Imperial College London
Department of Electrical and Electronic Engineering

A Rate-Splitting Approach to Multiple-Antenna Broadcasting

Hamdi Joudeh

2016

Supervised by Dr. Bruno Clerckx

Submitted in part fulfilment of the requirements for the degree of
Doctor of Philosophy in Electrical and Electronic Engineering of Imperial College London
and the Diploma of Imperial College London

Statement of Originality

I declare that this thesis is the result of my own work. Information and ideas derived from the work of others has been acknowledged in the text and a list of references is given in the bibliography. The material of this thesis has not been and will not be submitted for another degree at any other university or institution.

Copyright Declaration

To my parents

Abstract

Signal processing techniques for multiple-antenna transmission can exploit the spatial dimension of the wireless channel to serve multiple users simultaneously, achieving high spectral efficiencies. Realizing such gains; however, is strongly dependent on the availability of highly accurate and up-to-date Channel State Information at the Transmitter (CSIT). This stems from the necessity to deal with multiuser interference through preprocessing; as receivers cannot coordinate in general. In wireless systems, CSIT is subject to uncertainty due to estimation and quantization errors, delays and mismatches. This thesis proposes optimized preprocessing techniques for broadcasting scenarios where a multi-antenna transmitter communicates with single-antenna receivers under CSIT uncertainties.

First, we consider a scenario where the transmitter communicates an independent message to each receiver. The most popular preprocessing techniques in this setup are based on linear precoding (or beamforming). Despite their near-optimum rate performances when highly accurate CSIT is available, we show that such techniques exhibit severe losses under CSIT uncertainties, even when optimally designed. We depart from this conventional approach and adopt an unorthodox transmission strategy based on Rate-Splitting (RS), which relies on broadcasting a common data stream on top of the private data streams precoded using partial CSIT. We propose an average Weighted Minimum Mean Square Error (WMMSE) algorithm to maximize the ergodic sum-rate performance.

While the ergodic sum-rate measure captures the long-term overall performance, it is not well suited for delay-limited or fairness based transmissions. Hence, we generalize the RS strategy to formulate the problem of achieving robust max-min fairness over one random fading state under a bounded CSIT uncertainty model. We derive new performance limits in terms of the symmetric-DoF under heterogeneous CSIT qualities across users to identify the RS gains. Then, a robust WMMSE algorithm based on the cutting-set method is proposed to solve the semi-infinite optimization problem. This framework is extended to address the problem of power minimization under Quality of Service (QoS) constraints.

Finally, we consider the problem of achieving max-min fairness in a multigroup multicasting scenario, where each message is intended to a group of users. We assume perfect CSIT in this setup, where the presence of multiple users in each group is thought of as a source of (finite) uncertainty. The DoF performance of conventional beamforming techniques are derived from which their limitations are identified. The RS strategy is then extended to this scenario, where we show that significant DoF gains can be achieved. The RS precoder optimization problem in this setup is then solved using the WMMSE approach.

Acknowledgements

I would like to express my sincere gratitude and appreciation to my supervisor Dr. Bruno Clerckx. His insightful suggestions set me on the right path from the very beginning, and his continued support and encouragement guided me through this gratifying journey. Bruno's mentorship and dedication will always be a source of inspiration, and I feel very privileged to have worked under his supervision.

I would like to extend my sincere gratitude to my friends and colleagues from Bruno's research group: Chenxi for introducing me to the art of DoF analysis, Borzoo for simplifying the complicated information theoretic concepts, Mingbo for the oriental-canteen time and all the lemmas, and Yang for the fruitful discussions on optimization theory. This work would not have been possible without the help I received from each one of you.

I am also truly thankful to my friends from the Communications and Signal Processing group: Jon for taking me in when I first arrived at the glorious Level 8, John for all the useful and useless discussions and being my Gym/Basketball-buddy (or rather allowing me to be yours), and Zexi for all the help with producing elegant content using LaTeX.

Special heartfelt thanks go to my brothers and sisters, friends and loved ones in Gaza, London and around the globe. Each one of you played an irreplaceable role in my life and helped me become the person I am. Finally, words are not enough to express all my gratitude to my wonderful parents; an infinite *thank you* for your boundless love, countless sacrifices and endless support in every possible manner.

Hamdi M. Joudeh
London, May 2016.

Contents

List of Figures

1. Introduction

Wireless communication has become an integral part of our daily life. The emergence of mobile broadband has enabled a plethora of applications touching every aspect of modern society. Conversely, the development of new applications fuels an ever growing demand for ubiquitous high-quality connectivity. Meeting such demand requires the continuous development of highly efficient and robust systems.

Multiple-Input-Multiple-Output (MIMO) technology has been at the heart of modern advancements in the physical layer of wireless systems. MIMO has developed beyond the classical point-to-point channel, and nowadays includes multiuser setups where spatial dimensions are exploited to serve several receivers simultaneously. High multiplexing gains are realized in multiuser setups with far less restrictions on the scattering environment compared to the point-to-point channel [1]. However, this comes with a price of higher restrictions imposed on the quality of the Channel State Information (CSI) required at the tranmsitter, specifically in the downlink mode. This stems from the necessity to deal with multiuser interference through preprocessing, as the receivers are generally decentralized and cannot cooperate. The ability to provide highly accurate and up-to-date CSI at the Transmitter (CSIT) is questionable. Therefore, considerable effort has been devoted to characterize and improve the performance in the presence of CSIT uncertainties.

1.1. Background and Motivation

Consider a Multiuser Multiple-Input-Single-Output (MU-MISO) system operating in downlink, where a Base Station (BS) equipped with multiple antennas communicates with K single-antenna receivers. This system is illustrated in Figure 1.1. The signal received by the kth user is expressed by

$$y_k = \mathbf{h}_k^H \mathbf{x} + n_k \tag{1.1}$$

where $\mathbf{h}_k$ denotes the vector channel, $\mathbf{x}$ is the transmitted signal, and n_k is received noise.

A typical transmission scenario is one where the BS wishes to communicate independent data to each user. It is also typical to construct the transmitted signal as a superposition of the K users' data signals such that

$$\mathbf{x} = \mathbf{x}_1 + \mathbf{x}_2 + \ldots + \mathbf{x}_K \tag{1.2}$$

where $\mathbf{x}_k$ denotes the signal intended to the kth user. Due to the broadcast nature of

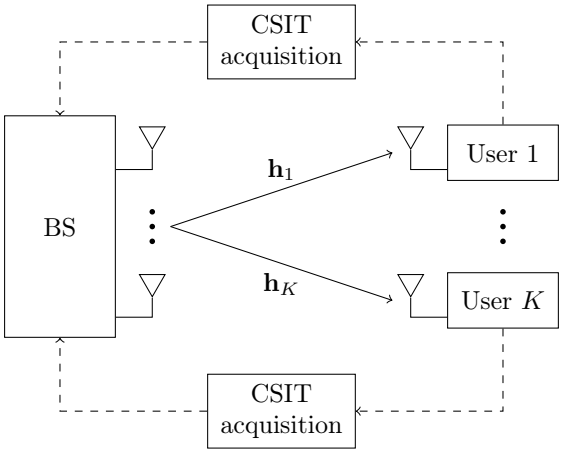


Figure 1.1.: MU-MISO system with K users operating in downlink. The acquisition of CSIT is typically done through quantized feedback or uplink training.

the wireless channel, each user receives a linear transformation of the transmitted signal with all its components. The undesired parts of the received signal is known as multiuser interference. Such interference can have a detrimental effect on the performance of the wireless system by limiting each receiver's ability to recover its desired signal.

One way to eliminate interference is to place each user's signal in the null space of all other users' channels. In particular, we have

$$\mathbf{x}_k \in \text{null}\left([\mathbf{h}_1, \ldots, \mathbf{h}_{k-1}, \mathbf{h}_{k+1}, \ldots, \mathbf{h}_K]^H\right) \tag{1.3}$$

which is known as the Zero Forcing (ZF) strategy. In this case, no user sees the signal of another user, and the received signal reduces to

$$y_k = \mathbf{h}_k^H \mathbf{x}_k + n_k. \tag{1.4}$$

By eliminating multiuser interference, each user is said to achieve a full Degree of Freedom (DoF), i.e. a DoF of 1. A DoF in this context can be understood as a an interference-free signal dimension, entirely reserved for the user throughout the transmission. Hence, the total DoF (sum-DoF) of the channel in this case is K.

One implicit condition required to achieve interference-free transmission is that the number of antennas should be at least equal to K. This is usually guaranteed by scheduling the *right* number of users in each transmission period[1]. Another condition is that the BS should know the CSI, characterized in this case by the channel matrix given by

$$\mathbf{H} \triangleq [\mathbf{h}_1, \ldots, \mathbf{h}_K]. \tag{1.5}$$

Typically, the BS acquires such information through feedback links in Frequency Division

[1]Another condition is that the channel matrix, comprising channel vectors of all users, should be well conditioned. This is assumed to be satisfied in multiuser scenarios.

Duplex (FDD) systems or uplink training pilots in Time Division Duplex (TDD) systems.

Assuming that $\mathbf{H}$ is perfectly known by the BS is somehow optimistic, as both methods are subject to errors. Moreover, wireless/mobile channels are subject to fading and are expected to change. Hence, even if the acquired CSI is of high quality, unless it was updated on a regular basis, it may be partially or fully outdated.

1.1.1. No CSIT

The essential value of CSIT is in the *directions* of $\mathbf{h}_1, \ldots, \mathbf{h}_K$. This enables the BS to place the signal of each user in the null space of all others as shown in (1.3). Assuming the extreme case of no CSIT[2], the BS loses such ability. In this case, the BS should resort to an alternative method of multiuser interference management.

One method is to transmit each of the signals $\mathbf{x}_1, \ldots, \mathbf{x}_K$ in an orthogonal *slot* (time or frequencey). This can be achieved using Time Division Multiple Access (TDMA), where each signal gets the entire channel to itself for a fraction of the transmission period. The received signal is still expressed by (1.4), as multiuser interference is eliminated. However, each user gets only a fraction of this interference-free transmission. By allocating time equally amongst users, a DoF of $1/K$ per user is achieved. The sum-DoF in this case is 1.

1.1.2. Partial CSIT

The absence of CSIT has a detrimental effect on the achievable DoF. However, this assumption is overly pessimistic, particularly in modern wireless communication systems where a form of CSIT acquisition is carried out as illustrated in Figure 1.1. The CSIT acquired by the BS, also known as the channel estimate, for the kth user can be expressed by

$$\widehat{\mathbf{h}}_k = \mathbf{h}_k - \widetilde{\mathbf{h}}_k \tag{1.6}$$

where $\widetilde{\mathbf{h}}_k$ denotes the channel estimation error. To evaluate the influence of CSIT uncertainty on the DoF, it is customary to look at the CSIT scaling laws with the Signal to Noise Ratio (SNR). Let P_{t} be the transmit SNR[3]. The CSIT error scales with SNR such that $\|\widetilde{\mathbf{h}}_k\|^2 = O\big(P_{\mathrm{t}}^{-\alpha}\big)$, for some scaling factor $\alpha \in [0, 1]$. This factor characterizes the quality of the CSIT as explained in the following discussion.

Consider the same ZF design in (1.3), except that now the BS uses imperfect CSIT as

$$\mathbf{x}_k \in \mathrm{null}\left(\left[\widehat{\mathbf{h}}_1, \ldots, \widehat{\mathbf{h}}_{k-1}, \widehat{\mathbf{h}}_{k+1}, \ldots, \widehat{\mathbf{h}}_K\right]^H\right). \tag{1.7}$$

[2]The BS may know the channel gains, i.e. $\|\mathbf{h}_1\|, \ldots, \|\mathbf{h}_K\|$, but not the directions.

[3]Assuming a normalized noise variance, this corresponds to the transmission power at the BS.

This yields a received signal given by

$$y_k = \overbrace{\mathbf{h}_k^H \mathbf{x}_k}^{O(P_\mathrm{t})} + \overbrace{\widetilde{\mathbf{h}}_k^H \sum_{i \neq k} \mathbf{x}_i}^{O(P_\mathrm{t}^{1-\alpha})} + \overbrace{n_k}^{O(P_\mathrm{t}^0)} \tag{1.8}$$

which bears *residual* multiuser interference, resulting from the imperfect interference nulling. Note that the SNR scalings of different received signal components are shown in (1.8). To explain such scalings, consider the noise component. This clearly does not scale with SNR, as the effects of noise can be eliminated by making P_t sufficiently large. The residual interference on the other hand scales as $O(P_\mathrm{t}^{1-\alpha})$, which follows from the CSIT error scaling. Such scaling is what causes the loss in DoF.

It is clear from (1.8) that the Signal to Interference plus Noise Ratio (SINR) scales as $O(P_\mathrm{t}^{\alpha})$. Hence, it follows that the data rate achieved by the kth user can be expressed by

$$\alpha \log_2(P_\mathrm{t}) + O(1). \tag{1.9}$$

The pre-log factor in (1.9) is in fact the achievable DoF[4]. The residual multiuser interference in (1.8) occupies a fraction $1-\alpha$ of the signal space, leaving the remaining α for the useful signal. Hence, ZF achieves a fraction α of the full DoF, i.e. αK, under partial CSIT.

Note that $\alpha = 1$ is considered perfect in the DoF sense, as the residual interference would scale as $O(P_\mathrm{t}^0)$. This is similar to the noise scaling, and can be eliminated by increasing P_t. On the other hand, $\alpha = 0$ corresponds to no CSIT in the DoF sense, as simultaneous interference-free multiuser transmission is not possible.

1.1.3. A Rate-Splitting Approach

Recall that from the BS's point of view, the channel in (1.6) is a superposition of a known component and an unknown component. Hence, blindly applying a scheme based on the known component only is in fact not the best strategy. Instead, a better suited strategy is one that bridges transmission with perfect CSIT (based on ZF) and transmission with no CSIT (e.g. TDMA). To achieve this, the transmitted signal is divided (or split) into two corresponding parts: one to exploit the known component of the channel, while the other is adapted to the unknown component. This transmission strategy is known in literature as Rate-Splitting (RS) [2,3], or signal space partitioning [4].

Let us first assume, without loss of generality, that this splitting is performed on the signal of user-1. The corresponding transmitted signal of user-1 writes as

$$\mathbf{x}_1 + \mathbf{x}_\mathrm{c} \tag{1.10}$$

where the notation $\mathbf{x}_1$ is retained for one of the parts for ease of the following analysis. The

[4]Also known as the multiplexing gain in the MIMO literature.

signals given by $\mathbf{x}_1, \ldots, \mathbf{x}_K$ are meant to exploit the known part of the channel, and hence they are designed according to (1.7). On the other hand, $\mathbf{x}_\mathrm{c}$ is broadcasted in an isotropic manner (with no preferred direction), resembling transmission with no CSIT. Now, let us allocate the average powers such that[5]

$$\mathbb{E}\left\{\|\mathbf{x}_1\|^2\right\} = \mathbb{E}\left\{\|\mathbf{x}_2\|^2\right\} = \ldots = \mathbb{E}\left\{\|\mathbf{x}_K\|^2\right\} = \frac{P_\mathrm{t}^\alpha}{K} \quad \text{and} \quad \mathbb{E}\left\{\|\widetilde{\mathbf{x}}_1\|^2\right\} = P_\mathrm{t} - P_\mathrm{t}^\alpha. \tag{1.11}$$

From this power allocation, it follows that the signal received by the kth user writes as

$$y_k = \overbrace{\mathbf{h}_k^H \mathbf{x}_\mathrm{c}}^{O(P_\mathrm{t})} + \overbrace{\mathbf{h}_k^H \mathbf{x}_k}^{O(P_\mathrm{t}^\alpha)} + \overbrace{\widetilde{\mathbf{h}}_k^H \sum_{i \neq k} \mathbf{x}_i}^{O(P_\mathrm{t}^0)} + \overbrace{n_k}^{O(P_\mathrm{t}^0)}. \tag{1.12}$$

First, it can be seen that $\mathbf{x}_\mathrm{c}$ is received by all users with a SINR that scales as $O\big(P_\mathrm{t}^{1-\alpha}\big)$. Hence, it can be decoded by all users achieving a DoF of $1-\alpha$. Note that while $\mathbf{x}_\mathrm{c}$ is not required by all users, decoding it allows them to remove it from the received signal, and then decode their own signals under reduced interference. After removing $\mathbf{x}_\mathrm{c}$, it can be seen that $\mathbf{x}_k$ is received with a SINR scaling as $O\big(P_\mathrm{t}^\alpha\big)$, achieving a DoF of α.

