\boldsymbol{\mathcal{G}}_i(\mathbf{x}^{(n)})\right) * \boldsymbol{\Phi}_i\right]^T(\mathbf{x}^{(n)} - \mathbf{x}^{(n+1)})\\ &\quad\le \mathrm{Tr}\left(\boldsymbol{\Phi}_i^{(n+1)}\left(\mathbf{S}_i^{(n)} - \mathbf{S}_i^{(n+1)}\right)\right)\\ &\quad\le \tau^{(n)}\,\mathrm{Tr}(\mathbf{S}_i^{(n)}) - \tau^{(n)}\,\mathrm{Tr}(\mathbf{S}_i^{(n+1)}). \qquad \text{(C.11)}\end{aligned}$$

Combining (C.2), (C.5) and (C.11), we have reached:

$$\hat{\varphi}^{(n)}(\mathbf{x}^{(n)}, \mathbf{S}^{(n)}) - \hat{\varphi}^{(n)}(\mathbf{x}^{(n+1)}, \mathbf{S}^{(n+1)}) \ge \frac{\rho_f + \rho_g}{2}\|\mathbf{x}^{(n+1)} - \mathbf{x}^{(n)}\|^2. \qquad \text{(C.12)}$$

The above inequality is indeed (5.37), which therefore proves the item 1). If either $f_0$ or $g_0$

is strongly convex, i.e., $\rho_f + \rho_g > 0$, then the statement in item 2) follows directly from the above inequality, i.e., for $\Delta \mathbf{x}^{(n)} = \mathbf{x}^{(n+1)} - \mathbf{x}^{(n)} \neq \mathbf{0}$, $\hat{\varphi}^{(n)}(\mathbf{x}^{(n+1)}, \mathbf{S}^{(n+1)}) < \hat{\varphi}^{(n)}(\mathbf{x}^{(n)}, \mathbf{S}^{(n)})$.

## C.2 Proof of Theorem 5.1

We first prove scenario 1). If Algorithm 5.1 terminates after a finite number of $\check{n}$ iterations, it follows from the termination criterion that $\mathbf{x}^{(\check{n}+1)} = \mathbf{x}^{(\check{n})}$ and $\mathbf{S}_i^{(\check{n}+1)} = \mathbf{0}$ for all $i$, i.e., $\check{\mathbf{x}}$ is a feasible solution to (5.20). Letting $n = \check{n}$ and substituting the above relations into the generalized KKT conditions (C.1), we obtain the following:

$$\mathbf{0} \in \nabla f_0(\mathbf{x}^{(\check{n})}) - \nabla g_0(\mathbf{x}^{(\check{n})}) + \sum_{i=1}^{I} \left( \left( \boldsymbol{\mathcal{D}}\boldsymbol{\mathcal{F}}_i(\mathbf{x}^{(\check{n})}) - \boldsymbol{\mathcal{D}}\boldsymbol{\mathcal{G}}_i(\mathbf{x}^{(\check{n})}) \right) * \boldsymbol{\Phi}_i^{(\check{n})} \right) + \mathcal{N}(\Omega, \mathbf{x}^{(\check{n})}) \quad \text{(C.13a)}$$

$$\boldsymbol{\mathcal{F}}_i(\mathbf{x}^{(\check{n})}) - \boldsymbol{\mathcal{G}}_i(\mathbf{x}^{(\check{n})}) \preceq \mathbf{0}, \ i \in \mathcal{I} \quad \text{(C.13b)}$$

$$\mathrm{Tr}\left( \boldsymbol{\Phi}_i^{(\check{n}+1)} \left( \boldsymbol{\mathcal{F}}_i(\mathbf{x}^{(\check{n})}) - \boldsymbol{\mathcal{G}}_i(\mathbf{x}^{(\check{n})}) \right) \right) = 0, \ i \in \mathcal{I} \quad \text{(C.13c)}$$

$$\mathbf{x}^{(\check{n})} \in \Omega, \ \boldsymbol{\Phi}_i^{(\check{n}+1)} \succeq \mathbf{0}. \quad \text{(C.13d)}$$

A careful examination reveals the equivalence between (C.13) and the KKT conditions of (5.20). Therefore, it is proved that $(\mathbf{x}^{(\check{n})}, \{\boldsymbol{\Phi}_i^{(\check{n})}\}$ is a KKT point of (5.20), where $\mathbf{x}^{(\check{n})}$ is called a stationary point of (5.20) and $\{\boldsymbol{\Phi}_i^{(\check{n})}\}$ are the corresponding Lagrange multipliers.

We now proceed to prove scenario 2). The key ingredients of the proof are to show that any limit point of $\{\mathbf{x}^{(n)}\}$, say, $\bar{\mathbf{x}}$, is feasible to (5.20) and the sequence of dual variable $\{\boldsymbol{\Phi}_i^{(n)}\}$ is bounded such that there exists limit points $\bar{\boldsymbol{\Phi}}_i$ of $\{\boldsymbol{\Phi}_i^{(n)}\}$. Then we show that any primal-dual pair of the limit point $(\bar{\mathbf{x}}, \{\bar{\boldsymbol{\Phi}}_i\})$ satisfies the KKT conditions of (5.20).

To prove that any limit point $\bar{\mathbf{x}}$ is a feasible point of (5.20), we will need to rely on the following claims, whose proof can be found in Appendices C.3 and C.4, respectively:

**Claim 1.** There exists a finite iteration index $\tilde{n}$ such that

$$\tau^{(n)} = \tau^{(\tilde{n})}, \quad \forall n \geq \tilde{n}. \quad \text{(C.14)}$$

**Claim 2.** The sequence of intermediate solutions $\{\mathbf{x}^{(n)}\}$ satisfies

$$\lim_{n \to \infty} \|\mathbf{x}^{(n+1)} - \mathbf{x}^{(n)}\| = 0. \quad \text{(C.15)}$$

As indicated by the updating rule (5.30), we have

$$\tau_i^{(\tilde{n})} \geq \lambda_{\max}[\mathbf{\Phi}_i^{(\tilde{n}+1)}] + \delta_1,\ i \in \mathcal{I} \tag{C.16}$$

or equivalently, $\tau^{(n)}\mathbf{I} \succeq \mathbf{\Phi}_i^{(n+1)} + \delta_1\mathbf{I}$ for all $n \geq \tilde{n}$. Then in view of the complementary slackness (C.1b), it straightforwardly follows that $\mathbf{Z}_i^{(n+1)} \succ \mathbf{0}$. By (C.1e), we obtain $\mathbf{S}_i^{(n+1)} = \mathbf{0}$ for all $n \geq \breve{n}$, which means that $\mathbf{x}^{(n)}$ is a feasible point of (5.20) for all $n \geq \tilde{n}$. Without loss of generality, considering a subsequence $\{\mathbf{x}^{(n_j)}\}$ of $\{\mathbf{x}^{(n)}\}$, its limit point $\lim_{j\to\infty}\mathbf{x}^{(n_j)} = \bar{\mathbf{x}}$ is feasible to (5.20). Furthermore, (C.16) implies that the subsequence $\{\mathbf{\Phi}_i^{(n_j)}\}$ is bounded, and therefore we can assume that

$$\lim_{j\to\infty} \mathbf{\Phi}_i^{(n_j)} = \bar{\mathbf{\Phi}}_i,\ i \in \mathcal{I}. \tag{C.17}$$