By virtue of the power allocation in (1.11), the signal space is divided into two partitions: one where transmission with no CSIT is carried out, and another where ZF is performed using partial CSIT without *leaking* significant interference. Hence, an additional DoF of $1-\alpha$ is achieved through $\mathbf{x}_\mathrm{c}$, while retaining the $K\alpha$ achieved through ZF. Note that in the above example, the DoF of $1-\alpha$ achieved with no CSIT is given entirely to user-1. This DoF can be further divided amongst users in a manner suitable for its no CSIT nature, e.g. TDMA. The sum-DoF is always given by $1 + (K-1)\alpha$.

1.1.4. Design and Optimization

The RS strategy described above provides a general guideline to a communication scheme suited for transmission with partial CSIT, with some insight into the signal design and power allocation in the high SNR regime. However, this is far from enough to achieve an optimized performance over the range of finite SNRs. As we can see from (1.9), while the DoF describes the rate scaling for high SNRs, it does not say much (if anything at all) about the actual achievable rates for specific finite (or low) SNR values.

To highlight the importance of signal optimization, consider $\mathbf{x}_1, \ldots, \mathbf{x}_K$ in (1.12) designed using the ZF strategy in (1.7). While such strategy achieves the aspired DoF, it is known to be suboptimal at finite SNRs. An optimized design on the other hand strikes a delicate balance between interference nulling and maximizing the desired signal power [5]. Moreover, $\mathbf{x}_\mathrm{c}$ was designed with no CSIT knowledge at all, as such knowledge does not provide DoF gains in this case. However, since $\mathbf{x}_\mathrm{c}$ is decoded by all users, partial CSIT can still be leveraged to enhance the performance at finite SNRs through a multicast-like design [6].

[5]Note that the powers in (1.11) satisfy a total power constraint of P_t.

The signal optimization at finite SNRs requires a comprehensive CSIT uncertainty model, as the SNR scaling laws are not sufficient to formulate such design problem. For example, CSIT uncertainty with known statistical properties can be used to formulate the problem in terms of average-based or outage-based performance metrics. Moreover, the objective of the design should be clearly identified, based on the considered scenario and application. The objective may be to maximize the sum-rate performance, where very little regard is given to the individual rates achieved by users. Another objective is maximizing the minimum rate amongst users to achieve fairness, where more attention is paid to individual rates. This may highly influence the manner in which resources are allocated, signals are divided and signal space partitions are accessed by different users.

This thesis proposes RS-based signal designs for a number MU-MISO scenarios. The different scenarios share certain features, making RS applicable to all of them. However, underlying assumptions and design objectives differ from one scenario to the other, yielding optimization problems with distinct structures and challenges.

1.2. Outline and Contributions

Preprocessing techniques based on linear precoding are adopted throughout this thesis. Hence, the transmitted signal model in (1.2) is expressed by

$$\mathbf{x} = \mathbf{p}_{\mathrm{c}} s_{\mathrm{c}} + \sum_{k=1}^{K} \mathbf{p}_k s_k. \tag{1.13}$$

This reduces the signal optimization problem to designing the precoding vectors in (1.13) such that a certain objective function is maximized, subject to some constraints, and under a well-defined CSIT uncertainty model. Such *robust*[6] precoder optimization, in the light of the RS transmission strategy, is the main theme of this thesis.

A summary of each chapter, highlighting the main contributions, is given as follows.

- **Chapter 2: Preliminaries** lays the foundations for this thesis by describing the signal and channel models, the main assumptions regarding CSIT uncertainty, in addition to introducing important concepts and definitions used throughout this thesis.

- **Chapter 3: Sum-Rate Maximization** considers a transmission that takes place over a long sequence of fading states. The objective is to design an adaptive precoding scheme that updates the precoders based on the available CSIT to maximize the ergodic sum-rate performance subject to a transmission power constraint. This is achieved by solving a conditional average rate optimization problem for each incoming CSIT state. To solve the stochastic optimization problem at hand, we propose an

[6]This comes from the uncertainty in the model due to partial CSIT. The notion of robustness will be discussed in detail in the following chapters.

algorithm based on the Weights Minimum Mean Square Error (WMMSE) approach [7, 8] coupled with means of Sample Average Approximation (SAA) [9].

- **Chapter 4: Fairness and Quality of Service** considers transmission over one random fading state under a bounded CSIT uncertainity model. We consider the classical robust design problem (in the worst-case sense) of achieving max-min fairness in the light of the new RS strategy. First, we derive new performance limits to identify the benefits of employing RS to achieve fairness amongst users with heterogeneous CSIT qualities. Such gains are characterized in terms of the symmetric-DoF. Then we develop a robust WMMSE algorithm based on the cutting-set method [10] to solve the semi-infinite optimization problem. Finally, the framework is extended to the problem of power minimization under worst-case Quality of Service (QoS) constraints.

- **Chapter 5: Multigroup Multicasting** departs from the unicast transmission scenario considered in two previous chapters Alternatively, we consider a multigroup multicast transmission, where each signal is intended to a group of users. Perfect CSIT is assumed in this chapter, and we consider the problem of achieving max-min fairness amongst groups. A resemblance to the MU-MISO setups considered in previous chapters is observed by viewing each group of users as a receiver, and each user in a group as a channel state in a finite uncertainty region. First, we show that conventional beamforming designs suffer from inter-group interference that limits the performance, particularly when the number of transmitting antennas is insufficient. Then, we formulate the RS multigroup multicasting problem, and we derive performance limits that demonstrate the superiority of this strategy. Finally, the design problem is solved using the WMMSE approach.

- **Chapter 6: Conclusion** provides the concluding remarks of the thesis and suggests a number of future research directions.

The schemes and algorithms proposed in the three core chapters, i.e. Chapter 3, Chapter 4 and Chapter 5, are evaluated and analysed through simulations under various parameters. For the sake of readability and to avoid interruptions, lengthy proofs of the main results are relegated to the appendices. However, important ideas in the proofs, discussions and insights drawn from such results are mentioned in the chapters.

1.3. Publications

The work in this thesis has resulted in a number of papers, that have been published, accepted or submitter for publication, or are still under preparation.

To be submitted

- **H. Joudeh** and B. Clerckx, "Multigroup Multicast Beamforming in Overloaded Systems: A Rate-Splitting Approach," (working title) *to be submitted*.

Peer-reviewed journals and magazines

- **H. Joudeh** and B. Clerckx, "Robust Transmission in Downlink Multiuser MISO Systems: A Rate-Splitting Approach," *submitted to IEEE Transactions on Signal Processing*, (*accepted with minor corrections*).

- **H. Joudeh** and B. Clerckx, "Sum-rate maximization for linearly precoded downlink multiuser MISO systems with partial CSIT: A Rate-Splitting Approach," *submitted to IEEE Transactions on Communications*, (*major corrections*).

- B. Clerckx, **H. Joudeh**, C. Hao, M. Dai, and B. Rassouli, "Rate Splitting for MIMO Wireless Networks: A Promising PHY-Layer Strategy for LTE Evolution," *IEEE Communication Magazine, special issue on LTE Evolution*, 2016.

Peer-reviewed conferences

- **H. Joudeh** and B. Clerckx, "A Rate-Splitting Strategy for Max-Min Fair Multigroup Multicasting," *IEEE International workshop on Signal Processing advances in Wireless Communications (SPAWC)*, 2016.

- **H. Joudeh** and B. Clerckx, "A Rate-Splitting Approach To Robust Multiuser MISO Transmission," *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, 2016.

- **H. Joudeh** and B. Clerckx, "Sum rate maximization for MU-MISO with partial CSIT using joint multicasting and broadcasting," *IEEE International Conference on Communications (ICC)*, 2015.

- **H. Joudeh** and B. Clerckx, "Achieving max-min fairness for MU-MISO with partial CSIT: A multicast assisted transmission," *IEEE International Conference on Communications (ICC)*, 2015.

1.4. Notations

The following notations are used throughout the thesis. Boldface uppercase denote matrices (e.g. $\mathbf{A}$), boldface lowercase denote vectors (e.g. $\mathbf{a}$), and standard letters denote scalars (e.g. a). The identity matrix is denoted by $\mathbf{I}$, where the dimensions should be clear from the context. $\mathbf{e}_i$ is the ith column of matrix $\mathbf{I}$ with the appropriate dimension. The superscrips $(\cdot)^T$, $(\cdot)^H$ and $(\cdot)^\dagger$ denote the transpose, conjugate-transpose (Hermitian), and

pseudo-inverse operators, respectively. $\mathrm{tr}(\cdot)$ denotes the trace operator. $\mathbf{A} \succeq 0$ means that the symmetric matrix $\mathbf{A}$ is positive semidefinite. $\mathrm{diag}(\mathbf{a})$ is a diagonal matrix, with diagonal entries given by the elements of $\mathbf{a}$. On the other hand, $\mathrm{diag}(\mathbf{A})$ is a vector with entries given by the diagonal elements of $\mathbf{A}$. The Euclidian norm is denoted by $\|\cdot\|$. The expectation with respect to a random variable X is denoted by $\mathbb{E}_{\mathrm{X}}\{\cdot\}$. Finally, we define the function $(x)^+ \triangleq \max\{x, 0\}$.

2. Preliminaries

This chapter lays the foundations for this thesis by: describing the system model and main assumptions, and introducing important terminology and definitions. This thesis focuses on setups where a single transmitter equipped with multiple antennas communicates with multiple single-antenna receivers under Additive White Gaussian Noise (AWGN). In information theoretic terms, such setups are known as Gaussian Multiple-Input-Single-Output (MISO) Broadcast Channels (BCs), or MISO-BCs for short. We start by describing a general signal model for the MISO-BC, while specifying particular classes of MISO-BCs which are relevant to this work. This is followed by describing the main channel fading and CSIT assumptions. The conditions under which reliable communication can take place are then reviewd. Moreover, we introduce the linear precoded signal model adopted throughout this thesis. Finally, we look at the influence of imperfect CSIT and introduce the RS scheme.

2.1. MISO Broadcast Channels

A BC models the downlink of a wireless system consisting of one transmitter (or information source) and multiple uncoordinated receivers[1]. The terminology reflects the fact that communication takes place over a shared broadcast medium. While the usage of the term broadcasting in common language is almost restricted to transmitting common information to all receivers, e.g. radio or TV broadcasting, the information theoretic terminology is in fact much more general [11, 12]. Indeed, a BC may refer to a system where the transmitter communicates independent messages to the receivers, or a mixture of independent and common messages [12], just to name a few.

A BC where the transmitter and (possibly) the receivers are equipped with multiple antennas is known as a MIMO-BC. This thesis focuses on the special case where each receiver is equipped with only one antenna, known as the MISO-BC[2]. This setup has received considerable attention in both academia and industry, as MIMO gains are realized using (multiple) simple and power efficient receiving devices with limited processing capabilities [1]. Next, the general input-output relationship is described.

[1]The transmitter and receiver are occasionally referred to as a Base Station (BS) and a user respectively.
[2]This is referred to as a MIMO-BC in some works, where the multiple outputs correspond to receivers.

2.1.1. Signal Model

Consider a general MISO-BC comprising a transmitter equipped with $N_\mathrm{t} > 1$ antennas and K single antenna receivers. Due to the frequent use of the user index set, it is denoted by $\mathcal{K} \triangleq \{1, \ldots, K\}$ in the following. For a transmission occurring over a sequence of T discrete uses of a narrow-band channel, the complex baseband signal at the kth receiver in the tth channel use, where $k \in \mathcal{K}$ and $t \in \{1, \ldots, T\}$, is expressed by

$$y_k(t) = \mathbf{h}_k^H(t)\mathbf{x}(t) + n_k(t) \tag{2.1}$$

where $\mathbf{h}_k(t) \in \mathbb{C}^{N_\mathrm{t}}$ is the channel vector between the transmitter and the kth receiver, $\mathbf{x}(t) \in \mathbb{C}^{N_\mathrm{t}}$ is the input (transmitted) signal and $n_k(t) \sim \mathcal{CN}(0, \sigma_{\mathrm{n},k}^2)$ is the AWGN. Without loss of generality, it is assumed that $\sigma_{\mathrm{n},1}^2, \ldots, \sigma_{\mathrm{n},K}^2 = \sigma_\mathrm{n}^2$, and any variation in noise powers among users is absorbed into the channel gains. We assume that $\sigma_\mathrm{n}^2 > 0$ and remains fixed over the entire SNR range. Hence, $\mathrm{SNR} \to \infty$ is equivalent to $P_\mathrm{t} \to \infty$.

The input signal is subject to an average power constraint per channel use given by

$$\mathbb{E}_\mathrm{x}\left\{\|\mathbf{x}(t)\|^2\right\} \leq P_\mathrm{t} \tag{2.2}$$

where the expectation is taken over the distribution of $\mathbf{x}(t)$. While it is common to consider P_t a given parameter in power-constrained transmission, it may correspond to a variable in performance-constrained transmission. Unless stated otherwise, power-constrained transmission is assumed where P_t is a given parameter.

The sequence of channel inputs over the transmission period writes as $\{\mathbf{x}(t)\}_{t=1}^T$, and is commonly referred to as a *codeword*. The signal model in (2.1) is general as it does not specify the number of information messages encoded into the transmitted signal, and the set of receivers each message is intended to, which define various classes of the MISO-BC. Moreover, assumptions regarding the employed encoding/decoding strategy and the fading process that describes the variation of the channel during the transmission, are also left open. Such assumptions are made explicit and described in detail when necessary.

2.1.2. Different Classes

The most common class of BCs is that with private messages, where the transmitter communicates independent information to each receiver. This is generally referred to in literature as a BC with no further specification. Such terminology is also adopted in this thesis where a MISO-BC with no specification is used to refer to the MISO-BC with private messages. This class is the main focus of the remainder of this chapter, and the two following chapters.

Another class of BCs is that where a common message is added on top of the private messages. This common message is intended to, and decoded by, all users in the system. This class is usually called a BC with a common message. This channel plays a role in establishing the RS concept as we see later in this chapter.