Now what remains to show is that any primal-dual pair of the limit point $(\bar{\mathbf{x}}, \{\bar{\mathbf{\Phi}}_i\})$ a KKT stationary point of (5.20). Let us replace $n$ with $n_j$ in (C.1) and let $j \to \infty$. By noting that $\mathbf{x}^{(n_j)}$ and $\mathbf{x}^{(n_j+1)}$ are asymptotically close as indicated by Claim 2, we obtain

$$\mathbf{0} \in \nabla f_0(\bar{\mathbf{x}}) - \nabla g_0(\bar{\mathbf{x}}) + \sum_{i=1}^{I} \left(\bar{\mathbf{\Phi}}_i * (\boldsymbol{\mathcal{D}}\boldsymbol{\mathcal{F}}_i(\bar{\mathbf{x}}) - \boldsymbol{\mathcal{D}}\boldsymbol{\mathcal{G}}_i(\bar{\mathbf{x}}))\right) + \mathcal{N}(\Omega, \bar{\mathbf{x}}) \tag{C.18a}$$

$$\boldsymbol{\mathcal{F}}_i(\bar{\mathbf{x}}) - \boldsymbol{\mathcal{G}}_i(\bar{\mathbf{x}}) \preceq \mathbf{0},\ \bar{\mathbf{\Phi}}_i \succeq \mathbf{0},\ i \in \mathcal{I} \tag{C.18b}$$

$$\mathrm{Tr}\left((\boldsymbol{\mathcal{F}}_i(\bar{\mathbf{x}}) - \boldsymbol{\mathcal{G}}_i(\bar{\mathbf{x}}))\,\bar{\mathbf{\Phi}}_i\right) = 0,\ i \in \mathcal{I} \tag{C.18c}$$

$$\bar{\mathbf{x}} \in \Omega \tag{C.18d}$$

which is exactly the KKT conditions of the PSD DC problem (5.20). Noting the boundness of $\{\mathbf{x}^{(n)}\}$ assumed in A.2), it readily follows that there exists at least one limit point of $\{\mathbf{x}^{(n)}\}$ and by (C.18), any limit point of $\left\{\mathbf{x}^{(n)}\right\}$ is a KKT stationary point of (5.20).

## C.3 Proof of Claim 1

We argue by contradiction. Assume the contrary, i.e., $\lim_{n\to\infty}\tau^{(n)} = +\infty$. From the updating rule (5.30), it follows, without loss of generality, that there exists infinitely many

indices $j$ such that

$$\tau^{(n_j)} < \lambda_{\max}\Big[\sum_{i=1}^{I} \mathbf{\Phi}_i^{(n_j+1)}\Big] + \delta_1, \tag{C.19}$$

and

$$\tau^{(n_j)} < \|\mathbf{x}^{(n_j+1)} - \mathbf{x}^{(n_j)}\|^{-1}. \tag{C.20}$$

By possibly restricting to a subsequence of $\{n_j\}$, without loss of generality, we can further assume that there exists at least some $i \in \mathcal{S}_{\mathcal{I}}$, where $\mathcal{S}_{\mathcal{I}}$ denotes a subset of $\mathcal{I}$, i.e., $\mathcal{S}_{\mathcal{I}} \subseteq \mathcal{I}$ such that

$$\lim_{j\to\infty} \lambda_{\max}[\mathbf{\Phi}_i^{(n_j+1)}] = +\infty \Leftrightarrow \lim_{j\to\infty} \|\mathbf{\Phi}_i^{(n_j+1)}\|_F = +\infty,\ i \in \mathcal{S}_{\mathcal{I}} \tag{C.21}$$

and

$$\lim_{j\to\infty} \|\mathbf{x}^{(n_j+1)} - \mathbf{x}^{(n_j)}\| = 0. \tag{C.22}$$

Let $\lim_{j\to\infty} \mathbf{x}^{(n_j)} = \bar{\mathbf{x}}$, and then we will show that

$$\boldsymbol{\mathcal{F}}_i(\bar{\mathbf{x}}) - \boldsymbol{\mathcal{G}}_i(\bar{\mathbf{x}}) \not\prec \mathbf{0}, i \in \mathcal{S}_{\mathcal{I}}. \tag{C.23}$$

Again we show by contradiction. If we assume that $\boldsymbol{\mathcal{F}}_i(\bar{\mathbf{x}}) - \boldsymbol{\mathcal{G}}_i(\bar{\mathbf{x}}) \prec \mathbf{0}$, then we must have, for sufficiently large $j$,

$$\boldsymbol{\mathcal{F}}_i(\mathbf{x}^{(n_j+1)}) - \boldsymbol{\mathcal{G}}_i(\mathbf{x}^{(n_j)}) - \boldsymbol{\mathcal{D}}\boldsymbol{\mathcal{G}}(\mathbf{x}^{(n_j)})(\mathbf{x}^{(n_j+1)} - \mathbf{x}^{(n_j)}) - \mathbf{S}_i^{(n_j+1)} \prec \mathbf{0}. \tag{C.24}$$

This is due to (C.1b) and subsequently

$$\lim_{j\to\infty} \Big(\boldsymbol{\mathcal{F}}_i(\mathbf{x}^{(n_j+1)}) - \boldsymbol{\mathcal{G}}_i(\mathbf{x}^{(n_j)}) - \boldsymbol{\mathcal{D}}\boldsymbol{\mathcal{G}}(\mathbf{x}^{(n_j)})(\mathbf{x}^{(n_j+1)} - \mathbf{x}^{(n_j)}) - \mathbf{S}_i^{(n_j+1)}\Big) = \boldsymbol{\mathcal{F}}_i(\bar{\mathbf{x}}) - \boldsymbol{\mathcal{G}}_i(\bar{\mathbf{x}}). \tag{C.25}$$

By the complementary slackness condition (C.1d), it readily follows that when $j$ becomes sufficiently large, $\mathbf{\Phi}_i^{(n_j+1)} = \mathbf{0}$ for $i \in \mathcal{S}_{\mathcal{I}}$, which contradicts the previous result of (C.19).

Therefore, we must have $\boldsymbol{\mathcal{F}}_i(\bar{\mathbf{x}}) - \boldsymbol{\mathcal{G}}_i(\bar{\mathbf{x}}) \not\succ \mathbf{0}$ for $i \in \mathcal{S}_{\mathcal{I}}$.

Now let us assume, without loss of generality, that

$$\lim_{j\to\infty} \frac{\boldsymbol{\Phi}_i^{(n_j+1)}}{\sum_{i=1}^{I} \|\boldsymbol{\Phi}_i^{(n_j+1)}\|_F} = \hat{\boldsymbol{\Phi}}_i \succeq \mathbf{0}. \tag{C.26}$$

and it is easy to observe that $\hat{\boldsymbol{\Phi}}_i = \mathbf{0}$ for $i \in \mathcal{I}\backslash\mathcal{S}_{\mathcal{I}}$ and $\hat{\boldsymbol{\Phi}}_i \neq \mathbf{0}$ for $i \in \mathcal{S}_{\mathcal{I}}$. We now replace $n$ with $n_j$ in (C.1a). Dividing the both sides of (C.1a) by $\sum_{i=1}^{I} \|\boldsymbol{\Phi}_i^{(n_j+1)}\|_F$, taking the limit as $j \to \infty$ and using the result of (C.24), we obtain

$$\mathbf{0} \in \sum_{i\in\mathcal{S}_{\mathcal{I}}} (\boldsymbol{\mathcal{DF}}_i(\bar{\mathbf{x}}) - \boldsymbol{\mathcal{DG}}_i(\bar{\mathbf{x}})) * \hat{\boldsymbol{\Phi}}_i + \mathcal{N}(\Omega, \bar{\mathbf{x}}). \tag{C.27}$$

Multiplying the both sides of the above by $(\mathbf{y} - \mathbf{x}), \mathbf{y} \in \Omega$ yields

$$\sum_{i\in\mathcal{S}_{\mathcal{I}}} \mathrm{Tr}\left(\hat{\boldsymbol{\Phi}}_i \left(\boldsymbol{\mathcal{DF}}_i(\bar{\mathbf{x}}) - \boldsymbol{\mathcal{DG}}_i(\bar{\mathbf{x}})\right)(\mathbf{y} - \bar{\mathbf{x}})\right) \geq 0. \tag{C.28}$$