2.2. Fading and Channel State Information

The matrix with columns constituting the K channel vectors is denoted by

$$\mathbf{H}(t) \triangleq \big[\mathbf{h}_1(t), \ldots, \mathbf{h}_K(t)\big] \tag{2.3}$$

which is referred to as the (instantaneous) channel state. We assume a *block-fading* model where $\mathbf{H}(t)$ remains fixed over T_b channel uses, and varies from one block to another according to some distribution. Let $\mathbf{H}[b]$ be the channel state in the bth block such that

$$\mathbf{H}(t) = \mathbf{H}[b], \forall t \in \big\{(b-1)T_\mathrm{b}+1, \ldots, bT_\mathrm{b}\big\}. \tag{2.4}$$

The transmitter obtains an estimate of the channel state in each block, typically through uplink training in TDD systems [13, 14] or quantized feedback in FDD systems [15]. Denoting the instantaneous CSIT by $\widehat{\mathbf{H}}(t) \triangleq \big[\widehat{\mathbf{h}}_1(t), \ldots, \widehat{\mathbf{h}}_K(t)\big]$ and the estimate obtained in the bth block by $\widehat{\mathbf{H}}[b]$, we have

$$\widehat{\mathbf{H}}(t) = \widehat{\mathbf{H}}[b], \forall t \in \big\{(b-1)T_\mathrm{b}+1, \ldots, bT_\mathrm{b}\big\}. \tag{2.5}$$

The joint channel state $\big\{\mathbf{H}[b], \widehat{\mathbf{H}}[b]\big\}$ is assumed to evolve according to some known ergodic stationary process with a joint probability density $f_{\mathrm{H},\widehat{\mathrm{H}}}\big(\mathbf{H}, \widehat{\mathbf{H}}\big)$ [16]. This fading process, which ticks at the block rate, is expressed by

$$\underbrace{\big\{\mathbf{H}[1], \widehat{\mathbf{H}}[1] \mid t = 1, \ldots, T_\mathrm{b}\big\}}_{\text{block } 1}, \underbrace{\big\{\mathbf{H}[2], \widehat{\mathbf{H}}[2] \mid t = T_\mathrm{b}+1, \ldots, 2T_\mathrm{b}\big\}}_{\text{block } 2},$$
$$\ldots, \underbrace{\big\{\mathbf{H}[b], \widehat{\mathbf{H}}[b] \mid t = (b-1)T_\mathrm{b}+1, \ldots, bT_\mathrm{b}\big\}}_{\text{block } b}, \ldots$$

The transmission period is taken as an integer multiple of the block period, i.e. $T = BT_\mathrm{b}$ where $B \in \mathbb{N}$. Under such assumptions, two different transmission scenarios are considered in this thesis:

1. **Non-ergodic transmission**: restricted to $T = T_\mathrm{b}$ (or $B = 1$), where the transmission is carried out over one random joint channel state.

2. **Ergodic transmission**: $T \gg T_\mathrm{b}$ (or $B \gg 1$), where the transmission takes place over a long sequence of blocks spanning almost all possible joint channel states.

For non-ergodic transmission, the channel state appears invariant throughout the entire transmission. This may be interpreted as a delay-limited transmission. However, T_b is assumed to be large enough such that the information-theoretic limits for reliable communication are approached [16,17]. On the other hand, ergodic transmission may be interpreted as delay-unlimited transmission. This is carried out to exploit some long-term properties of the channel as seen in Chapter 3.

Before proceeding, it should be highlighted that perfect CSIR is assumed throughout the thesis, i.e. each receiver is able to estimate and track its channel with very high accuracy. This is a common and well justified assumption as highly accurate CSIR can be obtained through downlink training at a small overhead cost [16, 17].

2.2.1. CSIT Uncertainity

The transmitter has perfect instantaneous CSI when $\widehat{\mathbf{H}}(t) = \mathbf{H}(t)$ for all $t \in \{1, \ldots, T\}$. This is generally not possible as the channel estimate is subject to uncertainties arising from estimation errors in TDD [13, 14], quantization errors in FDD [15] and *staleness* due to delays [18]. To model the CSIT uncertainty, the channel state is expressed by

$$\mathbf{H}(t) = \widehat{\mathbf{H}}(t) + \widetilde{\mathbf{H}}(t) \tag{2.6}$$

where $\widetilde{\mathbf{H}}(t) \triangleq \left[\widetilde{\mathbf{h}}_1(t), \ldots, \widetilde{\mathbf{h}}_K(t)\right]$ denotes the CSIT error not known to the transmitter[3]. The CSIT uncertainty is characterized by the conditional density $f_{\mathrm{H}|\widehat{\mathrm{H}}}\big(\mathbf{H} \mid \widehat{\mathbf{H}}\big)$, which takes non-zero values over the (possibly unbounded) set $\mathbb{H}\big(\widehat{\mathbf{H}}\big)$ known as the uncertainty region. This consists of all possible $\mathbf{H}$ that the channel state may take given a certain estimate $\widehat{\mathbf{H}}$. It is easy to see that $\mathbb{H}\big(\widehat{\mathbf{H}}[b]\big)$ reduces to $\mathbf{H}[b]$ under perfect CSIT. For brevity, the uncertainty region is denoted by $\mathbb{H}$, where the dependency on $\big(\widehat{\mathbf{H}}\big)$ is implicit.

2.2.2. CSIT Scaling with SNR

It is well established that CSIT uncertainties hinder the performance of the MISO-BC, as it may be impossible to eliminate or reduce the multiuser interference as seen in Section 1.1.2. This phenomenon was studies in [19] for errors arising from quantization and limited feedback. It was shown that maintaining the full DoF of the MISO-BC requires improving the CSIT quality with the SNR level through increasing the number of feedback bits. This idea of scaling quality was later adopted in a large number of works on the MISO-BC with partial and imperfect CSIT, for example see [2–4, 18, 20] and references therein.

Taking each user separately, the marginal density of the kth channel conditioned on its estimate writes as $f_{\mathrm{h}_k|\widehat{\mathrm{h}}_k}\big(\mathbf{h}_k \mid \widehat{\mathbf{h}}_k\big)$. Assuming that the mean of the conditional distribution is given by the estimate, i.e. $\mathbb{E}_{\mathrm{H}|\widehat{\mathrm{H}}}\{\mathbf{H} \mid \widehat{\mathbf{H}}\} = \widehat{\mathbf{H}}$, we have

$$\mathbb{E}_{\mathrm{h}_k|\widehat{\mathrm{h}}_k}\{\mathbf{h}_k\mathbf{h}_k^H \mid \widehat{\mathbf{h}}_k\} = \widehat{\mathbf{h}}_k\widehat{\mathbf{h}}_k^H + \mathbf{R}_{\mathrm{e},k} \tag{2.7}$$

where $\mathbf{R}_{\mathrm{e},k}$ is the kth user's CSIT error covariance matrix. The CSIT scaling is expressed by allowing the maximum entry of $\mathrm{diag}\,(\mathbf{R}_{\mathrm{e},k})$ to scale as $O(P_{\mathrm{t}}^{-\alpha_k})$ for some constant $\alpha_k \in [0, \infty)$. Equivalently, we have

$$\sigma_{\mathrm{e},k}^2 \triangleq \mathbb{E}_{\widetilde{\mathrm{h}}_k|\widehat{\mathrm{h}}_k}\left\{\|\widetilde{\mathbf{h}}_k\|^2\right\} = \mathrm{tr}(\mathbf{R}_{\mathrm{e},k}) = O(P_{\mathrm{t}}^{-\alpha_k}). \tag{2.8}$$

[3]The CSIT and CSIT error are occasionally referred to as the estimate and the estimation error respectively.

The constant $\alpha_k \triangleq \lim_{P_\mathrm{t}\to\infty} -\frac{\log(\sigma^2_{\mathrm{e},k})}{\log(P_\mathrm{t})}$ is known as the CSIT quality scaling factor (or exponent), which quantifies the CSIT quality as SNR grows large. For example, $\alpha_k \to \infty$ corresponds to perfect CSIT, as $\sigma^2_{\mathrm{e},k} \to 0$ for any SNR. The opposite extreme of $\alpha_k = 0$ represents a fixed quality w.r.t SNR, e.g. a constant number of quantization (and hence feedback) bits in FDD systems [19]. A finite non-zero α_k corresponds to CSIT quality that improves with increased SNR, e.g. by increasing the number of feedback bits.

Assumption 2.1. *Without loss of generality, an ascending order of CSIT quality factors is assumed throughout the thesis, i.e.*

$$\alpha_1 \leq \alpha_2 \leq \ldots \leq \alpha_K. \tag{2.9}$$

Moreover, the scaling factors are truncated such that $\alpha_k \in [0, 1]$, which is customary in DoF analysis as $\alpha_k = 1$ corresponds to perfect CSIT in the DoF sense [2, 20]. This follows from the fact that such quality is sufficient to reduce multiuser interference to the level of noise as seen in Section 1.1.2. It is also worth highlighting that α_k assumes various practical interpretations in addition to the limited feedback interpretation in [3, 19]. For example, it may also correspond to the Doppler process in delayed/outdated CSIT, where α_k is inversely proportional to the normalized Doppler frequency [2, 18].

2.3. Coding and Performance Limits

The transmitter *encodes* information messages it wishes to send to the receivers into a sequence of channel inputs $\{\mathbf{x}(t)\}_{t=1}^{T}$, broadcasted over T channel uses. The input sequence is commonly referred to as a *codeword*. At the other end of the channel, each receiver *decodes* the message(s) intended to it from the received sequence $\{y_k(t)\}_{t=1}^{T}$.

Consider a MISO-BC where the transmitter communicates private messages to the receivers as shown in Figure 2.1. The messages intended to receivers $1, \ldots, K$ are denoted by $W_1, \ldots, W_K$, which are uniformly distributed over the message sets $\mathcal{W}_1, \ldots, \mathcal{W}_K$, respectively. The corresponding rate of the kth user in bits per channel use is given by $r_k = \frac{\log_2(|\mathcal{W}_k|)}{T}$, from which a rate tuple consisting of all rates writes as $[r_1, \ldots, r_K]$. A *code* of rate $[r_1, \ldots, r_K]$ and length T, also known as a $\left(2^{Tr_1}, \ldots, 2^{Tr_K}, T\right)$ code, consists of:

1. K message sets $\mathcal{W}_1, \ldots, \mathcal{W}_K$.
2. A sequence of *encoding* functions $\{\phi(t)\}_{t=1}^{T}$, where $\phi(t)$ maps any CSIT and message tuple $[W_1, \ldots, W_K]$ into the input signal $\mathbf{x}(t)$.
3. K *decoding* functions $\psi_1, \ldots, \psi_K$, such that ψ_k maps received sequences $\{y_k(t)\}_{t=1}^{T}$ into the decoded message $\widehat{W}_k \in \mathcal{W}_k$.

The set of $2^{T\sum_{k=1}^{K} r_k}$ codewords associated with such code is denoted by $\mathbb{X}_T$, and is referred to as a *codebook*. Codes are subject to some input constraint, e.g. the average power

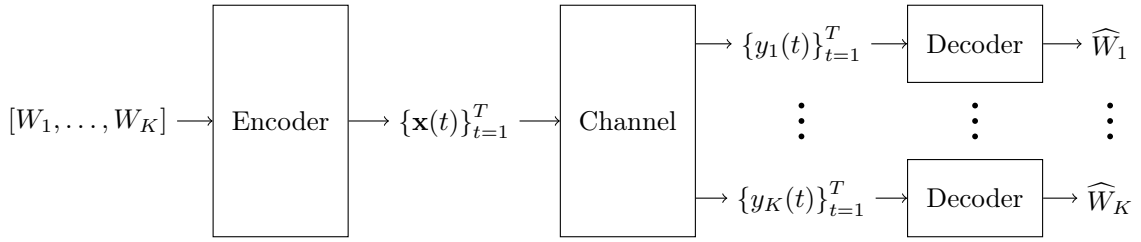


Figure 2.1.: K-user Broadcast Channel with private messages.

constraint in (2.2) which the codebook should comply with. The constraint could also be a certain codeword structure, as seen later with linear precoding.

The performance of a code is characterized by the decoding error probability, where a decoding error occurs when $\widehat{W}_k \neq W_k$ at any of the K receivers. Under a prescribed set of constraints, the rate tuple $[r_1, \ldots, r_K]$ is deemed achievable if there exists a sequence of $\left(2^{Tr_1}, \ldots, 2^{Tr_K}, T\right)$ codes such that every receiver is able to decode its message with an error probability that can be made arbitrarily small by making T sufficiently large, while satisfying the constraints. The coding strategy and achievable rates are highly dependent on the fading process and the accuracy of CSIT. For example, it is known that the availability of instantaneous CSIT in the MISO-BC allows for codes that eliminate (or reduce) multiuser interference such that higher rates can be achieved.

2.3.1. Capacity

The *capacity region* of the MISO-BC under a power constraint is defined as the closure of the set of all achievable rate tuples $[r_1, \ldots, r_K]$ such that the power constraint is not violated. This is denoted by $\mathcal{C}(P_\mathrm{t})$ for a power constraint P_t. The significance of such region is that it characterizes the optimum tradeoff between the rates that can be simultaneously achieved by different receivers. The capacity region is influenced by the type of transmission. For example, $\mathcal{C}(P_\mathrm{t})$ is a function of the instantaneous channel state in the non-ergodic case, while it normally depends on the long-term properties in ergodic transmission [1]. Moreover, $\mathcal{C}(P_\mathrm{t})$ is also highly influenced by the availability of CSIT and its accuracy [20–22]. Under perfect CSIT, the capacity region of the MISO-BC is achieved by a non-linear coding strategy known as Dirty Paper Coding (DPC) [23]. Under imperfect CSIT, the capacity region is unknown in general.

Several meaningful scalar performance measures can be derived from the capacity region. One popular example is the sum-capacity defined as

$$C_\mathrm{sum} \triangleq \max_{[r_1,\ldots,r_K]\in\mathcal{C}} \sum_{k=1}^{K} r_k \tag{2.10}$$

which corresponds to the maximum achievable sum-rate. While operating at the sum-capacity guarantees the optimum system throughput, it does not guarantee any form of

fairness amongst users. For instance, a receiver with a poor channel condition may be left to *starve* while receivers with better channel conditions are allocated all resources for the sake of the overall throughput. Another performance measure that accounts for fairness amongst users is the Max-Min Fair (MMF)-capacity defined as

$$C_{\text{MMF}} \triangleq \max_{[r_1,\ldots,r_2]\in\mathcal{C}} \min_{k\in\mathcal{K}} r_k. \tag{2.11}$$

This is also known as the symmetric-capacity [17], as the symmetric rate tuple given by $[C_{\text{MMF}},\ldots,C_{\text{MMF}}]$ is optimum in the MMF sense. The measures in (2.10) and (2.11), or suboptimal versions of them that do not necessarily satisfy the maximization operators, are the main design objectives considered in this thesis.

2.3.2. Degrees of Freedom

The characterization of the MISO-BC's capacity region under imperfect CSIT in still an open problem. A tractable alternative is the DoF region, a first order approximation of the capacity region in the high SNR regime. As opposed to the capacity region, considerable progress has been made towards characterizing achievable and optimum DoF regions of the MISO-BC under a variety of imperfect CSIT scenarios [2, 24–27].

To better understand the DoF concept, let us introduce the notion of a *coding scheme*. A coding scheme of length T can be thought of as a family of $\left(2^{Tr_1(P_\text{t})},\ldots,2^{Tr_K(P_\text{t})},T\right)$ codes, with one code for each SNR level [28]. The associated sequences of rate tuples and codebooks are denoted by $\left\{[r_1(P_\text{t}),\ldots,r_K(P_\text{t})]\right\}_{P_\text{t}}$ and $\left\{\mathbb{X}_T(P_\text{t})\right\}_{P_\text{t}}$ respectively. A coding scheme is also associated with a DoF tuple $[d_1,\ldots,d_K]$, where d_k is defined as

$$d_k \triangleq \lim_{P_\text{t}\to\infty} \frac{r_k(P_\text{t})}{\log_2(P_\text{t})}. \tag{2.12}$$

This metric, also known as the multiplexing gain, is independent of the actual transmit power (or SNR) and captures the rate's asymptotic slope with respect to $\log_2(P_\text{t})$.

Similar to the capacity region, a DoF region is defined as the closure of the set of all achievable DoF tuples [24]. This is expressed by

$$\mathcal{D} \triangleq \Bigg\{ [d_1,\ldots,d_K] \mid d_k \geq 0 \text{ and } \exists \big[r_1(P_\text{t}),\ldots,r_K(P_\text{t})\big] \in \mathcal{C}(P_\text{t}) \text{ such that } d_k \triangleq \lim_{P_\text{t}\to\infty} \frac{r_k(P_\text{t})}{\log_2(P_\text{t})}\ \forall k\in\mathcal{K} \Bigg\}. \tag{2.13}$$

It is clear that each achievable DoF tuple is associated with at least one sequence of achievable rate tuples, and $\mathcal{D}$ is associated with the sequence $\left\{\mathcal{C}(P_\text{t})\right\}_{P_\text{t}}$.

Now consider the interference-free transmission of only one data stream to a single user. The maximum achievable rate of this transmission is given by the capacity of a point-to-point MISO channel [29], which scales with P_t as $\log_2(P_\text{t}) + O(1)$. Hence, the maximum

DoF achieved by any user in the MISO-BC is upper-bounded by 1. It follows that the DoF can be interpreted as a fraction of an interference-free data stream that can be reliably communicated per channel use. The DoF region characterizes optimum tradeoff between fractions that are simultaneously assessable by the K receivers.

Similar to the scalar measures derived from the capacity region, one can derive scalar measures from the DoF region. The most common measure is the sum-DoF given by

$$d_{\text{sum}} \triangleq \max_{[d_1,\ldots,d_K]\in\mathcal{D}} \sum_{k=1}^{K} d_k. \tag{2.14}$$

A less common one is the MMF-DoF, also known as the symmetric-DoF, defined as

$$d_{\text{MMF}} \triangleq \max_{[d_1,\ldots,d_K]\in\mathcal{D}} \min_{k\in\mathcal{K}} d_k. \tag{2.15}$$

The following result characterizes such measures under perfect CSIT. An implicit assumption is that the channel is generic, i.e. channel vectors have entries bounded away from zero and infinity and are drawn from continuous distributions, and hence they are linearly independent almost surely.

Lemma 2.1. *For the MISO-BC with $N_{\text{t}} \geq K$ under perfect CSIT, we have*

$$d_{\text{sum}} = K \quad \textit{and} \quad d_{\text{MMF}} = 1.$$

The sum-DoF result is a very well known one frequently reported in literature [1,21]. The MMF-DoF follows directly, as achieving the sum-DoF result implies that the DoF tuple $[d_1, \ldots, d_K] = [1, \ldots, 1]$ is achievable. Lemma 2.1 tell us that multiuser interference can be completely dealt with by the coding scheme under perfect CSIT.

2.4. Linear Precoding

DPC is merely a theoretical concept, and its practical implementations is deemed highly complicated [30, 31]. Linear precoding, also known as Beamforming (BF), is a suboptimal strategy that has emerged as the most attractive alternative. In BF, each message is first encoded independently into a data stream. Each stream is then mapped to the N_{t} transmit antennas through a vector of BF weights, known as the precoding vector. This structure makes the problems of designing MISO-BC codes more tractable and less complex. We start be considering linear precoding for non-ergodic transmission, then the concept is generalized to ergodic transmission. Moreover, perfect CSIT is assumed in the following discussion. This is relaxed in the next section.

2.4.1. The Non-Ergodic Regime (Fixed Channel)

The message intended to the kth user is encoded into a stream of data symbols such that $W_k \mapsto \{s_k(t)\}_{t=1}^{T}$, where $s_k(t) \in \mathbb{C}$ denotes an encoded data symbol in the tth channel use. Since the channel is invariant, it is sufficient to focus on an arbitrary channel use by dropping (t) in the following discussion. Let $\mathbf{s}_{\mathrm{p}} \triangleq [s_1, \ldots, s_K]^T$ be the K data symbols in an arbitrary channel use, the linearly precoded channel input signal is constructed as

$$\mathbf{x} = \mathbf{P}_{\mathrm{p}}\mathbf{s}_{\mathrm{p}} = \sum_{k=1}^{K} \mathbf{p}_k s_k. \tag{2.16}$$

where $\mathbf{P}_{\mathrm{p}} = [\mathbf{p}_1, \ldots, \mathbf{p}_K]$ is the precoding matrix, and $\mathbf{p}_k \in \mathbb{C}^{N_{\mathrm{t}}}$ is the kth user's precoding vector consisting of N_{t} BF weights. The precoding matrix is assumed to remain fixed throughout the transmission. Assume that each of the independent data streams has a normalized average power. It follows that $\mathbb{E}_{\mathrm{s_p}}\{\mathbf{s}_{\mathrm{p}}\mathbf{s}_{\mathrm{p}}^H\} = \mathbf{I}$, and the power constraint reduces to a constraint on the precoding matrix such that

$$\mathbb{E}_{\mathrm{x}}\{\mathbf{x}^H\mathbf{x}\} = \mathrm{tr}\big(\mathbf{P}_{\mathrm{p}}\mathbf{P}_{\mathrm{p}}^H\big) = \sum_{k=1}^{K} \|\mathbf{p}_k\|^2 \leq P_{\mathrm{t}}. \tag{2.17}$$

The kth received signal in (2.1) now has a special structure, described as

$$y_k = \overbrace{\mathbf{h}_k^H \mathbf{p}_k s_k}^{\text{desired signal}} + \overbrace{\mathbf{h}_k^H \sum_{i \neq k} \mathbf{p}_i s_i}^{\text{interference}} + n_k \tag{2.18}$$

from which the average received signal power is expressed by

$$T_k \triangleq \mathbb{E}\{|y_k|^2\} = \overbrace{|\mathbf{h}_k^H \mathbf{p}_k|^2}^{S_k} + \overbrace{\sum_{i \neq k} |\mathbf{h}_k^H \mathbf{p}_i|^2 + \sigma_{\mathrm{n}}^2}^{I_k} \tag{2.19}$$

where S_k and I_k denote the desired power and interference plus noise power respectively. It follows that the SINR is given by

$$\gamma_k \triangleq \frac{S_k}{I_k} = \frac{\left|\mathbf{h}_k^H \mathbf{p}_k\right|^2}{\sum_{i \neq k} \left|\mathbf{h}_k^H \mathbf{p}_i\right|^2 + \sigma_{\mathrm{n}}^2}. \tag{2.20}$$

For a given precoding matrix, each receiver can view its channel as an equivalent scalar channel with a power gain $\left|\mathbf{h}_k^H \mathbf{p}_k\right|^2$ and additive noise of some distribution with variance $\sum_{i \neq k} \left|\mathbf{h}_k^H \mathbf{p}_i\right|^2 + \sigma_{\mathrm{n}}^2$. Under Gaussian coding, the interference plus noise is seen as AWGN, and the kth user's achievable rate writes as

$$R_k = \log_2(1 + \gamma_k). \tag{2.21}$$

It follows that for a given precoding matrix, the rate tuple $[R_1, \ldots, R_K]$ is achievable using scalar point-to-point AWGN codes. Hence, the problem of designing a MISO-BC code reduces to designing the precoding matrix $\mathbf{P}_{\mathrm{p}}$ that achieves such rates.

In some analysis, it is helpful to emphasise the structure of $\mathbf{P}_{\mathrm{p}}$. Hence, each precoding vector is expressed as $\mathbf{p}_k = \sqrt{q_k}\widehat{\mathbf{p}}_k$, where $\widehat{\mathbf{p}}_k$ is a unit vector that denotes the BF direction, and $q_k = \|\mathbf{p}_k\|^2$ is the power allocated to the kth data stream. The power allocation vector is given by $\mathbf{q} \triangleq [q_1, \ldots, q_K]^T$, while the matrix of BF directions writes as $\widehat{\mathbf{P}}_{\mathrm{p}} \triangleq \big[\widehat{\mathbf{p}}_1, \ldots, \widehat{\mathbf{p}}_K\big]$. It follows that the power constraint in (2.17) reduces to

$$\sum_{k=1}^{K} q_k \leq P_{\mathrm{t}}. \tag{2.22}$$

The precoder design, which includes adjusting the BF directions and the power allocation, is highly dependent on the considered design objective. For example, $\mathbf{P}_{\mathrm{p}}$ that yields the sum-rate maximizing rate tuple is generally different to the one that achievers a MMF performance amongst streams. This is discussed in more detail throughout the thesis. For now, we look at a simple and popular design of precoders.

2.4.2. Zero Forcing Beamforming

Zero Forcing (ZF)-BF vectors are those that satisfy the zero-interference condition expressed by $\mathbf{h}_k^H \mathbf{p}_i = 0$ for all $i, k \in \mathcal{K}$ and $i \neq k$. This is achieved by placing each BF vector in the null-space of all receivers it is not intended to, i.e. $\mathbf{p}_k \in \mathrm{null}\big(\bar{\mathbf{H}}_k^H\big)$, where

$$\bar{\mathbf{H}}_k \triangleq \big[\mathbf{h}_1, \ldots, \mathbf{h}_{k-1}, \mathbf{h}_{k+1}, \ldots, \mathbf{h}_K\big] \tag{2.23}$$

is the matrix consisting of all columns in $\mathbf{H}$ except for $\mathbf{h}_k$. Such design is feasible given that: 1) the transmitter knows $\mathbf{H}$ (perfect CSIT), 2) $N_{\mathrm{t}} \geq K$, and 3) $\mathbf{H}$ has full column-rank, which is guaranteed almost surely when the channel vectors are generic. It was shown in [32] that under the sum power constraints in (2.17), the optimum ZF-BF solution is based on the pseudo-inverse regardless of the performance measure to be maximized. In particular, the solution is given by

$$\mathbf{P}_{\mathrm{p}}^{\mathrm{ZF}}(\mathbf{H}, \mathbf{q}) = \big(\mathbf{H}^H\big)^{\dagger}\mathbf{B}\ \mathrm{diag}(\sqrt{q_1}, \ldots, \sqrt{q_K}) \tag{2.24}$$

where $\big(\mathbf{H}^H\big)^{\dagger} \triangleq \mathbf{H}\big(\mathbf{H}^H\mathbf{H}\big)^{-1}$ is the pseudo-inverse of $\mathbf{H}^H$, and $\mathbf{B}$ is a diagonal matrix given by $\mathrm{diag}\big(\sqrt{1/b_1}, \ldots, \sqrt{1/b_K}\big)$, with $b_1, \ldots, b_K$ being the diagonal entries of $\big((\mathbf{H}^H)^{\dagger}\big)^H(\mathbf{H}^H)^{\dagger}$. In words, the optimum $\widehat{\mathbf{P}}_{\mathrm{p}}$ is always given by the pseudo-inverse (with normalized columns), while the optimum $\mathbf{q}$ is determined by the design objective. The power allocation is easily computed due to the absence of inter-stream interference. For example, $\mathbf{q}$ that maximizes the sum-rate is obtained by the water-filling solution [32]. It should be noted that while the solution in (2.24) is optimum under zero-interference constraints, it is generally suboptimal

when such constraints are relaxed [5], making ZF-BF a suboptimal precoding strategy.

2.4.3. Degrees of Freedom

One of the virtues of linear precoding is that it achieves the optimum DoF of the MISO-BC under prefect CSIT. DoF analysis in linearly precoded systems is facilitated by introducing the notion of a *precoding scheme*, defined as a family a precoding matrices with one precoding matrix for each SNR level. This is denoted by $\{\mathbf{P}_{\mathrm{p}}(P_{\mathrm{t}})\}_{P_{\mathrm{t}}}$, where $\mathbf{P}_{\mathrm{p}}(P_{\mathrm{t}})$ complies with the power constraint in (2.17). Since Gaussian coding is assumed for data streams, the precoding scheme alongside its sequence of achievable rate tuples carries the necessary information about its corresponding coding scheme.

The optimality of linear precoding in a DoF sense can be easily shown using ZF-BF. Consider the design in (2.24) with equal power allocation, i.e. $q_1, \ldots, q_K = P_{\mathrm{t}}/K$. The kth rate is given by $R_k = \log_2\left(1 + \frac{P_{\mathrm{t}}}{K b_k \sigma_{\mathrm{n}}^2}\right)$. It can be easily seen that such rate scales as $\log_2(P_{\mathrm{t}}) + O(1)$, and the DoF tuple $[d_1, \ldots, d_K] = [1, \ldots, 1]$ is achievable. Hence, all users can receive interference free streams simultaneously. On important point to highlight here is that while this ZF-BF is DoF-optimum, it is not rate optimum in general as the DoF is a first order approximation that only captures the rate growth.

2.4.4. The Ergodic Regime (Fading Channel)

When transmission is carried out over B blocks, a precoder can be thought of as a sequence of precoding matrices $\{\mathbf{P}_{\mathrm{p}}[b]\}_{b=1}^{B}$, with one for each block. A given precoder is associated with a sequence of achievable rate tuples given by $\{[R_1[b], \ldots, R_K[b]]\}_{b=1}^{B}$, where $[R_1[b], \ldots, R_K[b]]$ is achieved using $\mathbf{P}_{\mathrm{p}}[b]$ in the bth block as described for a fixed channel. Achievable rates over the entire transmission period are given by the ensemble averages [17], for example the kth user achieves

$$\frac{1}{B}\sum_{b=1}^{B} R_k[b]. \tag{2.25}$$

As seen from ZF-BF, eliminating multiuser interference relies on instantaneous CSIT. Hence, designs used in non-ergodic transmission can be reused for ergodic transmission under perfect CSIT in an adaptive manner. In particular, $\mathbf{P}_{\mathrm{p}}[b]$, and (possibly) the code rates of the encoded streams, are adjusted from one block to another.

For sufficiently large B such that all fading states are experienced, the precoder can be written as $\{\mathbf{P}_{\mathrm{p}}(\mathbf{H})\}_{\mathbf{H}}$, with one precoding matrix for each $\mathbf{H}$. This is associated with $\{[R_1(\mathbf{H}), \ldots, R_K(\mathbf{H})]\}_{\mathbf{H}}$, where $[R_1(\mathbf{H}), \ldots, R_K(\mathbf{H})]$ is achieved using $\mathbf{P}_{\mathrm{p}}(\mathbf{H})$. Moreover, the achievable rate in (2.25) converges almost surely to the ER given by

$$R_k^{\mathrm{E}} \triangleq \mathbb{E}_{\mathrm{H}}\{R_k(\mathbf{H})\}. \tag{2.26}$$

The corresponding ER tuple is denoted by $[R_1^{\mathrm{E}}, \ldots, R_K^{\mathrm{E}}]$, achieved almost surely over a long sequence of blocks using a variable-rate adaptive-coding strategy, where coding is performed

independently for each block in the manner described for non-ergodic transmission. Interestingly, the same ergodic performance is achieved using a fixed-rate adaptive-precoding strategy. In particular, encoding is carried out at fixed rates that correspond to the ERs, while the adaptive component is restricted to the precoding matrix. The kth receiver sees the precoder as part of an equivalent scalar channel with instantaneous SINR given by $\gamma_k\big(\mathbf{H}[b]\big)$, which varies from one block to another. Hence, R_k^{E} is achievable using a fixed code of length $T = BT_{\mathrm{b}}$ given that B is sufficiently large [16, 17]. This strategy is favored in this work for reasons that will become clear in Chapter 3.

The nature of the transmission allows for the long-term power constraint given by

$$\frac{1}{B}\sum_{b=1}^{B}\mathrm{tr}\Big(\mathbf{P}_{\mathrm{p}}\big(\mathbf{H}[b]\big)\mathbf{P}_{\mathrm{p}}\big(\mathbf{H}[b]\big)^{H}\Big) \leq P_{\mathrm{t}} \tag{2.27}$$

When B is sufficiently large, this is equivalent to

$$\mathbb{E}_{\mathrm{H}}\Big\{\mathrm{tr}\big(\mathbf{P}_{\mathrm{p}}(\mathbf{H})\mathbf{P}_{\mathrm{p}}(\mathbf{H})^{H}\big)\Big\} \leq P_{\mathrm{t}} \tag{2.28}$$

taken over the distribution of $\mathbf{P}_{\mathrm{p}}(\mathbf{H})$, seen as a random variable which typically depends on $\mathbf{H}$ under perfect CSIT. The long-term constraint in (2.27) makes it possible to carry out inter-block (temporal or frequency) power allocation, in addition to the inter-stream (spatial) power allocation within each block. However, this couples all precoding matrices across various fading states, making design problems highly complicated. It is common to replace the long-term constraint with its short-term counterpart given in (2.17), where the BS should satisfy the power constraint for each fading state (or block) [18,33,34]. By doing so, the inter-block power allocation disappears and problems become more tractable.