However, the MFCQ in Assumption 5.1 indicates that there exists some feasible direction $\mathbf{h} \in \mathtt{cone}(\Omega - \bar{\mathbf{x}})$ such that

$$(\boldsymbol{\mathcal{DF}}_i(\bar{\mathbf{x}}) - \boldsymbol{\mathcal{DG}}_i(\bar{\mathbf{x}}))\,\mathbf{h} \prec \mathbf{0},\ \forall i \in \mathcal{U}(\bar{\mathbf{x}}), \tag{C.29}$$

where we recall that $\mathcal{U}(\bar{\mathbf{x}})$ is the set of *active* constraints at $\bar{\mathbf{x}}$:

$$\mathcal{U}(\bar{\mathbf{x}}) \triangleq \left\{ i \in \mathcal{I} \middle| \boldsymbol{\mathcal{F}}_i(\bar{\mathbf{x}}) - \boldsymbol{\mathcal{G}}_i(\bar{\mathbf{x}}) \not\succ \mathbf{0} \right\}. \tag{C.30}$$

Considering $\hat{\boldsymbol{\Phi}}_i \succeq \mathbf{0}$, it is obvious (C.28) contradicts the MFCQ in Assumption 5.1. Therefore, we can assume that there exists an finite index $\tilde{n}$ such that

$$\tau^{(n)} = \tau^{(\tilde{n})},\ \forall n \geq \tilde{n}. \tag{C.31}$$

## C.4 Proof of Claim 2

Following directly from Lemma 5.2, we have

$$\hat{\varphi}^{(n)}(\mathbf{x}^{(n)}, \mathbf{S}^{(n)}) - \hat{\varphi}^{(n)}(\mathbf{x}^{(n+1)}, \mathbf{S}^{(n+1)}) \geq \frac{\rho_f + \rho_g}{2} \|\mathbf{x}^{(n+1)} - \mathbf{x}^{(n)}\|^2. \tag{C.32}$$

Then we evaluate the summation of (C.32) over $n$ from 0 to $\bar{n}$ and we obtain

$$\begin{aligned}
&\sum_{n=0}^{\bar{n}} \frac{\rho_f + \rho_g}{2} \|\mathbf{x}^{(n+1)} - \mathbf{x}^{(n)}\|^2 \\
&\quad \leq \sum_{n=0}^{\tilde{n}} \left( \hat{\varphi}^{(n)}(\mathbf{x}^{(n)}, \mathbf{S}^{(n)}) - \hat{\varphi}^{(n)}(\mathbf{x}^{(n+1)}, \mathbf{S}^{(n+1)}) \right) + \hat{\varphi}^{(\tilde{n})}(\mathbf{x}^{(\tilde{n})}, \mathbf{S}^{(\tilde{n})}) - \hat{\varphi}^{(\bar{n})}(\mathbf{x}^{(\bar{n}+1)}, \mathbf{S}^{(\bar{n}+1)})
\end{aligned} \tag{C.33}$$

where $\tilde{n}$ is a finite index as defined in (C.31), such that $\tau^{(n)} = \tau^{(\tilde{n})},\ \forall n \geq \tilde{n}$, i.e., the value of $\tau$ remains constant for all indices $n \geq \tilde{n}$. By Assumption 5.2, $\varphi(\mathbf{x})$ is bounded from below and hence $\hat{\varphi}^{(n)}(\cdot)$ is also lower-bounded. Taking the limit as $\bar{n} \to \infty$ on both sides of (C.33), we obtain

$$\begin{aligned}
&\sum_{n=0}^{\infty} \frac{\rho_f + \rho_g}{2} \|\mathbf{x}^{(n+1)} - \mathbf{x}^{(n)}\|^2 \\
&\quad \leq \sum_{n=0}^{\tilde{n}} \left( \hat{\varphi}^{(n)}(\mathbf{x}^{(n)}, \mathbf{S}^{(n)}) - \hat{\varphi}^{(n)}(\mathbf{x}^{(n+1)}, \mathbf{S}^{(n+1)}) \right) + \hat{\varphi}^{(\tilde{n})}(\mathbf{x}^{(\tilde{n})}, \mathbf{S}^{(\tilde{n})}) - \hat{\varphi}^{(\infty)}(\bar{\mathbf{x}}, \bar{\mathbf{S}}).
\end{aligned} \tag{C.34}$$

The right hand side of the above inequality must be of finite value. Therefore, it must hold that

$$\lim_{n \to \infty} \|\mathbf{x}^{(n+1)} - \mathbf{x}^{(n)}\| = 0. \tag{C.35}$$

## C.5 Proof of Proposition 5.2

Firstly, expanding all the quadratic terms of $\mathbf{W}$ in (5.62a) and (5.62b), and invoking the identities

$$\mathrm{Tr}\left(\mathbf{A}^H\mathbf{B}\mathbf{C}\mathbf{D}^H\right) = \mathtt{vec}\left(\mathbf{A}\right)^H\left(\mathbf{D}^T\otimes\mathbf{B}\right)\mathtt{vec}\left(\mathbf{C}\right) \tag{C.36}$$

$$\mathrm{Tr}\left(\mathbf{A}^H\mathbf{B}\mathbf{A}\right) = \mathtt{vec}\left(\mathbf{A}\right)^H\left(\mathbf{I}\otimes\mathbf{B}\right)\mathtt{vec}\left(\mathbf{A}\right), \tag{C.37}$$

(5.62a) and (5.62b) can be recast as (5.67a) and (5.67b), respectively, with $\mathbf{X} = \mathbf{w}\mathbf{w}^H$.

Next, we transform (5.62c). Recall that $\boldsymbol{\Theta}(\mathbf{W},\boldsymbol{\Psi}) = \sigma_{\mathrm{S}}^2\mathbf{W}\mathbf{h}_1\mathbf{h}_1^H\mathbf{W}^H - \gamma\sigma_{\mathrm{R}}^2\mathbf{W}\mathbf{W}^H - \gamma\boldsymbol{\Psi}$, and its first term on the right hand side is equivalent to

$$\sigma_{\mathrm{S}}^2\mathbf{W}\mathbf{h}_1\mathbf{h}_1^H\mathbf{W}^H = \sigma_{\mathrm{S}}^2(\mathbf{h}_1^T\otimes\mathbf{I}_{N_{\mathrm{R}}})\mathbf{w}\mathbf{w}^H(\mathbf{h}_1^T\otimes\mathbf{I}_{N_{\mathrm{R}}})^H, \tag{C.38}$$

by using $\mathtt{vec}(\mathbf{ABC}) = (\mathbf{C}^T\otimes\mathbf{A})\,\mathtt{vec}(\mathbf{B})$. To transform the second term, we express $\mathbf{W} = [\boldsymbol{w}_1,\cdots,\boldsymbol{w}_l,\cdots,\boldsymbol{w}_{N_{\mathrm{R}}}]$, where $\boldsymbol{w}_l$ denotes the $l^{\mathrm{th}}$ column of $\mathbf{W}$. Then $\mathbf{W}\mathbf{W}^H$ can be equivalently expressed as

$$\mathbf{W}\mathbf{W}^H = \sum_{l=1}^{N_{\mathrm{R}}}\boldsymbol{w}_l\boldsymbol{w}_l^H \tag{C.39}$$

By establishing the connection between $\boldsymbol{w}_l$ and $\mathbf{w}$ by $\boldsymbol{w}_l = \mathbf{E}_l\mathbf{w}$, (C.39) can further be written as

$$\mathbf{W}\mathbf{W}^H = \sum_{l=1}^{N_{\mathrm{R}}}\mathbf{E}_l\mathbf{w}\mathbf{w}^H\mathbf{E}_l^H. \tag{C.40}$$

Using (C.38) and (C.40), $\boldsymbol{\Theta}(\mathbf{W},\boldsymbol{\Psi})$ is equivalent to

$$\boldsymbol{\Theta}(\mathbf{W},\boldsymbol{\Psi}) = \sigma_{\mathrm{S}}^2\mathbf{H}_1\mathbf{w}\mathbf{w}^H\mathbf{H}_1^H - \gamma\sigma_{\mathrm{R}}^2\sum_{l=1}^{N_{\mathrm{R}}}\mathbf{E}_l\mathbf{w}\mathbf{w}^H\mathbf{E}_l^H - \gamma\boldsymbol{\Psi}. \tag{C.41}$$

Invoking $\mathbf{X} = \mathbf{w}\mathbf{w}^H$ and $\mathrm{Rank}(\mathbf{X}) = 1$, (5.62) is readily re-expressed as (5.67).

## C.6 Proof of Proposition 5.3

We prove the rank-one optimality of the solution to (5.69) by examining its KKT conditions. Let $y_1$, $y_2$ and $\mathbf{Y}_k$ denote the Lagrange multipliers associated with (5.69b)–(5.69d),

respective, and let $\mathbf{Z}_1$ and $\mathbf{Z}_2$ denote the Lagrange multipliers associated with $\mathbf{X} \succeq \mathbf{0}$ and $\mathbf{\Psi} \succeq \mathbf{0}$, respectively. The Lagrangian function of (5.69) can then be written as

$$\begin{aligned}\mathcal{L} = \operatorname{Tr}(\mathbf{Q}_2\mathbf{X}) + y_1(\operatorname{Tr}(\mathbf{Q}_2\mathbf{X}) + \operatorname{Tr}(\mathbf{\Psi})) - y_2\left(\operatorname{Tr}(\mathbf{Q}_0\mathbf{X}) - \tau^*\operatorname{Tr}(\mathbf{Q}_1\mathbf{X}) - \tau^*\operatorname{Tr}(\mathbf{h}_2\mathbf{h}_2^H\mathbf{\Psi})\right) \\ + \sum_{k=1}^{K}\operatorname{Tr}\left(\mathbf{P}_k^H\mathbf{Q}_3(\mathbf{X},\mathbf{\Psi})\mathbf{P}_k\mathbf{Y}_k\right) - \operatorname{Tr}(\mathbf{X}\mathbf{Z}_1) - \operatorname{Tr}(\mathbf{\Psi}\mathbf{Z}_2),\end{aligned} \tag{C.42}$$

where we have neglected the terms, which are independent of $\mathbf{X}$ and $\mathbf{\Psi}$. Now we exploit the first-order KKT conditions with respect to $\mathbf{X}$ and $\mathbf{\Psi}$, which can be given by

$$\begin{aligned}\frac{\partial\mathcal{L}}{\partial\mathbf{X}} &= \mathbf{Q}_2 + y_1\mathbf{Q}_2 - y_2\mathbf{Q}_0 + y_2\tau^*\mathbf{Q}_1 \\ &\quad + \sum_{k=1}^{K}\sigma_{\mathtt{S}}^2\mathbf{H}_1^H\mathbf{P}_k\mathbf{Y}_k\mathbf{P}_k^H\mathbf{H}_1 \\ &\quad - \gamma\sigma_{\mathtt{R}}^2\sum_{k=1}^{K}\sum_{l=1}^{N_{\mathtt{R}}}\mathbf{E}_l^H\mathbf{P}_k\mathbf{Y}_k\mathbf{P}_k^H\mathbf{E}_l - \mathbf{Z}_1 = \mathbf{0}\end{aligned} \tag{C.43}$$

$$\frac{\partial\mathcal{L}}{\partial\mathbf{\Psi}} = y_1\mathbf{I} + y_2\tau^*\mathbf{h}_2\mathbf{h}_2^H - \sum_{k=1}^{K}\gamma\mathbf{P}_k\mathbf{Y}_k\mathbf{P}_k^H - \mathbf{Z}_2 = \mathbf{0} \tag{C.44}$$

Re-arranging (C.44) and using the associativity of the Kronecker product, we obtain

$$\mathbf{I}\otimes\mathbf{Z}_2 = y_1\mathbf{I} + y_2\tau^*\mathbf{I}\otimes(\mathbf{h}_2\mathbf{h}_2^H) - \mathbf{I}\otimes\sum_{k=1}^{K}\gamma\mathbf{P}_k\mathbf{Y}_k\mathbf{P}_k^H. \tag{C.45}$$

A simple calculation reveals that the right hand side of (C.45) is equivalent to

$$\mathbf{I}\otimes\mathbf{Z}_2 = y_1\mathbf{I} + y_2\tau^*\mathbf{Q}_1/\sigma_{\mathtt{R}}^2 - \gamma\sum_{k=1}^{K}\sum_{l=1}^{N_{\mathtt{R}}}\mathbf{E}_l^H\mathbf{P}_k\mathbf{Y}_k\mathbf{P}_k^H\mathbf{E}_l. \tag{C.46}$$

Substituting the above relation into (C.43), we further obtain

$$\underbrace{\mathbf{Q}_2 + y_1\mathbf{Q}_2 + \sigma_{\mathtt{R}}^2\mathbf{Z}_2 + \sum_{k=1}^{K}\sigma_{\mathtt{S}}^2\mathbf{H}_1^H\mathbf{P}_k\mathbf{Y}_k\mathbf{P}_k^H\mathbf{H}_1}_{\triangleq\mathbf{\Theta}} - y_2\mathbf{Q}_0 = \mathbf{Z}_1. \tag{C.47}$$

Since $\mathbf{Q}_2 \succ \mathbf{0}$, $\boldsymbol{\Theta}$ must be a positive definite matrix, which has full rank, i.e., $\mathrm{Rank}(\boldsymbol{\Theta}) = N_{\mathtt{R}}^2$. It is further implied by (C.47) that

$$\mathrm{Rank}(\mathbf{Z}_1) \geq \mathrm{Rank}(\boldsymbol{\Theta}) - \mathrm{Rank}(\mathbf{Q}_0), \tag{C.48}$$

where $\mathrm{Rank}(\mathbf{Q}_0) = 1$. Then it is clear that the rank of $\mathbf{Z}_1$ is either $N_{\mathtt{R}}^2$ or $N_{\mathtt{R}}^2 - 1$. If $\mathrm{Rank}(\mathbf{Z}_1) = N_{\mathtt{R}}^2$, we must have $\mathbf{X} = \mathbf{0}$ due to the complementary slackness condition $\mathrm{Tr}(\mathbf{X}\mathbf{Z}_1) = 0$. However, it is obvious that $\mathbf{X} = \mathbf{0}$ is not the optimal solution. Then the rank of $\mathbf{Z}_1$ must be $N_{\mathtt{R}}^2 - 1$ and in this case, $\mathbf{X}$ must lie in the nullspace of $\mathbf{Z}_1$, whose dimension is one. Therefore, $\mathbf{X}$ must be of rank one.

# References


[1] S. W. Peters, A. Y. Panah, K. T. Truong, and R. W. Heath, "Relay architecures for 3GPP LTE-advanced," *EURASIP J. Wireless Commun. Network.*, pp. 1–14, 2009:618787.

[2] 3GPP TR 36.843 V 12.0.1, "3rd generation partnership project; technical specification group radio access network; study on LTE device to device proximity services; radio aspects (Release 12)," Mar. 2014.

[3] T. Cover and A. E. Gamal, "Capacity theorems for the relay channel," *IEEE Trans. Inf. Theory*, vol. IT-25, pp. 572–584, Sep. 1979.