Assumption 2.2. *A short-term precoder power constraint is assumed in this thesis.*

Finally, a precoding scheme that scales with SNR is given by $\big\{\mathbf{P}_{\mathrm{p}}(\mathbf{H}, P_{\mathrm{t}})\big\}_{\mathbf{H},P_{\mathrm{t}}}$, with one precoding matrix for each possible fading state and power level. It can be seen that an adaptive ZF-BF scheme achieves the full DoF in a fading MISO-BC under perfect CSIT.

2.5. Fundamentals of Rate-Splitting

In this section, we look at the concept of RS in the MISO-BC with partial CSIT. We start by examining the influence of CSIT imperfection on the ZF-BF scheme.

2.5.1. Influence of Imperfect CSIT

As seen in Section 1.1.2, imperfect CSIT can damage the achievable DoF in the MISO-BC. The analysis in Section 1.1.2 is revisited in the context of linear precoding, and different CSIT qualities across users. Consider a ZF-BF scheme with uniform power allocation

designed using partial CSIT such that

$$\mathbf{P}_{\mathrm{p}}\big(\widehat{\mathbf{H}}\big) = \mathbf{P}_{\mathrm{p}}^{\mathrm{ZF}}\big(\widehat{\mathbf{H}}, [P_{\mathrm{t}}/K, \ldots, P_{\mathrm{t}}/K]\big) \tag{2.29}$$

This is a special case of the ZF strategy seen in (1.7) realized through linear precoding. The signal received by the kth user in this case writes as

$$y_k = \overbrace{\mathbf{h}_k^H \mathbf{p}_k s_k}^{O(P_{\mathrm{t}})} + \overbrace{\sum_{i \neq k} \widetilde{\mathbf{h}}_k^H \mathbf{p}_i s_i}^{O(P_{\mathrm{t}}^{1-\alpha_k})} + \overbrace{n_k}^{O(P_{\mathrm{t}}^0)} \tag{2.30}$$

with the power scalings of different components highlighted. The imperfect ZF-BF yields residual interference scaling as $I_k = O(P_{\mathrm{t}}^{1-\alpha_k})$. It follows that the SINR of the kth users scales as $\gamma_k = O(P_{\mathrm{t}}^{\alpha_k})$, from which the achievable DoF is given by $d_k = \alpha_k$. Hence, the achievable DoF tuple is given by $[\alpha_1, \ldots, \alpha_K]$ from which the sum-DoF writes as

$$\sum_{k=1}^{K} d_k = \sum_{k=1}^{K} \alpha_k. \tag{2.31}$$

The loss in DoF compared to the prefect CSIT is observed by recalling Lemma 2.1 and the fact that $\alpha_k \in [0, 1]$. It can be seen that the DoF achieved by a given user depends on the corresponding CSIT quality. This is due to the fact that having a better estimate of a given user's channel enables better interference nulling in that direction.

Now, consider modifying the power allocation in (2.29) such that

$$\mathbf{P}_{\mathrm{p}}\big(\widehat{\mathbf{H}}\big) = \mathbf{P}_{\mathrm{p}}^{\mathrm{ZF}}\big(\widehat{\mathbf{H}}, [P_{\mathrm{t}}^{\alpha}/K, \ldots, P_{\mathrm{t}}^{\alpha}/K]\big) \tag{2.32}$$

where $\alpha = \max_k \alpha_k = \alpha_K$. It can be seen that the total power used in (2.32) is given by $P_{\mathrm{t}}^{\alpha} \leq P_{\mathrm{t}}$, where equality holds for $\alpha = 1$. The received signal in this case writes as

$$y_k = \overbrace{\mathbf{h}_k^H \mathbf{p}_k}^{O(P_{\mathrm{t}}^{\alpha})} + \overbrace{\sum_{i \neq k} \widetilde{\mathbf{h}}_k^H \mathbf{p}_i}^{O(P_{\mathrm{t}}^{\alpha-\alpha_k})} + \overbrace{n_k}^{O(P_{\mathrm{t}}^0)}. \tag{2.33}$$

The SINR still scales as $\gamma_k = O(P_{\mathrm{t}}^{\alpha_k})$, and $d_k = \alpha_k$ is achieved. It can be seen that the sum-DoF in (2.31) is maintained using a potentially lower power scaling. In particular, the transmitted power scaling is such that the interference received by the user with the highest CSIT quality is at the same level of noise. In the special case where $\alpha_1 = \ldots = \alpha_K = \alpha$, the residual interference is received at the level of noise by all users as seen in (1.12).

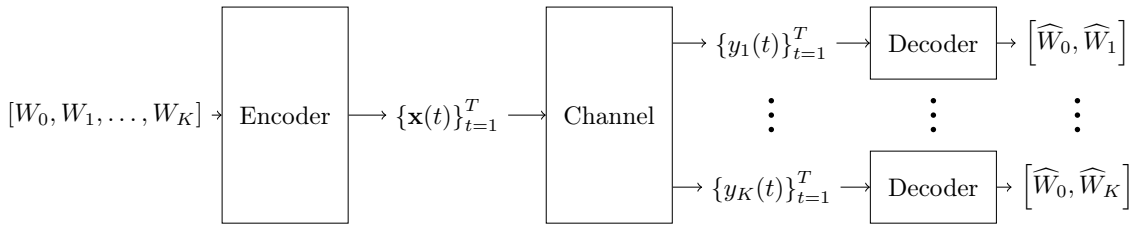


Figure 2.2.: K-user Broadcast Channel with a common message.

2.5.2. Broadcast Channel with a Common Message

Consider the BC with a common message shown in Figure 2.2. This class of BCs plays a vital role in constructing the RS scheme. A code for such channel is described as:

1. $K+1$ message sets $\mathcal{W}_{\mathrm{c}}, \mathcal{W}_1, \ldots, \mathcal{W}_K$.
2. A sequence of encoding functions $\left\{\phi(t)\right\}_{t=1}^{T}$, where $\phi(t)$ maps any CSIT and message tuple $[W_0, W_1, \ldots, W_K]$ into the input signal $\mathbf{x}(t)$.
3. K *decoding* functions $\psi_1, \ldots, \psi_K$, such that ψ_k maps received sequences $\left\{y_k(t)\right\}_{t=1}^{T}$ into a pair of decoded messages $[\widehat{W}_0, \widehat{W}_k] \in \mathcal{W}_0 \times \mathcal{W}_k$.

In a MISO-BC with linear precoding, the $K+1$ messages are first encoded into symbol streams as $W_{\mathrm{c}} \mapsto \left\{s_{\mathrm{c}}(t)\right\}_{t=1}^{T}$ and $W_k \mapsto \left\{s_k(t)\right\}_{t=1}^{T}$ for all $k \in \mathcal{K}$. This is followed by linear precoding, where the precoding matrix is given by $\mathbf{P} = [\mathbf{p}_{\mathrm{c}}, \mathbf{p}_1, \ldots, \mathbf{p}_K]$, with $\mathbf{p}_{\mathrm{c}} \in \mathbb{C}^{N_{\mathrm{t}}}$ as the common stream's precoding vector. The transmitted signal is expressed by

$$\mathbf{x} = \mathbf{P}\mathbf{s} = \mathbf{p}_{\mathrm{c}} s_{\mathrm{c}} + \sum_{k=1}^{K} \mathbf{p}_k s_k \tag{2.34}$$

where $\mathbf{s} \triangleq [s_{\mathrm{c}}, s_1, \ldots, s_K]^T$ consists of the $K+1$ data symbols. This transmission is clearly a superposition of multicast beamforming [6] and the unicast beamforming in (2.16). In an ergodic scenario, a sequence of precoding matrices is denoted by $\left\{\mathbf{P}(\mathbf{H})\right\}_{\mathbf{H}}$, with one precoding matrix for each fading state. The short-term power constraint writes as

$$\mathrm{tr}\big(\mathbf{P}\mathbf{P}^H\big) \leq P_{\mathrm{t}} \tag{2.35}$$

The signal received by the kth user in a given channel use is expressed as

$$y_k = \overbrace{\mathbf{h}_k^H \mathbf{p}_{\mathrm{c}} s_{\mathrm{c}}}^{\text{common signal}} + \overbrace{\mathbf{h}_k^H \mathbf{p}_k s_k}^{\text{desired private signal}} + \overbrace{\mathbf{h}_k^H \sum_{i \neq k} \mathbf{p}_i s_i}^{\text{interference}} + n_k \tag{2.36}$$

from which the average received power is given by

$$T_{\mathrm{c},k} \triangleq \mathrm{E}\{|y_k|^2\} = \overbrace{|\mathbf{h}_k^H \mathbf{p}_\mathrm{c}|^2}^{S_{\mathrm{c},k}} + \underbrace{\overbrace{|\mathbf{h}_k^H \mathbf{p}_k|^2}^{S_k} + \overbrace{\sum_{i\neq k} |\mathbf{h}_k^H \mathbf{p}_i|^2 + \sigma_\mathrm{n}^2}^{I_k}}_{I_{\mathrm{c},k}=T_k} \tag{2.37}$$

where $S_{\mathrm{c},k}$ and $I_{\mathrm{c},k}$ are the common stream's desired signal power and the interference plus noise power respectively. Recall that the common message is intended to, and decoded by, all receivers. Hence, Successive Interference Cancellation (SIC) can be used to remove he common signal from the received signals to enhance the decodability of the private messages. This is followed by decoding the private message in the presence of the remaining interfering signals and noise. After SIC, S_k and I_k denote the private stream's desired signal power and the interference plus noise power respectively.

Next, we look at the achievable rates and DoF in an ergodic setting. The same results apply to non-ergodic transmission by considering the instantaneous rates. The instantaneous SINR of the common stream at the kth receiver is given by

$$\gamma_{\mathrm{c},k} \triangleq \frac{S_{\mathrm{c},k}}{I_{\mathrm{c},k}} = \frac{\left|\mathbf{h}_k^H \mathbf{p}_\mathrm{c}\right|^2}{\sum_{i=1}^K \left|\mathbf{h}_k^H \mathbf{p}_i\right|^2 + \sigma_\mathrm{n}^2} \tag{2.38}$$

from which the instantaneous rate of the common message from the kth user's point of view writes as

$$R_{\mathrm{c},k} = \log_2(1+\gamma_{\mathrm{c},k}). \tag{2.39}$$

In an ergodic setting, transmitting W_c at a rate of $R_{\mathrm{c},k}^\mathrm{E} \triangleq \mathbb{E}_\mathrm{H}\{R_{\mathrm{c},k}\}$ guarantees that it is decodable by the kth user. To guarantee that W_c is successfully decoded by all users, it should be transmitted at the common ER given by

$$R_\mathrm{c}^\mathrm{E} \triangleq \min_j \left\{ \mathbb{E}_\mathrm{H}\{R_{\mathrm{c},k}\} \right\}_{j=1}^K. \tag{2.40}$$

On the other hand, as SIC is employed, the private instantaneous SINRs and rates are given by (2.20) and (2.21) respectively, while the ERs are given by (2.26).

It is clear that an ER tuple for such setup consists of $K+1$ rates and writes as $\left[R_\mathrm{c}^\mathrm{E}, R_1^\mathrm{E}, \ldots, R_K^\mathrm{E}\right]$. Similarly, a DoF tuple is denoted by $[d_\mathrm{c}, d_1, \ldots, d_K]$, where d_c denotes the DoF of the common message. An achievable DoF tuple is given next.

Lemma 2.2. *For the MISO-BC with a common message under linear precoding and scaling CSIT, the DoF tuple given by* $[d_\mathrm{c}, d_1, \ldots, d_K] = [1-\alpha_K, \alpha_1, \ldots, \alpha_K]$ *is achievable.*

This result is given for $K=2$ and equal CSIT qualities in [2, Lemma 2]. The extension to arbitrary K and CSIT qualities is straightforward. The achievability scheme is described as follows. For given $\widehat{\mathbf{H}}$ and P_t, the private precoders are given by (2.32). On the other hand, the common precoder is generated randomly with power $q_\mathrm{c} = P_\mathrm{t} - P_\mathrm{t}^\alpha$. The achievable DoF

follows by observing the power scaling of different received signal components given by

$$y_k = \overbrace{\mathbf{h}_k^H \mathbf{p}_{\mathrm{c}}}^{O(P_{\mathrm{t}})} + \overbrace{\mathbf{h}_k^H \mathbf{p}_k}^{O(P_{\mathrm{t}}^{\alpha})} + \overbrace{\sum_{i \neq k} \widetilde{\mathbf{h}}_k^H \mathbf{p}_i}^{O(P_{\mathrm{t}}^{\alpha - \alpha_k})} + \overbrace{n_k}^{O(P_{\mathrm{t}}^{0})}. \tag{2.41}$$

We have $\gamma_{\mathrm{c},k} = O(P_{\mathrm{t}}^{1-\alpha})$ and $\gamma_k = O(P_{\mathrm{t}}^{\alpha_k})$, from which the DoF in Lemma (2.2) follows.

2.5.3. Rate-Splitting

From Lemma 2.2, the MISO-BC with a common message achieves a sum-DoF of

$$d_{\mathrm{c}} + \sum_{k=1}^{K} d_K = 1 + \sum_{k=1}^{K-1} \alpha_k. \tag{2.42}$$

By comparing this to the sum-DoF in (2.31), it can be seen that the incorporation of the common message achieves a DoF gain of $1 - \alpha_K$. This follows from the fact that the maximum DoF achieved through ZF-BF is maintained using a fraction of the total power as seen in (2.32), while the rest of the power is utilized to achieve extra DoF through the common message. This gives rise to the following question: could this gain be utilized in a MISO-BC with only private messages? This is answered in the affirmative through the RS strategy as seen in Section 1.1.3. This is revisited here in the context of a linear precoding and different CSIT qualities across users. RS transmission is formulated as follows.

Assume that the message of the Kth user is split into two parts, a common part W_{K0} and a private part W_{K1}. The common part is encoded into a common data stream such that $W_{K0} \mapsto \left\{s_{\mathrm{c}}(t)\right\}_{t=1}^{T}$, while the private part alongside other messages are encoded into private streams such that $W_1, W_2, \ldots, W_{K1} \mapsto \left\{s_1(t)\right\}_{t=1}^{T}, \left\{s_2(t)\right\}_{t=1}^{T}, \ldots, \left\{s_K(t)\right\}_{t=1}^{T}$. The resulting $K+1$ data streams are precoded and transmitted as shown in (2.34).

Each receiver decodes W_{K0}, removes the corresponding part from the received signal using SIC, before decoding its private stream. It follows from Lemma 2.2 that the common stream achieves a DoF of $1 - \alpha_K$ and private streams achieve $\alpha_1, \ldots, \alpha_K$. Note that while the common stream is decoded by all users in the RS scheme, it is only indented to user-K, in contrast to the MISO-BC with a common message. Hence, the RS scheme achieves the DoF tuple $[\alpha_1, \ldots, \alpha_{K-1}, 1]$, from which the sum-Dof is given by (2.42). Note that when all users experience the same CSIT quality, the sum-DoF becomes $1 + (K-1)\alpha$.

While the splitting was restricted to the Kth user here, it may be extended to other users. This is required particularly when the objective is to achieve fairness. In this case, different users share the benefits of the common message as seen in Chapter 4.