[4] J. N. Laneman, D. N. Tse, and G. W. Wornell, "Cooperative diversity in wireless networks: Efficient protocols and outage behavior," *IEEE Trans. Inf. Theory*, vol. 50, pp. 3062–3080, Dec. 2004.

[5] D. Kim, H. Lee, and D. Hong, "A survey of in-band full-duplex transmission: From the perspective of PHY and MAC layers," *IEEE Commun. Surveys Tuts.*, vol. 17, pp. 2017–2046, Fourthquater 2015.

[6] L. Sanguinetti, A. A. D'Amico, and Y. Rong, "A tutorial on the optmization of amplify-and-forward MIMO relay systems," *IEEE J. Sel. Areas Commun.*, vol. 30, pp. 1331–1346, Sep. 2012.

[7] 3GPP TS 36.106, "Evolved universal terrestrial radio access (E-UTRA); FDD repeater radio transmission and reception (Release 8)," Jan. 2011.

[8] 3GPP TS 36.216, "Evolved universal terrestrial radio access (E-UTRA); Physical layer for relaying operation (Release 10)," Sep. 2011.

[9] M. Sawahashi, Y. Kishiyama, A. Morimoto, D. Nishikawa, and M. Tanno, "Coordinated multipoint transmission/reception techniques for LTE-advanced," *IEEE Wireless Commun.*, vol. 17, pp. 26–34, Jun. 2010.

[10] G. P. Fettweis and E. Zimmermann, "ICT energy consumption - trends and challenges," in *Proc. 11th Int. Symp. Wireless Pers. Multimedia Commun.*, (Lapland, Finland), Sep. 2008.

[11] "Mobile cyber threats: Kaspersky lab & INTERPOL joint report," tech. rep., INTERPOL and Kaspersky Lab, Oct. 2014.

[12] C. Timberg, "German researchers discover a flaw that could let anyone listen to your cell calls," *Washington Post: The Switch*, Dec. 2014. Available online: https://www.washingtonpost.com/news/the-switch/wp/2014/12/18/german-researchers-discover-a-flaw-that-could-let-anyone-listen-to-your-cell-calls-and-read-your-texts/.

[13] A. D. Wyner, "The wire-tap channel," *Bell Syst. Tech. J.*, vol. 54, pp. 1355–1387, Jan. 1975.

[14] A. Mukherjee, S. A. A. Fakoorian, J. Huang, and A. L. Swindlehurst, "Principles of physical layer security in multiuser wireless networks: A survey," *IEEE Commun. Surveys Tuts.*, vol. 16, pp. 1550–1573, Third Quarter 2014.

[15] L. Lai and H. E. Gamal, "The relay eavesdropper channel: Cooperation for secrecy," *IEEE Trans. Inf. Theory*, vol. 54, pp. 4005–4019, Sep. 2008.

[16] Y. Rong, "Optimal linear non-regnerative multi-hop MIMO relays with MMSE-DFE receiver at the destination," *IEEE Trans. Wireless Commun.*, vol. 9, pp. 2268–2279, Jul. 2010.

[17] Y. Rong, X. Tang, and Y. Hua, "A unified framework for optimizing linear nonregenerative multicarrier MIMO relay communication systems," *IEEE Trans. Signal Process.*, vol. 57, pp. 4837–4851, Aug. 2009.

[18] D. P. Palomar, J. M. Cioffi, and M. A. Lagunas, "Joint Tx-Rx beamforming design for multicarrier MIMO channels: A unified framework for convex optimization," *IEEE Trans. Signal Process.*, vol. 51, pp. 2381–2401, Sep. 2003.

[19] A. Wiesel, Y. C. Eldar, and S. Shamai, "Linear precoding via conic optimization for fixed MIMO receivers," *IEEE Trans. Signal Process.*, vol. 54, pp. 161–176, Jan. 2005.

[20] R. H. Y. Louie, Y. Li, and B. Vucetic, "Zero forcing in general two-hop relay networks," *IEEE Trans. Veh. Technol.*, vol. 59, pp. 191–202, Jan. 2010.

[21] C. Song, K.-J. Lee, and I. Lee, "MMSE based transceiver designs in closed-loop non-regenerative MIMO relaying systems," *IEEE Trans. Wireless Commun.*, vol. 9, pp. 2310–2319, Jul. 2010.

[22] W. Zhang, U. Mitra, and M. Chiang, "Optimization of amplify-and-forward multicarrier two-hop transmission," *IEEE Trans. Wireless Commun.*, vol. 59, pp. 1434–1445, May 2011.

[23] R. Cendrillon, W. Yu, M. Moonen, J. Verlinden, and T. Bostoen, "Optimal multiuser spectrum balancing for digital subscriber lines," *IEEE Trans. Commun.*, vol. 54, pp. 922–933, May 2006.

[24] C. Xing, S. Ma, and Y.-C. Wu, "Robust joint design of linear relay precoder and destination equalizer for dual-hop amplify-and-forward MIMO relay systems," *IEEE Trans. Signal Process.*, vol. 58, pp. 2273–2283, Apr. 2010.

[25] M. R. A. Khandaker and Y. Rong, "Interference MIMO relay channel: Joint power control and transceiver-relay beamforming," *IEEE Trans. Signal Process.*, vol. 60, pp. 6509–6518, Dec. 2012.

[26] Y. Rong, "Joint source and relay optimization for two-way MIMO multi-relay networks," *IEEE Commun. Lett.*, vol. 15, pp. 1329–1331, Dec. 2011.

[27] L. Gopal, Y. Rong, and Z. Zang, "Robust MMSE transceiver design for nonregenerative multicasting MIMO relay systems," *IEEE Trans. Veh. Technol.*, vol. 66, pp. 8979–8989, Oct. 2017.

[28] K. T. Truong, P. Sartori, and R. W. Heath, "Cooperative algorithms for MIMO amplify-and-forward relay networks," *IEEE Trans. Signal Process.*, vol. 61, pp. 1272–1287, Mar. 2013.

[29] M. Zhang, H. Yi, H. Yu, H. Luo, and W. Chen, "Joint optimization in bidirectional multi-user multi-relay MIMO systems: Non-robust and robust cases," *IEEE Trans. Veh. Technol.*, preprint, Mar. 2013. doi:10.1109/TVT.2013.2255898.

[30] J. Yang, B. Champagne, Y. Zou, and L. Hanzo, "Joint optimization of transceiver matrices for MIMO-aided multiuser AF relay networks: Improving the QoS in the presence of CSI errors," *IEEE Trans. Veh. Technol.*, vol. 65, pp. 1434–1451, Mar. 2016.

[31] A. Beck and L. Tetruashvili, "On the convergence of block coordinate descent type methods," *SIAM J. Control Optim.*, vol. 23, no. 4, pp. 2037–2060, 2013.

[32] P. Tseng, "Convergence of a block coordinate descent method for nondifferentiable minimization," *J. Optim. Theory Appl.*, vol. 109, pp. 475–494, Jun. 2001.

[33] Z.-Q. Luo and P. Tseng, "On the convergence of the coordinate descent method for convex differentiable minimization," *J. Optim. Theory Appl.*, vol. 72, pp. 7–35, Jan. 1992.

[34] M. Razaviyayn, M. Hong, and Z.-Q. Luo, "A unified convergence analysis of block successive minimization methods for nonsmooth opimization," *SIAM J. Optim.*, vol. 23, no. 2, pp. 1126–1153, 2013.

[35] Z. Qin, K. Scheinberg, and D. Glodfarb, "Efficient block-coordinate descent algorithms for the group lasso," *Math. Prog. Comp.*, no. 5, pp. 143–169, 2013.

[36] M. R. K. A. Toding and Y. Rong, "Joint source and relay optimization for parallel MIMO relay networks," *EURASIP J. Adv. Signal Process.*, pp. 1–7, 2012:174.

[37] W.-P. Z. Y. Fu, L. Yang and C. Liu, "Optimum linear design of two-hop MIMO relay networks with Qos requirements," *IEEE Trans. Signal Process.*, vol. 59, pp. 2257–2269, May 2011.

[38] O. Oyman and A. J. Paulraj, "Design and analysis of linear distributed MIMO relaying algorithms," *IEE Proc. Commun.*, vol. 153, pp. 565–572, Apr. 2006.

[39] C. Zhao and B. Champagne, "Joint design of multiple non-regenerative MIMO relaying matrices with power constraints," *IEEE Trans. Signal Process.*, vol. 61, pp. 4861–4873, Oct. 2013.

[40] H. Shi, T. Abe, T. Asai, and H. Yoshino, "Relaying schemes using matrix triangularization for MIMO wireless networks," *IEEE Trans. Commun.*, vol. 55, pp. 1683–1688, Sep. 2007.

[41] H. Shi, T. Abe, T. Asai, and H. Yoshino, "A relaying scheme using QR decomposition with phase control for MIMO wireless networks," in *Proc. IEEE Int. Conf. Commun. (ICC), 2005*, (Seoul, Korea), pp. 2705–2711, May 2005.

[42] K.-J. Lee, H. Sung, E. Park, and I. Lee, "Joint optimization for one and two-way MIMO AF multiple-relay systems," *IEEE Trans. Wireless Commun.*, vol. 9, pp. 3671–3681, Dec. 2010.

[43] A. Toding, M. R. Khandaker, and Y. Rong, "Joint source and relay design for MIMO multi-relay systems using projected gradient approach," *EURASIP J. Wireless Commun. Network.*, pp. 1–9, 2014:151.

[44] C. Zhao and B. Champagne, "A unified approach to optimal transceiver design for nonregenerative MIMO relaying," *IEEE Trans. Veh. Technol.*, vol. 64, pp. 2938–2951, Jul.. 2015.

[45] S. Berger, M. Kuhn, A. Wittneben, T. Unger, and A. Klein, "Recent advances in amplify-and-forward two-hop relaying," *IEEE Commun. Mag.*, vol. 47, pp. 50–56, Jul. 2009.

[46] Y. Liu and A. P. Petropulu, "On the sumrate of amplify-and-forward relay networks with multiple source-destination pairs," *IEEE Trans. Wireless Commun.*, vol. 10, pp. 3732–3742, Nov. 2011.

[47] A. El-Keyi and B. Champagne, "Adaptive linearly constrained minimum variance beamforming for multiuser cooperative relaying using the Kalman filter," *IEEE Trans. Wireless Commun.*, vol. 9, pp. 641–651, Feb. 2010.

[48] P. Ubaidulla and A. Chockalingam, "Relay precoder optimization in MIMO-relay networks with imperfect CSI," *IEEE Trans. Signal Process.*, vol. 59, pp. 5473–5484, Nov. 2011.

[49] B. K. Chalise and L. Vandendorpe, "MIMO relay design for multipoint-to-multipoint communications with imperfect channel state information," *IEEE Trans. Signal Process.*, vol. 57, pp. 2785–2796, Jul. 2009.

[50] B. K. Chalise and L. Vandendorpe, "Optimization of MIMO relays for multipoint-to-multipoint communications: Nonrobust and robust designs," *IEEE Trans. Signal Process.*, vol. 58, pp. 6355–6368, Dec. 2010.

[51] M. Fadel, A. El-Keyi, and A. Sultan, "QOS-constrained multisuer peer-to-peer amplify-and-forward relay beamforming," *IEEE Trans. Signal Process.*, vol. 60, pp. 1397–1408, Mar. 2012.

[52] W. Liu, L.-L. Yang, and L. Hanzo, "SVD-assisted multiuser transmitter and multiuser detector design for MIMO systems," *IEEE Trans. Veh. Technol.*, vol. 58, pp. 1016–1021, Feb. 2009.

[53] D. P. Bertsekas, *Nonlinear Programming*. Singapore: Athena Scientific, second ed., 1999.

[54] Z.-Q. Luo, W.-K. Ma, A. M.-C. So, Y. Ye, and S. Zhang, "Semidefinite relaxation of quadratic optimization problems," *IEEE Signal Process. Mag.*, vol. 27, pp. 20–34, May 2010.

[55] K. X. Nguyen, Y. Rong, and S. Nordholm, "MMSE-based transceiver design algorithms for interference MIMO relay systems," *IEEE Trans. Wireless Commun.*, vol. 14, pp. 6414–6424, Nov. 2015.

[56] C. Xing, S. Li, Z. Fei, and J. Kuang, "How to understand linear minimum mean-square-error transceiver design for multiple-input-multiple-output systems from quadratic matrix programming," *IET Commun.*, vol. 7, pp. 1231–1242, Aug. 2013.

[57] "C-RAN: The road towards green RAN," white paper, ver. 2.5, China Mobile Research Institute, Bejing, China, Oct. 2011.

[58] X. Zhang, D. P. Palomar, and B. Ottersten, "Statistically robust design of linear MIMO transecivers," *IEEE Trans. Signal Process.*, vol. 56, pp. 3678–3689, Aug. 2008.

[59] A. Pascual-Iserte, D. P. Palomar, A. I. Perez-Neira, and M. A. Lagunas, "A robust maximin approach for MIMO communications with imperfect channel state information based on convex optimization," *IEEE Trans. Signal Process.*, vol. 54, pp. 346–360, Jan. 2006.

[60] D. Shiu, G. J. Foschini, M. J. Gans, and J. M. Kahn, "Fading correlation and its effect on the capacity of multielement antenna systems," *IEEE Trans. Commun.*, vol. 48, pp. 502–513, Mar. 2000.

[61] L. Musavian, M. R. Nakhi, M. Dohler, and A. H. Aghvami, "Effect of channel uncertainty on the mutual information of MIMO fading channels," *IEEE Trans. Veh. Technol.*, vol. 56, pp. 2798–2806, Sep. 2007.

[62] M. Ding and S. D. Blostein, "MIMO minimum total MSE transceiver design with imperfect CSI at both ends," *IEEE Trans. Signal Process.*, vol. 57, pp. 1141–1150, Mar. 2009.

[63] A. D. Dabbagh and D. J. Love, "Multiple antenna MMSE based downlink precoding with quantized feedback or channel mismatch," *IEEE Trans. Wireless Commun.*, vol. 56, pp. 1859–1868, Nov. 2008.

[64] D. Zheng, J. Liu, K.-K. Wong, H. Chen, and L. Chen, "Robust peer-to-peer collaborative-relay beamfomring with ellipsoidal CSI uncertainties," *IEEE Signal Process. Lett.*, vol. 16, pp. 442–445, Apr. 2012.

[65] L. Sanguinetti, A. A. D'Amico, and Y. Rong, "On the design of amplify-and-forward MIMO-OFDM relay systems with QoS requirements specified as Schur-convex functions of the MSEs," *IEEE Trans. Veh. Technol.*, vol. 62, pp. 1871–1877, May 2013.

[66] Y. Cheng and M. Pesavento, "Joint optimization of source power allocation and distributed relay beamforming in multiuser peer-to-peer relay networks," *IEEE Trans. Signal Process.*, vol. 60, pp. 2962–2973, Jun. 2012.

[67] U. Rashid, H. D. Tuan, H. H. Kha, and H. H. Nguyen, "Joint optimization of source precoding and relay beamforming in wireless MIMO relay networks," *IEEE Trans. Commun.*, vol. 62, pp. 488–499, Feb. 2014.