2.6. Conclusion

This chapter provided the main background necessary for the remainder of this thesis. A general MISO-BC signal model was described, which will be used in the following chapters. The main assumptions regarding the CSIT uncertainty and the coding and preprocessing strategy were given. We also introduced the performance limits and measures used throughout this thesis. Finally, the RS strategy was revisited and its sum-DoF gain over conventional ZF-BF transmission was highlighted.

3. Sum-Rate Maximization

RS was shown to boost the sum-DoF of the MISO-BC under CSIT errors that decay with increased SNR. Such gains were demonstrated assuming a simple design of precoders where ZF-BF was used for the private streams. However, ZF-BF is known to be suboptimal, particularly for finite SNRs. Moreover, simple non-optimized precoders are usually used for the common stream in DoF-motivated designs and analysis.

In this chapter, the RS strategy is married with precoder design and optimization techniques to achieve a maximized Ergodic Sum-Rate (ESR) performance over the entire range of SNRs. Precoders are designed using partial CSIT by solving a stochastic rate optimization problem using two robust algorithms based on the Weighted Minimum Mean Square Error (WMMSE) approach. The first algorithm uses a conservative approximation to tackle the stochastic nature of the problem, while the second algorithm employs means of Sample Average Approximation (SAA). Before going into details, we start by formulating the ESR maximization problem. This is followed by an investigation of the general performance limits of the formulated problems through DoF analysis.

3.1. The Precoder Optimization Problem

In this chapter, we consider a MU-MISO system operating in downlink with the signal model described in (2.1). The BS wishes to communicate one private message to each user, i.e. a standard MISO-BC. It is assumed that the number of users scheduled during a transmission period may not exceed the number of transmit antennas, i.e. $K \leq N_t$.

We saw in Chapter 2 that linear precoding is an efficient coding strategy for MU-MISO transmission. One popular choice of linear precoders is ZF-BF, shown in Section 2.4.2. While ZF-BF is optimum in the DoF sense under perfect CSIT, it is generally a suboptimal precoding strategy [5]. Optimum precoders are ones that strike a delicate balance between nulling multiuser interference and maximizing the desired signal power at each receiver.

For a fixed channel $\mathbf{H}$ and under perfect CSIT, the precoding matrix that maximizes the Sum-Rate (SR) under a power constraint P_{t} is obtained by solving

$$\begin{aligned} \max_{\mathbf{P}_{\mathrm{p}}} \quad & \sum_{k=1}^{K} R_k \\ \text{s.t.} \quad & \mathrm{tr}\big(\mathbf{P}_{\mathrm{p}}\mathbf{P}_{\mathrm{p}}^{H}\big) \leq P_{\mathrm{t}}. \end{aligned} \tag{3.1}$$

This optimization problem is non-convex and known to be very challenging to solve in its

current form. In fact, it was shown to be NP-hard in [35]. The optimum solution can be obtained using a branch-reduce-and-bound algorithm [36], with a huge computational complexity that grows exponentially with K. Alternatively, a number of polynomial time methods that obtain suboptimal solutions have been proposed in [7, 37–39].

The WMMSE algorithm proposed in [7] is arguably the most popular method due to its efficiency and effectiveness. First, the domain of the original problem in (3.1) is extended by introducing new optimization variables, namely receive filters and MSE weights. Then, the extended WMSE problem is solved using an Alternating Optimization (AO) algorithm, also known as the Block Coordinate Descent (BCD) method [40]. The convergence of the algorithm to a stationary solution of the original SR problem in (3.1) was rigorously established in [8]. This method is widely used throughout this thesis.

3.1.1. The Ergodic Regime

For ergodic transmission under perfect CSIT, the short-term power constraint in Assumption 2.2 decouples $\left\{\mathbf{P}_{\mathrm{p}}(\mathbf{H})\right\}_{\mathbf{H}}$ across fading states. The problem writes as

$$\mathbb{E}_{\mathrm{H}}\left\{\max_{\mathrm{tr}\left(\mathbf{P}_{\mathrm{p}}(\mathbf{H})\mathbf{P}_{\mathrm{p}}(\mathbf{H})^{H}\right)\leq P_{\mathrm{t}}}\sum_{k=1}^{K}R_k(\mathbf{H})\right\}. \tag{3.2}$$

The optimum solution can be obtained by solving a sequence of problem (3.1), where $\mathbf{P}_{\mathrm{p}}$ is optimized separately for each $\mathbf{H}$. In a transmission taking place over a sequence of blocks with $\left\{\mathbf{H}[b]\right\}_{b=1}^{B}$, $\mathbf{P}_{\mathrm{p}}[b]$ is updated in each block according to $\mathbf{H}[b]$. This adaptation of the precoding matrix is enabled by the availability of the instantaneous CSIT in each block.

3.1.2. Robustness Under Imperfect CSIT

Now what if the BS has imperfect CSIT as described in Section 2.2.1? A naive approach would be to solve the problem assuming that $\mathbf{H}=\widehat{\mathbf{H}}$, i.e. perfect CSIT. This design is not only unable to cope with the resulting multiuser interference, it is also unaware of it. In addition, this may lead to an overestimation of the instantaneous and ergodic rates by the BS, yielding transmission at undecodable rates. A robust approach on the other hand employs the available CSIT knowledge to:

1. Design an informed precoding scheme that enhances the instantaneous channel condition experienced by the receivers.

2. Perform transmission at reliable rates and hence guarantee decodability.

The two points are related, as the characterization of the reliable transmission rates enables us to formulate, and ultimately solve, the robust precoder design problem.

In the non-ergodic regime, $\mathbf{H}$ and $\widehat{\mathbf{H}}$ do not change, i.e. transmission is carried out over one joint fading state $\left\{\mathbf{H},\widehat{\mathbf{H}}\right\}$. The BS has no access to the exact instantaneous

rates experienced by the receivers given in (2.21). Hence, the BS may only guarantee reliable communication with a certain outage derived from the CSIT statistics [41, 42]. Precoders would be designed such that the rate of communication is maximized without exceeding a given outage probability [43]. If the uncertainty region is bounded, zero-outage communication can be guaranteed [44, 45]. On the other hand, if the transmission is delay-unlimited, the ergodic nature of the fading process and knowledge of its long-term properties can be exploited to guarantee communication at reliable rates as we see in what follows. The question that arises is, how can we utilize the imperfect instantaneous CSIT to maximize the long-term (ergodic) performance?

3.1.3. Average Sum-Rate Problem

To answer the previous question, let us define an adaptive precoder under imperfect CSIT. This is given by $\{\mathbf{P}_{\mathrm{p}}(\widehat{\mathbf{H}})\}_{\widehat{\mathbf{H}}}$ with one precoding matrix for each channel state estimate, as the precoder is adapted by the BS. In a transmission taking place over a sequence of blocks with joint channel states $\{\mathbf{H}[b], \widehat{\mathbf{H}}[b]\}_{b=1}^{B}$, $\mathbf{P}_{\mathrm{p}}[b]$ is updated in each block according to $\widehat{\mathbf{H}}[b]$. The instantaneous rates achieved in each block however depend on $\mathbf{H}$ as seen from (2.21). There is also an implicit dependency of the rates on $\widehat{\mathbf{H}}$ now, which follows from the dependency on $\mathbf{P}_{\mathrm{p}}(\widehat{\mathbf{H}})$. While the BS is unable to predict the instantaneous rates, it has access to the Average Rates (ARs) defined as follows.

Definition 3.1. *For a given channel estimate $\widehat{\mathbf{H}}$ and precoding matrix $\mathbf{P}_{\mathrm{p}}(\widehat{\mathbf{H}})$, the kth user's AR is defined as*

$$\bar{R}_k \triangleq \mathbb{E}_{\mathrm{H}|\widehat{\mathrm{H}}}\left\{R_k(\mathbf{H}) \mid \widehat{\mathbf{H}}\right\}. \tag{3.3}$$

The AR in (3.3) and the ER in (2.26) should not be confused: while the latter describes the long-term performance over all fading states and precoding matrices, the former is a short-term (instantaneous) measure that captures the expected performance over the CSIT uncertainty given an estimate and precoding matrix.

It turns out that the ERs can be characterized by averaging the ARs over the variation in $\widehat{\mathbf{H}}$. In particular, the ER experienced by the kth user is expressed by

$$\mathbb{E}_{\{\mathrm{H},\widehat{\mathrm{H}}\}}\left\{R_k\big(\mathbf{H},\widehat{\mathbf{H}}\big)\right\} = \mathbb{E}_{\widehat{\mathrm{H}}}\left\{\mathbb{E}_{\mathrm{H}|\widehat{\mathrm{H}}}\left\{R_k\big(\mathbf{H},\widehat{\mathbf{H}}\big) \mid \widehat{\mathbf{H}}\right\}\right\} = \mathbb{E}_{\widehat{\mathrm{H}}}\left\{\bar{R}_k\big(\widehat{\mathbf{H}}\big)\right\} \tag{3.4}$$

which follows from the law of total expectation. Hence, the ESR maximization problem under imperfect instantaneous CSIT writes as

$$\mathbb{E}_{\widehat{\mathrm{H}}}\left\{\max_{\mathrm{tr}\left(\mathbf{P}_{\mathrm{p}}(\widehat{\mathbf{H}})\mathbf{P}_{\mathrm{p}}(\widehat{\mathbf{H}})^{H}\right)\leq P_{\mathrm{t}}} \sum_{k=1}^{K} \bar{R}_k\big(\widehat{\mathbf{H}}\big)\right\} \tag{3.5}$$

where for each channel estimate, the precoders are optimized such that the Average SR (ASR) is maximized. For a class of information theoretic channels under imperfect CSIT with a Markov property, the ergodic capacity is formulated following the same philosophy,

while replacing the ASR with the average mutual-information and maximizing over the input distribution rather than the precoding matrix [16].

By focusing on one channel state estimate, the ASR optimization problem writes as

$$\mathcal{R}(P_{\mathrm{t}}) : \begin{cases} \max\limits_{\mathbf{P}_{\mathrm{p}}} & \sum_{k=1}^{K} \bar{R}_k \\ \text{s.t.} & \mathrm{tr}\big(\mathbf{P}_{\mathrm{p}}\mathbf{P}_{\mathrm{p}}^{H}\big) \leq P_{\mathrm{t}} \end{cases} \tag{3.6}$$

where the dependency of $\bar{R}_k$ and $\mathbf{P}_{\mathrm{p}}$ on a particular $\widehat{\mathbf{H}}$ is removed from the notation for brevity. The maximized ESR given by $\mathbb{E}_{\widehat{\mathrm{H}}}\{\mathcal{R}(P_{\mathrm{t}})\}$, which also corresponds to (3.5), is achievable using the fixed-rate/adaptive-precoding strategy described in Section 2.4.4. In particular, the encoding is carried out at the fixed ERs given by $\mathbb{E}_{\widehat{\mathrm{H}}}\{\bar{R}_1\}, \ldots, \mathbb{E}_{\widehat{\mathrm{H}}}\{\bar{R}_K\}$, which only depend on the CSIT statistics. On the other hand, precoders are optimized and adapted in each block according to the instantaneous CSIT.

Remark 3.1. *Under perfect CSIT, the AR in* (3.3) *reduces to the instantaneous rate in* (2.21)*. It follows the ASR problem in* (3.6) *reduces to the SR problem in* (3.1)*, and the ESR problem in* (3.5) *reduces to the one in* (3.2)*.*

In [46,47], the authors considered a fading model where $\widehat{\mathbf{H}}$ remains fixed over all channel states, representing a Line-of-Sight (LoS) component. In such scenario, the ASR problem in (3.6) and the ESR problem in (3.5) are equivalent. The model considered here is more general as it allows for a changing $\widehat{\mathbf{H}}$. On the other hand, [48] considered a fast-fading model where $\widehat{\mathbf{H}}$ is completely suppressed. In this case, which is another special case of the general fading model considered here, the ASR and ESR problems also coincide.

3.1.4. The Rate-Splitting Problem Formulation

Now we turn to ESR maximization for the RS strategy described in Section 2.5.3. First, we denote a RS precoder with imperfect CSIT as $\big\{\mathbf{P}(\widehat{\mathbf{H}})\big\}_{\widehat{\mathbf{H}}}$ with one precoding matrix for each channel state estimate. We start with defining the common ARs.

Definition 3.2. *For a given channel estimate $\widehat{\mathbf{H}}$ and precoding matrix $\mathbf{P}_{\mathrm{p}}(\widehat{\mathbf{H}})$, the kth user's common AR is defined as*

$$\bar{R}_{\mathrm{c},k} \triangleq \mathbb{E}_{\mathrm{H}|\widehat{\mathrm{H}}}\left\{R_{\mathrm{c},k}\big(\mathbf{H}\big) \mid \widehat{\mathbf{H}}\right\}. \tag{3.7}$$

As in (3.4), the kth user's common ER is characterized by averaging the corresponding common AR over the variation in $\widehat{\mathbf{H}}$ such that

$$\mathbb{E}_{\{\mathrm{H},\widehat{\mathrm{H}}\}}\left\{R_{\mathrm{c},k}\big(\mathbf{H},\widehat{\mathbf{H}}\big)\right\} = \mathbb{E}_{\widehat{\mathrm{H}}}\left\{\bar{R}_{\mathrm{c},k}\big(\widehat{\mathbf{H}}\big)\right\} \tag{3.8}$$

It follows that transmitting the common message at the ER given in (3.8) guarantees its decodability by the kth user. To guarantee decodability by all users, the common message

should be transmitted at the common ER given by: $\min_{j\in\mathcal{K}} \mathbb{E}_{\widehat{\mathrm{H}}}\left\{\bar{R}_{\mathrm{c},j}\big(\widehat{\mathbf{H}}\big)\right\}$. It follows that the RS ESR problem is formulated as

$$\max_{\mathrm{tr}\left(\mathbf{P}(\widehat{\mathbf{H}})\mathbf{P}(\widehat{\mathbf{H}})^H\right)\leq P_{\mathrm{t}},\forall\widehat{\mathbf{H}}} \ \min_{j\in\mathcal{K}} \mathbb{E}_{\widehat{\mathrm{H}}}\left\{\bar{R}_{\mathrm{c},j}\big(\widehat{\mathbf{H}}\big)\right\} + \sum_{k=1}^{K} \mathbb{E}_{\widehat{\mathrm{H}}}\left\{\bar{R}_{k}\big(\widehat{\mathbf{H}}\big)\right\} \tag{3.9}$$

Comparing the NoRS problem in (3.5) to its RS counterpart in (3.9), it can be seen that the latter poses an extra challenge. In particular, the fact that the pointwise minimum of the common ERs is outside the expectation couples the precoder design across fading states, making it no longer possible to carry out the optimization separately for each $\widehat{\mathbf{H}}$. To eliminate this coupling, we employ the following lower-bound

$$\min_{j\in\mathcal{K}}\left\{\mathbb{E}_{\widehat{\mathrm{H}}}\left\{\bar{R}_{\mathrm{c},j}\big(\widehat{\mathbf{H}}\big)\right\}\right\} \geq \mathbb{E}_{\widehat{\mathrm{H}}}\left\{\min_{j\in\mathcal{K}} \bar{R}_{\mathrm{c},j}\big(\widehat{\mathbf{H}}\big)\right\} \tag{3.10}$$

where the inequality is from the fact that moving the minimization inside the expectation does not increase the value. Now we can write the following ESR lower-bound problem

$$\mathbb{E}_{\widehat{\mathrm{H}}}\left\{\max_{\mathrm{tr}\left(\mathbf{P}_{\mathrm{p}}(\widehat{\mathbf{H}})\mathbf{P}_{\mathrm{p}}(\widehat{\mathbf{H}})^H\right)\leq P_{\mathrm{t}}} \ \min_{j\in\mathcal{K}} \bar{R}_{\mathrm{c},j}\big(\widehat{\mathbf{H}}\big) + \sum_{k=1}^{K} \bar{R}_{k}\big(\widehat{\mathbf{H}}\big)\right\}. \tag{3.11}$$

Now ASR maximization can be carried out separately for each $\widehat{\mathbf{H}}$ by considering the problem inside the expectation in (3.11) expressed by

$$\mathcal{R}_{\mathrm{RS}}(P_{\mathrm{t}}): \begin{cases} \max\limits_{\bar{R}_c,\mathbf{P}} & \bar{R}_{\mathrm{c}} + \sum_{k=1}^{K} \bar{R}_k \\ \text{s.t.} & \bar{R}_{\mathrm{c},k} \geq \bar{R}_{\mathrm{c}},\ \forall k \in \mathcal{K} \\ & \mathrm{tr}\big(\mathbf{P}\mathbf{P}^H\big) \leq P_{\mathrm{t}} \end{cases} \tag{3.12}$$

where the dependency on $\widehat{\mathbf{H}}$ is omitted for brevity. The pointwise minimization of the common ARs is replaced with the set of inequalities involving the common AR auxiliary variable denoted by $\bar{R}_{\mathrm{c}}$. It is easy to see that at optimality, we have $\bar{R}_{\mathrm{c}} = \min_j \left\{\bar{R}_{\mathrm{c},j}\right\}_{j=1}^{K}$.