[68] M. Peng, Y. Sun, X. Li, Z. Mao, and C. Wang, "Recent advances in cloud radio access networks: System architectures, key techniques, and open issues," *IEEE Commun. Surveys Tuts.*, vol. 18, pp. 2282–2308, Third Quarter 2016.

[69] Y. Shi, J. Zhang, and K. B. Letaief, "Group sparse beamforming for green cloud-RAN," *IEEE Trans. Wireless Commun.*, vol. 13, pp. 2809–2823, May 2014.

[70] Y. Shi, J. Zhang, and K. B. Letaief, "Robust group sparse beamforming for multicast green cloud-RAN with imperfect CSI," *IEEE Trans. Signal Process.*, vol. 63, pp. 4647–4659, Sep. 2015.

[71] S. Luo, R. Zhang, and T. J. Lim, "Downlink and uplink energy minimization through user association and beamforming in C-RAN," *IEEE Trans. Wireless Commun.*, vol. 14, pp. 494–508, Jan. 2015.

[72] B. Dai and W. Yu, "Sparse beamforming and user-centric clustering for downlink cloud radio access network," *IEEE Access*, vol. 2, pp. 1326–1339, Oct. 2014.

[73] Y. Shi, J. Zhang, K. B. Letaief, B. Bai, and W. Chen, "Large-scale convex optimization for ultra-dense cloud-RAN," *IEEE Wireless Commun.*, vol. 22, pp. 84–91, Jun. 2015.

[74] L. Dong, Z. Han, A. P. Petropulu, and H. V. Poor, "Improving wireless physical layer security via cooperating relays," *IEEE Trans. Signal Process.*, vol. 58, pp. 1875–1888, Mar. 2010.

[75] J. Li, A. P. Petropulu, and S. Weber, "On cooperative relaying schemes for wireless physical layer security," *IEEE Trans. Signal Process.*, vol. 59, pp. 4985–4997, Oct. 2011.

[76] Y. Yang, Q. Li, W.-K. Ma, J. Ge, and P. C. Ching, "Cooperative secure beamforming for AF relay networks with multiple eavesdroppers," *IEEE Signal Process. Lett.*, vol. 20, pp. 35–38, Jan. 2013.

[77] H.-M. Wang, F. Liu, and X.-G. Xia, "Joint source-relay precoding and power allocation for secure amplify-and-forward MIMO relay networks," *IEEE Trans. Inf, Forensics Security*, vol. 9, pp. 1240–1250, Aug. 2014.

[78] C. Jeong, I.-M. Kim, and D. I. Kim, "Joint secure beamforming design at the source and the relay for an amplify-and-forward MIMO untrusted relay system," *IEEE Trans. Signal Process.*, vol. 60, pp. 310–325, Jan. 2012.

[79] K. Jayasinghe, P. Jayasinghe, N. Rajatheva, and M. Latva-aho, "Secure beamforming design for physical layer network coding based MIMO two-way relaying," *IEEE Commun. Lett.*, vol. 18, pp. 1270–1273, Jul. 2014.

[80] J. Yang, B. Champagne, Y. Zou, and L. Hanzo, "MIMO AF relaying security: Robust transceiver design in the presence of multiple eavesdroppers," in *Proc. IEEE Int. Conf. Commun. (ICC), 2015*, (London, U.K.), pp. 4937–4942, Jun. 2015.

[81] J. Yang, B. Champagne, Q. Li, and L. Hanzo, "Secure MIMO AF relaying design: An intercept probability constrained approach," in *Proc. IEEE Global Commun. Conf. (Globecom), 2015*, (San Diego, CA), pp. 1–6, Dec. 2015.

[82] G. Zheng, L.-C. Choo, and K.-K. Wong, "Optimal cooperative jamming to enhance physical layer security using relays," *IEEE Trans. Signal Process.*, vol. 59, pp. 4985–4997, Oc. 2011.

[83] M. Lin, J. Ge, and Y. Yang, "An effective secure transmission scheme for AF relay networks with two-hop information leakage," *IEEE Commun. Lett.*, vol. 17, pp. 1676–1679, Aug. 2013.

[84] L. Wang, C. Cao, M. Song, and Y. Cheng, "Joint cooperative relaying and jamming for maximum secrecy capacity in wireless networks," in *Proc. 2014 IEEE Int. Conf. Commun. (ICC)*, (Sydney, Australia), pp. 4448–4453, Jun. 2014.

[85] Y. Liu, J. Li, and A. P. Petropulu, "Destination assisted cooperative jamming for wireless physical-layer security," *IEEE Trans. Inf. Forensics Security*, vol. 8, pp. 682–694, Apr. 2013.

[86] S. Vishwakarma and A. Chockalingam, "Amplify-and-forward relay beamforming for secrecy with cooperative jamming and imperfect CSI," in *Proc. IEEE Int. Conf. Commun. (ICC), 2013*, (Budapest, Hungry), pp. 3047–3052, Jun. 2013.

[87] Q. Li, Y. Yang, W.-K. Ma, M. Lin, J. Ge, and J. Lin, "Robust cooperative beamforming and artificial noise design for physical-layer secrecy in AF multi-antenna multi-relay networks," *IEEE Trans. Signal Process.*, vol. 63, pp. 206–220, Jan. 2015.

[88] C. Zhang, H. Gao, H. Liu, and T. Lv, "Robust beamforming and jamming for secure af relay networks with multiple eavesdroppers," in *Proc. IEEE Military Commun. Conf. (MILCOM), 2014*, (Baltimore, MD), pp. 495–500, Oct. 2014.

[89] X. Wang, K. Wang, and X.-D. Zhang, "Secure relay beamforming with imperfect channel side information," *IEEE Trans. Veh. Technol.*, vol. 62, pp. 2140–2155, Jun. 2013.

[90] M. Zhang, J. Huang, H. Yu, H. Luo, and W. Chen, "QoS-based source and relay secure optimization design with presence of channel uncertainty," *IEEE Commun. Lett.*, vol. 17, pp. 1544–1547, Aug. 2013.

[91] Z. Chu, K. Cumanan, M. Xu, and Z. Ding, "Robust secrecy rate optimisations for multiuser multiple-input-single-output channel with device-to-device communications," *IET Commun.*, vol. 9, pp. 396–403, Feb. 2015.

[92] J. Yang, Q. Li, H. Li, and B. Champagne, "Intercept probability-constrained secure MIMO AF relaying with arbitrarily distributed ECSI errors," in *Proc. 2016 IEEE Int. Sym. Pers. Indoor Mobile Radio Commun. (PIMRC)*, (Valencia, Spain), pp. 1–6, Sep. 2016.

[93] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge Univ. Press, 2004.

[94] L. Vandenberghe and S. Boyd, "Semidefinite programming," *SIAM Rev.*, vol. 38, pp. 49–95, Mar. 1996.

[95] M. S. Lobo, L. Vandenberghe, S. Boyd, and H. Lebret, "Applications of second-order cone programming," *Linear Algebra Appl.*, vol. 284, pp. 193–228, Nov. 1998.

[96] E. J. Candès, M. B. Wakin, and S. P. Boyd, "Enhancing sparsity by reweighted $l_1$ minimization," *J. Fourier Anal. Appl.*, pp. 877–905, Oct. 2008.

[97] J. Yang and B. Champagne, "Joint transceiver optimization for MIMO multiuser relaying networks with channel uncertainties," in *2014 IEEE 80th Veh. Technol. Conf., (VTC)*, (Vancouver, Canada), pp. 1–6, Sep. 2014.

[98] Y. Xu and W. Yin, "A block corrdinate descent method for regularized multiconvex optimization with applications to nonnegative tensor factorization and completion," *SIAM J. Imaging Sci.*, vol. 6, pp. 1758–1789, Sep. 2013.

[99] Y. C. Eldar, A. Ben-Tal, and A. Nemirovski, "Robust mean-squared error estimation in the presence of model uncertainties," *IEEE Trans. Signal Process.*, vol. 53, pp. 168–181, Jan. 2005.

[100] M. Grant and S. Boyd, "CVX: Matlab software for disciplined convex programming, version 2.0 beta." `http://cvxr.com/cvx`, Sep 2013.

[101] J. F. Sturm, "Using SeDuMi 1.02, a MATLAB toolbox for optimization over symmetric cones," *Optim. Methods Softw.*, vol. 11, pp. 625–653, Jan. 1999.

[102] MOSEK ApS, *The MOSEK optimization toolbox for MATLAB manual. V 7.1 (R 49)*, 2016.

[103] Y. Nesterov and A. Nemirovski, *Interior Point Polynomial Time Methods in Convex Programming: Theory and Applications*. Philadelphia, PA: SIAM, 1994.

[104] S. Boyd, L. El Ghaoui, E. Feron, and V. Balakrishnan, *Linear Matrix Inequalities in System and Control Theory*. Philadelphia, PA: SIAM, 1994.

[105] S. Liu, "Matrix results on the Khatri-Rao and Tracy-Singh products," *Elsevier Linear Algebra Appl.*, vol. 289, pp. 267–277, Mar. 1999.

[106] J. Yang, B. Champagne, Y. Zou, and L. Hanzo, "Centralized energy-efficient multiuser multiantenna relaying in next-generation radio access networks," *IEEE Trans. Veh. Technol.*, vol. 66, pp. 7913–7924, Sep. 2017.

[107] O. Arnold, F. Richter, G. Fettweis, and O. Blume, "Power consumption modeling of different base station types in heterogeneous cellular networks," in *in Proc. Future Network Mobile Summit, 2010*, (Florence, Italy), pp. 1–8, Jun. 2010.

[108] D. R. Hunter and K. Lange, "A tutorial on MM algorithms," *Amer. Statistician*, vol. 58, pp. 30–37, Feb. 2004.

[109] J. Yang, Y. Cai, B. Champagne, and L. Hanzo, "Multi-user beamforming-aided AF relaying: A low-complexity adaptive design approach," in *Asilomar Conf. Signals. Sys. Comput.*, (Pacific Grove, CA), pp. 983–988, Nov. 2015.

[110] H. Ning, C. Ling, and K. K. Leung, "Relay-aided interference alignment: Feasibility conditions and algorithm," in *Proc. IEEE Int. Sym. Inf. Theory*, (Austin, TX), pp. 390–394, Jun. 2010.

[111] B. Nourani, S. A. motahari, and A. K. Khandani, "Relay-aided interference alignment for the quasi-static interference channel," in *in Proc. IEEE Int. Sym. Inf. Theory, 2010*, (Austin, TX), pp. 405–409, Mar. 2010.

[112] K. P. Murphy, *Machine Learning: A Probabilistic Perspective*. Cambridge, MA: The MIT Press, 2012.

[113] J. Löfberg, "YALMIP Wiki: Solvers." [Online]. Available: http://users.isy.liu.se/johanl/yalmip/pmwiki.php?n=Solvers.Solvers.

[114] Q. Li, M. Hong, H.-T. Wai, Y.-F. Liu, W.-K. Ma, and Z.-Q. Luo, "Transmit solutions for MIMO wiretap channels using alternating optimization," *IEEE J. Sel. Areas Commun.*, vol. 31, pp. 1714–1727, Sep. 2013.

[115] Y. Shi, J. Cheng, J. Zhang, B. Bai, W. Chen, and K. B. Letaief, "Smoothed $l_p$-minimization for green cloud-RAN with user admission control." *IEEE J. Sel. Areas Commun.*, IEEEXplore, Early Access.

[116] J. Löfberg, "YALMIP: a toolbox for modeling and optimization in MATLAB," in *Proc. 2004 IEEE Int. Symp. Comput. Aided Control Syst. Design*, (Taipei, Taiwan), pp. 284–289, Sep. 2004.

[117] J. Yang, Q. Li, Y. Cai, Y. Zou, L. Hanzo, and B. Champagne, "Joint secure AF relaying and artificial noise optimization: A penalized difference-of-convex programming framework," *IEEE Access*, vol. 4, pp. 10076–10095, 2016.

[118] B. K. Sriperumbudur and G. R. G. Lanckriet, "On the convergence of the concave-convex procedure," *Advances Neural Inf. Process. Syst. 22*, pp. 1759–1767, 2009.

[119] A. Goldsmith, *Wireless Communications*. Cambridge, U.K.: Cambridge Univ. Press, first ed., Aug. 2005.

[120] R. Horst and N. V. Thoai, "DC programming: Overview," *J. Optim. Theory Appl.*, vol. 103, pp. 1–43, Oct. 1999.

[121] Q. T. Dinh, S. Gumussoy, W. Michiels, and M. Diehl, "Combining convex-concave decompositions and linearization approaches for solving BMIs, with application to static ouput feedback," *IEEE Trans. Autom. Control*, vol. 57, pp. 1377–1390, Jun. 2012.

[122] S.-P. Han and O. L. Mangasarian, "Exact penalty functions in nonlinear programming," *Math. Programming*, vol. 17, no. 1, pp. 251–269, 1979.

[123] G. Di Pillo and L. Grippo, "Exact penalty functions in constrained optimization," *SIAM J. Control Optim.*, vol. 26, pp. 1333–1360, Nov. 1989.

[124] O. L. Mangasarian and S. Fromovitz, "The Fritz John necessay optoptimal conditions in the presence of equality and inequality constraints," *J. Math. Anal. Appl.*, vol. 17, pp. 37–47, 1967.

[125] A. Charnes and W. W. Cooper, "Programming with linear fractional functionals," *Naval Res. Logist. Quart.*, vol. 9, pp. 181–186, Dec. 1962.

[126] Y. Huang and D. P. Palomar, "Randomized algorithms for optimal solutions of double-sided QCQP witha pplications in signal processing," *IEEE Trans. Signal Process.*, vol. 62, pp. 1093–1108, Mar. 2014.

[127] W. Ai, Y. Huang, and S. Zhang, "New results on Hermitian matrix rank-one decomposition," *Math. Programming*, vol. 128, pp. 253–283, Aug. 2009.

[128] N. J. T. Yoo and A. Goldsmith, "Multi-antenna downlink channels with limited feedback and user selection," *IEEE J. Sel. Areas Commun.*, vol. 25, pp. 1478–1491, Sep. 2007.

[129] N. Jindal, "MIMO broadcast channels with finite-rate feedback," *IEEE Trans. Inf. Theory*, vol. 52, pp. 5045–5060, Nov. 2006.

[130] W. X. C. Zhang and M. Chen, "Robust MMSE beamforming for multiuser MISO systems with limited feedback," *IEEE Signal Process. Lett.*, vol. 16, pp. 588–591, Jul. 2009.

[131] W. Wang, A. Harada, and H. Kayama, "Enhanced limited feedback schemes for DL MU-MIMO ZF precoding," *IEEE Trans. Wireless Commun.*, vol. 12, pp. 1554–1561, Apr. 2013.

[132] D. J. Love, R. W. Heath, and T. Strohmer, "Grassmannian beamforming for multiple-input multiple-output wireless systems," *IEEE Trans. Inf. Theory*, vol. 49, pp. 2735–2747, Oct. 2003.

[133] A. Medra and T. N. Davidson, "Flexible codebook design for limited feedback systems via sequential smooth optmization on the Grassmannian manifold," *IEEE Trans. Signal Process.*, vol. 62, pp. 1305–1318, Mar. 2014.