The ESR given in (3.11), also denoted by $\mathbb{E}_{\widehat{\mathrm{H}}}\left\{\mathcal{R}_{\mathrm{RS}}(P_{\mathrm{t}})\right\}$, is achievable using the fixed-rate/adaptive-precoding strategy. The achievability of the private ER is similar to the previous subsection, while the common ER is achievable from (3.10).

3.2. Performance Limits

After formulating the NoRS and RS instantaneous ASR problems in the previous section, the next two questions that come to mind are:

1. How does the instantaneous CSIT quality influence the long-term performances given by $\mathbb{E}_{\widehat{\mathrm{H}}}\left\{\mathcal{R}(P_{\mathrm{t}})\right\}$ and $\mathbb{E}_{\widehat{\mathrm{H}}}\left\{\mathcal{R}_{\mathrm{RS}}(P_{\mathrm{t}})\right\}$?

2. How do the NoRS and RS performances compare?

Such questions can be partially answered through DoF analysis. Before we proceed, we recall that the CSIT uncertainty model is as described in Section 2.2.1. Moreover, we assume throughout this chapter that the CSIT qualities are equal across users, i.e. $\alpha_1 = \alpha_2 = \ldots = \alpha_K = \alpha$. While this assumption is made to simplify the analysis, it is not necessary for the design algorithms proposed later in the chapter.

3.2.1. DoF Analysis

The DoFs for the optimized NoRS and RS strategies are defined as

$$d^* \triangleq \lim_{P_{\mathrm{t}}\to\infty} \frac{\mathbb{E}_{\widehat{\mathrm{H}}}\{\mathcal{R}(P_{\mathrm{t}})\}}{\log_2(P_{\mathrm{t}})} \quad \text{and} \quad d^*_{\mathrm{RS}} \triangleq \lim_{P_{\mathrm{t}}\to\infty} \frac{\mathbb{E}_{\widehat{\mathrm{H}}}\{\mathcal{R}_{\mathrm{RS}}(P_{\mathrm{t}})\}}{\log_2(P_{\mathrm{t}})}. \tag{3.13}$$

As highlighted before, the DoF is roughly interpreted as the total number of interference-free streams that can be simultaneously supported in a single channel use. The significance of the DoF in such setups comes from the detrimental effects multiuser interference may have, and the fundamental role of the CSIT in dealing with such effects.

Insight into the DoFs of the optimized NoRS and RS schemes can be gained from the discussion in Section 2.5. In particular, the ZF-BF based precoding schemes there are feasible for the problems at hand. Hence, the DoFs in (2.31) and (2.42) should be achievable by the optimized schemes. This indeed plays a role in the proof of the next result. First, we make the following assumption.

Assumption 3.1. *We assume that the channel estimate $\widehat{\mathbf{H}}$ is of full column rank with probability one. Moreover, for any given estimate, we assume isotropically distributed CSIT error vectors, i.e. we have $\mathbf{R}_{\mathrm{e},k} = \frac{\sigma^2_{\mathrm{e},k}}{N_{\mathrm{t}}}\mathbf{I}$ for all $k \in \mathcal{K}$.*

Note that these assumptions are made for the proof of the next result, and they are not necessary for the optimization in the following sections.

Theorem 3.1. *The DoF achieved by optimally solving the NoRS problem is given by*

$$d^* \triangleq \lim_{P_{\mathrm{t}}\to\infty} \frac{\mathbb{E}_{\widehat{\mathrm{H}}}\{\mathcal{R}(P_{\mathrm{t}})\}}{\log_2(P_{\mathrm{t}})} = \max\{1, K\alpha\} \tag{3.14}$$

while the DoF of an optimally designed RS scheme is given by

$$d^*_{\mathrm{RS}} \triangleq \lim_{P_{\mathrm{t}}\to\infty} \frac{\mathbb{E}_{\widehat{\mathrm{H}}}\{\mathcal{R}_{\mathrm{RS}}(P_{\mathrm{t}})\}}{\log_2(P_{\mathrm{t}})} = 1 + (K-1)\alpha. \tag{3.15}$$

Results of this nature, which characterize the optimum DoF performance of a given transmission strategy, appear in each of the three core chapters in this thesis. Such results are obtained through two steps: converse and achievability. Take the NoRS result in (3.14) for example. The converse shows that for any feasible precoding schemes, the sum-DoF for

the NoRS strategy is upper-bounded by the one given in (3.14). The achievability on the other hand shows that this upper-bound is in fact attainable through a feasible precoding scheme. The full proof is relegated to Appendix A.1.

3.2.2. Insight

When $\alpha = 1$, the maximum attainable DoF is achieved, i.e. K. Although CSIT may still be erroneous, it is sufficient to design private precoders that reduce the multiuser interference to the level of additive noise. However, this is not possible when α drops below 1, as private streams start leaking substantial interference. The NoRS scheme exhibits loss in DoF until $\alpha = 1/K$. As α drops below $1/K$, the CSIT quality is not good enough to support multiuser transmission, and it is preferred from a DoF perspective to switch to single-user transmission by allocating all power to one stream, hence achieving a DoF of 1. In naive designs as the one seen in Section 2.5.1, the DoF reduces to zero for $\alpha = 0$. For such designs, gains from switching to single-user transmission are not realized due to the fixed and uniform power allocation across users [18, 19].

Now looking at RS, as soon as α drops below 1, a DoF-optimum design carefully allocates powers to the private streams such that multiuser interference is always received at the level of additive noise. This is achieved by scaling down private powers to $O(P_{\mathrm{t}}^{\alpha})$, hence maintaining a DoF of $K\alpha$. The rest of the power, which scales as $O(P_{\mathrm{t}})$, is allocated to the common stream which is broadcasted to all users. This achieves a DoF gain of $1 - \alpha$, as interference from private messages is treated as noise. It should be noted that the RS DoF is strictly greater than the NoRS DoF for all $\alpha \in (0, 1)$.

The RS DoF in (3.15) is inline with the results in [2, 3, 25] (also in Section 2.5.3). Although optimizing the precoders does not improve the achievable DoF, simulation results presented later in this chapter show that the optimized scheme is superior from a rate performance perspective. It is worth highlighting that the converse result reported in [20] proves the optimality of the DoF in (3.15) for the real Gaussian MISO-BC with scaling CSIT uncertainty where one of the users has perfect CSIT. In case this can be extended to complex channels, (3.15) follows by combining with the achievability in Section 2.5.3.

We conclude this section by highlighting that in addition to the high SNR gains manifested through the DoF in Theorem 3.1, $\mathcal{R}_{\mathrm{RS}}(P_{\mathrm{t}}) \geq \mathcal{R}(P_{\mathrm{t}})$ is guaranteed over the entire range of SNRs. This is seen by noting that solving (3.6) is equivalent to solving (3.12) over a subset of its domain characterized by restricting $\|\mathbf{p}_{\mathrm{c}}\|^2$ to zero. In the following, we focus on solving the RS design problem, i.e. solving the ASR problem in (3.12).

3.3. The WMMSE Approach

This approach reformulates SR problems into equivalent, and more tractable, WMMSE problems [7, 8], which differ from their classical MMSE counterparts [49, 50] as:

1. Weights are introduced and considered as optimization variables.

2. The cost function is augmented to incorporate the logarithms of the weights.

A Rate-WMMSE relationship is built upon these two features, enabling the formulation of the equivalent problem. Next, equalization and MSE estimation are introduced into the signal model, paving the way for the WMMSE approach.

3.3.1. MSEs and MMSE Receivers

At the kth receiver, let $\widehat{s}_{\mathrm{c},k} = g_{\mathrm{c},k} y_k$ be an estimate of the common symbol $s_{\mathrm{c},k}$ obtained by applying the scalar equalizer $g_{\mathrm{c},k}$. Given that the common message is transmitted at an adequate rate and successfully decoded by all receivers, an estimate of s_k is obtained after removing the common stream from the received signal such that $\widehat{s}_k = g_k(y_k - \mathbf{h}_k^H \mathbf{p}_{\mathrm{c}} s_{\mathrm{c},k})$, where g_k is the corresponding equalizer. At the output of the kth receiver, the common and private MSEs are defined as:

$$\varepsilon_{\mathrm{c},k} \triangleq \mathbb{E}\left\{ |\widehat{s}_{\mathrm{c},k} - s_{\mathrm{c}}|^2 \right\} \tag{3.16a}$$

$$\varepsilon_k \triangleq \mathbb{E}\left\{ |\widehat{s}_k - s_k|^2 \right\} \tag{3.16b}$$

where the expectations are taken over the distributions of the input signals and noise. By substituting the expressions of the received signals in (2.36) and (2.18) into (3.16a) and (3.16b) respectively, the MSEs are expressed by:

$$\varepsilon_{\mathrm{c},k} = |g_{\mathrm{c},k}|^2 T_{\mathrm{c},k} - 2\Re\{ g_{\mathrm{c},k} \mathbf{h}_k^H \mathbf{p}_{\mathrm{c}} \} + 1 \tag{3.17a}$$

$$\varepsilon_k = |g_k|^2 T_k - 2\Re\{ g_k \mathbf{h}_k^H \mathbf{p}_k \} + 1 \tag{3.17b}$$

where $T_{\mathrm{c},k}$ and T_k are defined in (2.37). Optimizing the MSEs over their respective equalizers yields the MMSEs given by

$$\varepsilon_{\mathrm{c},k}^{\mathrm{MMSE}} \triangleq \min_{g_{\mathrm{c},k}} \varepsilon_{\mathrm{c},k} = T_{\mathrm{c},k}^{-1} I_{\mathrm{c},k} \tag{3.18a}$$

$$\varepsilon_k^{\mathrm{MMSE}} \triangleq \min_{g_k} \varepsilon_k = T_k^{-1} I_k. \tag{3.18b}$$

The corresponding optimum equalizers are the well known MMSE equalizers given by

$$g_{\mathrm{c},k}^{\mathrm{MMSE}} = \mathbf{p}_{\mathrm{c}}^H \mathbf{h}_k T_{\mathrm{c},k}^{-1} \quad \text{and} \quad g_k^{\mathrm{MMSE}} = \mathbf{p}_k^H \mathbf{h}_k T_k^{-1}. \tag{3.19}$$

It turns out that MMSE estimation is closely related to the achievable rates. In particular, it can be easily seen that the MMSEs and the SINRs are related such that

$$\gamma_{\mathrm{c},k} = \left(1/\varepsilon_{\mathrm{c},k}^{\mathrm{MMSE}}\right) - 1 \quad \text{and} \quad \gamma_k = \left(1/\varepsilon_k^{\mathrm{MMSE}}\right) - 1 \tag{3.20}$$

from which the instantaneous achievable rates are expressed by

$$R_{\mathrm{c},k} = -\log_2(\varepsilon_{\mathrm{c},k}^{\mathrm{MMSE}}) \quad \text{and} \quad R_k = -\log_2(\varepsilon_k^{\mathrm{MMSE}}). \tag{3.21}$$

3.3.2. Rate-WMMSE Relationship

Now, we are ready to introduce the augmented WMSEs from which the Rate-WMMSE relationship is derived.

Definition 3.3. *For given $\varepsilon_{\mathrm{c},k}$ and ε_k, and the auxiliary positive weights denoted by $u_{\mathrm{c},k}$ and u_k, the augmented WMSEs are defined as*

$$\xi_{\mathrm{c},k} \triangleq u_{\mathrm{c},k}\varepsilon_{\mathrm{c},k} - \log_2(u_{\mathrm{c},k}) \tag{3.22a}$$

$$\xi_k \triangleq u_k\varepsilon_k - \log_2(u_k). \tag{3.22b}$$

In the following, the augmented WMSEs in (3.22) are simply referred to as the WMSEs, where "augmented" is dropped for brevity. The Rate-WMMSE relationship is established by optimizing the equalizers and weights in (3.22). This is shown as follows.

Lemma 3.1. *The Rate-WMMSE relationship is given by*

$$\xi_{\mathrm{c},k}^{\mathrm{MMSE}} \triangleq \min_{u_{\mathrm{c},k}, g_{\mathrm{c},k}} \xi_{\mathrm{c},k} = 1 - R_{\mathrm{c},k} \tag{3.23a}$$

$$\xi_k^{\mathrm{MMSE}} \triangleq \min_{u_k, g_k} \xi_k = 1 - R_k \tag{3.23b}$$

where $\xi_{\mathrm{c},k}^{\mathrm{MMSE}}$ and ξ_k^{MMSE} denote the optimized WMSEs, i.e. the WMMSEs.

This is a well known result expressed either implicitly or explicitly in the WMMSE literature [7, 8, 51]. The structure of the optimum solution is highlighted through the following proof, as it is important for the development of the WMMSE algorithm.

Proof. From $\frac{\partial \xi_{\mathrm{c},k}}{\partial g_{\mathrm{c},k}} = 0$ and $\frac{\partial \xi_k}{\partial g_k} = 0$, the optimum equalizers are given by the MMSE equalizers in (3.19), i.e. $g_{\mathrm{c},k}^{\mathrm{MMSE}}$ and g_k^{MMSE}. Substituting this back into (3.22) yields

$$\xi_{\mathrm{c},k}\big(g_{\mathrm{c},k}^{\mathrm{MMSE}}\big) = u_{\mathrm{c},k}\varepsilon_{\mathrm{c},k}^{\mathrm{MMSE}} - \log_2(u_{\mathrm{c},k}) \tag{3.24a}$$

$$\xi_k\big(g_k^{\mathrm{MMSE}}\big) = u_k\varepsilon_k^{\mathrm{MMSE}} - \log_2(u_k). \tag{3.24b}$$

Now, from $\frac{\partial \xi_{\mathrm{c},k}(g_{\mathrm{c},k}^{\mathrm{MMSE}})}{\partial u_{\mathrm{c},k}} = 0$ and $\frac{\partial \xi_k(g_k^{\mathrm{MMSE}})}{\partial u_k} = 0$, we obtain the optimum MMSE weights:

$$u_{\mathrm{c},k}^{\mathrm{MMSE}} = \big(\varepsilon_{\mathrm{c},k}^{\mathrm{MMSE}}\big)^{-1} \quad \text{and} \quad u_k^{\mathrm{MMSE}} = \big(\varepsilon_k^{\mathrm{MMSE}}\big)^{-1} \tag{3.25}$$

where a scaling factor of $\big(\ln(2)\big)^{-1}$ has been omitted as it has no effect on the solution. Substituting this back into (3.24) yields the relationship in (3.23). □

The rational behind employing the Rate-WMMSE relationship is explained in the following. Let us start with an observation on the variable-wise convexity of the WMSEs.

Remark 3.2. *While the WMSEs in* (3.22) *are not convex in their joint set of variables, they are convex in each variable while fixing the other two. For example,* $\xi_{\mathrm{c},k}$ *is convex in one of the variables* $\mathbf{P}$*,* $g_{\mathrm{c},k}$ *and* $u_{\mathrm{c},k}$*, while fixing the other two variables. This can be easily confirmed by checking* (3.22) *and* (3.17)*.*

This property is exploited to solve a wide range of non-convex rate optimization problems. To give a better idea, consider the following optimization problem

$$\max_{\mathbf{P}\in\mathbb{P}} \; U_{\mathrm{R}}\big(R_{\mathrm{c},1},\ldots,R_{\mathrm{c},K},R_1,\ldots,R_K\big) \tag{3.26}$$

where U_{R} is a utility function of the rates and $\mathbf{P}\in\mathbb{P}$ is usually a power constraint, from which a convex precoder feasible set $\mathbb{P}$ is characterized. We assume that U_{R} is concave and non-decreasing in each of its arguments. Now, consider the problem formulated as

$$\max_{\mathbf{g},\mathbf{u},\mathbf{P}\in\mathbb{P}} \; U_{\mathrm{W}}\big(\xi_{\mathrm{c},1},\ldots,\xi_{\mathrm{c},K},\xi_1,\ldots,\xi_K\big) \tag{3.27}$$

where $\mathbf{g}\triangleq\{g_{\mathrm{c},k},g_k \mid k\in\mathcal{K}\}$ and $\mathbf{u}\triangleq\{u_{\mathrm{c},k},u_k \mid k\in\mathcal{K}\}$ are the sets of equalizers and weights respectively, and the objective function in (3.27) is defined as

$$U_{\mathrm{W}}\big(\xi_{\mathrm{c},1},\ldots,\xi_{\mathrm{c},K},\xi_1,\ldots,\xi_K\big)\triangleq U_{\mathrm{R}}\big(1-\xi_{\mathrm{c},1},\ldots,1-\xi_{\mathrm{c},K},1-\xi_1,\ldots,1-\xi_K\big). \tag{3.28}$$

The next result builds upon Lemma 3.1 and establishes the equivalence between rate and WMSE problems.

Proposition 3.1. *The rate problem in* (3.26) *and the WMSE problem in* (3.27) *are equivalent in the sense that their global optimum solutions coincide.*

Proof. Since U_{R} is non-decreasing in its arguments, then U_{W} is non-increasing in its arguments. Combining this with (3.23), it follows that

$$U_{\mathrm{W}}\big(\xi_{\mathrm{c},1},\ldots,\xi_{\mathrm{c},K},\xi_1,\ldots,\xi_K\big)\leq U_{\mathrm{W}}\big(\xi_{\mathrm{c},1}^{\mathrm{MMSE}},\ldots,\xi_{\mathrm{c},K}^{\mathrm{MMSE}},\xi_1^{\mathrm{MMSE}},\ldots,\xi_K^{\mathrm{MMSE}}\big) \tag{3.29}$$

which holds for all $\mathbf{P}\in\mathbb{P}$. The rate problem in (3.26) is equivalent to maximizing the right-hand side in (3.29). This maximization in turn is an upper-bound on the WMSE problem in (3.27). Since this upper-bound is attainable through a feasible solution, i.e. the MMSE solution in Lemma (3.1), then it is optimum for the WMSE problem in (3.27). □

The variable-wise convexity highlighted in Remark 3.2 is inherited by the WMSE problem in (3.27), enabling a solution through AO as we see next.

3.3.3. Alternating Optimization Algorithm

Since U_{R} is concave in each of its arguments, then U_{W} is concave in each of the blocks $\mathbf{P}$, $\mathbf{g}$ and $\mathbf{u}$, while fixing the other two. This follows from Remark 3.2 and the vector composition rules of concave functions [52]. This block-wise concavity can be exploited to obtain a solution using AO, where each block is updated by solving a convex optimization problem. Moreover, at least two of the blocks, namely $\mathbf{g}$ and $\mathbf{u}$, have close-form solutions given by the MMSE formulations. This procedure is summarized in Algorithm 3.1.

Algorithm 3.1 WMMSE Alternating Optimization.

1: **Initialize**: $n \leftarrow 0$, $U_{\mathrm{W}}^{[n]} \leftarrow 0$, $\mathbf{P}$
2: **repeat**
3: $\quad n \leftarrow n+1$, $\mathbf{P}^{[n-1]} \leftarrow \mathbf{P}$
4: $\quad \mathbf{g} \leftarrow \mathbf{g}^{\mathrm{MMSE}}\big(\mathbf{P}^{[n-1]}\big)$
5: $\quad \mathbf{u} \leftarrow \mathbf{u}^{\mathrm{MMSE}}\big(\mathbf{P}^{[n-1]}\big)$
6: $\quad \mathbf{P} \leftarrow \arg\max_{\mathbf{P}\in\mathbb{P}} U_{\mathrm{W}}^{[n]}$
7: **until** $\left|U_{\mathrm{W}}^{[n]} - U_{\mathrm{W}}^{[n-1]}\right| < \epsilon_R$

The precoding matrix obtained in the previous (or $n-1$th) iteration is denoted by $\mathbf{P}^{[n-1]}$. On the other hand, $\mathbf{g}^{\mathrm{MMSE}}\big(\mathbf{P}^{[n-1]}\big)$ and $\mathbf{u}^{\mathrm{MMSE}}\big(\mathbf{P}^{[n-1]}\big)$ are the optimum equalizers and weights updated using $\mathbf{P}^{[n-1]}$. Plugging $\mathbf{g}^{\mathrm{MMSE}}\big(\mathbf{P}^{[n-1]}\big)$ and $\mathbf{u}^{\mathrm{MMSE}}\big(\mathbf{P}^{[n-1]}\big)$ into U_{W} yields $U_{\mathrm{W}}^{[n]}$, which is explicitly expressed as

$$U_{\mathrm{W}}^{[n]} = U_{\mathrm{W}}\Big(\xi_{\mathrm{c},1}^{\mathrm{MMSE}}\big(\mathbf{P}^{[n-1]}\big),\ldots,\xi_{\mathrm{c},K}^{\mathrm{MMSE}}\big(\mathbf{P}^{[n-1]}\big),\xi_{1}^{\mathrm{MMSE}}\big(\mathbf{P}^{[n-1]}\big),\ldots,\xi_{K}^{\mathrm{MMSE}}\big(\mathbf{P}^{[n-1]}\big)\Big)$$

where $\xi_{\mathrm{c},k}^{\mathrm{MMSE}}\big(\mathbf{P}^{[n-1]}\big)$ and $\xi_{k}^{\mathrm{MMSE}}\big(\mathbf{P}^{[n-1]}\big)$ are obtained by substituting $\mathbf{P}^{[n-1]}$ into (3.23). The precoder optimization in Step 6 varies depending on the utility function. For instance, $\mathbf{P}$ is simply obtained in closed-form, subject to an efficient bisection search, for the sum-rate utility in [8]. On the other hand, a utility with a max-min fair component yields a problem with convex quadratic constraints with no known closed-form solution [51].

The convergence of Algorithm 3.1 is guaranteed through the following result.

Proposition 3.2. *Given that U_{R} is bounded above, then Algorithm 3.1 monotonically increases U_{R} until convergence.*

Proof. Recall that U_{W} is non-increasing in each of its arguments. It follows that for a given iteration n, steps 4 and 5 increase (or do not decrease) U_{W} by minimizing each of its arguments for given $\mathbf{P}$. Moreover, step 6 further increases U_{W} by optimizing $\mathbf{P}$. As a direct result, the sequence $\big\{U_{\mathrm{W}}^{[n]}\big\}_{n=1}^{\infty}$ is monotonically increasing (or non decreasing) in each iteration. Since U_{R} and U_{W} are bounded above, then convergence is guaranteed. □

The condition that U_{R} is bounded above in usually satisfied. In particular, $\mathbb{P}$ is usually a compact set, as it is characterized by a power constraint on $\mathbf{P}$. Moreover, assuming

that channel gains are finite and $\sigma_{\mathrm{n}}^2 > 0$, then SINRs are finite and the rate functions are continuous on $\mathbb{P}$. It follows that U_{R} is continuous on $\mathbb{P}$, as it is continuous in its rate arguments (due to concavity). As a result, U_{R} is bounded and attains its maximum on $\mathbb{P}$.

Since the optimization problems of interest are usually non-convex, the convergence to a global optimum cannot be guaranteed in general. A stronger convergence result would be to show that the limit point of the iterates generated by Algorithm 3.1 is in fact a stationary point (i.e. satisfies the KKT conditions) of the WMSE problem in (3.27), and ultimately the rate problem in (3.26). This has been established for smooth sum utility functions in [8], and non-smooth max-min fair utility functions in [51]. It was later shown that the WMMSE algorithm belongs to a class of inexact BCDs, known as Successive Upper-bound Minimization (SUM), and more general analysis and proofs were established in [53, 54]. The KKT optimality of limit points obtained by the WMMSE procedures used throughout this thesis can be established based on [53, 54].

3.3.4. AWMSE Formulation

Motivated by the rather nice properties of WMSE problems, we propose an Average WMSE (AWMSE) formulation of the RS ASR problem in (3.12). Taking the expectation of the WMSEs in (3.22), we start by defining the AWMSEs as

$$\bar{\xi}_{\mathrm{c},k} \triangleq \mathbb{E}_{\mathrm{H}|\widehat{\mathrm{H}}}\left\{\xi_{\mathrm{c},k} \mid \widehat{\mathbf{H}}\right\} \quad \text{and} \quad \bar{\xi}_{k} \triangleq \mathbb{E}_{\mathrm{H}|\widehat{\mathrm{H}}}\left\{\xi_{k} \mid \widehat{\mathbf{H}}\right\}. \tag{3.30}$$

from which an average version of the Rate-WMMSE relationship in Lemma 3.1, namely the AR-AWMMSE relationship, is defined as

$$\bar{\xi}_{\mathrm{c},k}^{\mathrm{MMSE}} \triangleq \mathbb{E}_{\mathrm{H}|\widehat{\mathrm{H}}}\left\{\min_{u_{\mathrm{c},k},g_{\mathrm{c},k}} \xi_{\mathrm{c},k} \mid \widehat{\mathbf{H}}\right\} = 1 - \bar{R}_{\mathrm{c},k} \tag{3.31a}$$

$$\bar{\xi}_{k}^{\mathrm{MMSE}} \triangleq \mathbb{E}_{\mathrm{H}|\widehat{\mathrm{H}}}\left\{\min_{u_k,g_k} \xi_{k} \mid \widehat{\mathbf{H}}\right\} = 1 - \bar{R}_{k}. \tag{3.31b}$$

Note that the expectations in (3.31) are taken outside the minimizations to account for the dependencies of the optimum equalizers and weights on $\mathbf{H}$. It is important to keep these dependencies in mind when WMSE formulation are used under imperfect CSIT. As we see in this chapter and the next chapter, such dependencies are usually overlooked yielding more tractable, yet conservative, formulations.

The Average Weighted Sum MSE (AWSMSE) problem that corresponds to the ASR problem in (3.12) is formulated as

$$\mathcal{A}_{\mathrm{RS}}(P_{\mathrm{t}}): \begin{cases} \min\limits_{\bar{\xi}_{\mathrm{c}},\mathbf{P},\mathrm{u},\mathrm{g}} & \bar{\xi}_{\mathrm{c}} + \sum_{k=1}^{K} \bar{\xi}_{k} \\ \text{s.t.} & \bar{\xi}_{\mathrm{c},k} \leq \bar{\xi}_{\mathrm{c}},\ \forall k \in \mathcal{K} \\ & \mathrm{tr}\big(\mathbf{P}\mathbf{P}^{H}\big) \leq P_{\mathrm{t}} \end{cases} \tag{3.32}$$

where $\bar{\xi}_{\mathrm{c}}$ is an auxiliary variable representing the common AWMSE. This is a stochastic optimization problem, where g and u are random optimization variables corresponding to the $\mathbf{H}$-dependent equalizers and weights, respectively. In particular, single realizations of such random variables are given by $\mathbf{g}(\mathbf{H})$ and $\mathbf{u}(\mathbf{H})$ respectively, where the dependencies on $\mathbf{H}$ are highlighted. The equivalence and block-wise convexity properties discussed in the previous subsection hold between (3.12) and (3.32) such that

$$\mathcal{A}_{\mathrm{RS}}(P_{\mathrm{t}}) = K + 1 - \mathcal{R}_{\mathrm{RS}}(P_{\mathrm{t}}) \tag{3.33}$$

where $\bar{\xi}_{\mathrm{c}} = 1 - \bar{R}_{\mathrm{c}}$ holds at optimality. This follows from Proposition 3.1 by extending the set of achievable rates to include all realizations of $\mathbf{H}$ given $\widehat{\mathbf{H}}$, and taking the expectations as part of the concave utility function. Next, we look at approximate methods from which deterministic counterparts of (3.32) are obtained and solved.

3.4. Conservative Approximation

In this section, problem (3.32) is approximated by relaxing the dependencies of the equalizers and weights on the channel state $\mathbf{H}$. As a result, closed form expressions of the expectations are obtained. In particular, the conservative AWMSEs are given by

$$\widehat{\xi}_{\mathrm{c},k} \triangleq \mathbb{E}_{\mathrm{H}|\widehat{\mathrm{H}}}\left\{\xi_{\mathrm{c},k} \mid \widehat{\mathbf{H}}, \widehat{g}_{\mathrm{c},k}, \widehat{u}_{\mathrm{c},k}\right\} = \widehat{u}_{\mathrm{c},k}\left(|\widehat{g}_{\mathrm{c},k}|^2 \bar{T}_{\mathrm{c},k} - 2\Re\{\widehat{g}_{\mathrm{c},k}\widehat{\mathbf{h}}_k^H \mathbf{p}_{\mathrm{c}}\} + 1\right) - \log_2(\widehat{u}_{\mathrm{c},k}) \tag{3.34a}$$

$$\widehat{\xi}_k \triangleq \mathbb{E}_{\mathrm{H}|\widehat{\mathrm{H}}}\left\{\xi_k \mid \widehat{\mathbf{H}}, \widehat{g}_k, \widehat{u}_k\right\} = \widehat{u}_k\left(|\widehat{g}_k|^2 \bar{T}_k - 2\Re\{\widehat{g}_k\widehat{\mathbf{h}}_k^H \mathbf{p}_k\} + 1\right) - \log_2(\widehat{u}_k) \tag{3.34b}$$

where $(\widehat{g}_{\mathrm{c},k}, \widehat{g}_k)$ and $(\widehat{u}_{\mathrm{c},k}, \widehat{u}_k)$ denote the relaxed equalizers and weights respectively, and

$$\bar{T}_{\mathrm{c},k} \triangleq \mathbb{E}_{\mathrm{H}|\widehat{\mathrm{H}}}\left\{T_{\mathrm{c},k} \mid \widehat{\mathbf{H}}\right\} = \mathbf{p}_{\mathrm{c}}^H\big(\widehat{\mathbf{h}}_k\widehat{\mathbf{h}}_k^H + \mathbf{R}_{\mathrm{e},k}\big)\mathbf{p}_{\mathrm{c}} + \bar{T}_k \tag{3.35a}$$

$$\bar{T}_k \triangleq \mathbb{E}_{\mathrm{H}|\widehat{\mathrm{H}}}\left\{T_k \mid \widehat{\mathbf{H}}\right\} = \sum_{i=1}^{K} \mathbf{p}_i^H\big(\widehat{\mathbf{h}}_k\widehat{\mathbf{h}}_k^H + \mathbf{R}_{\mathrm{e},k}\big)\mathbf{p}_i + \sigma_{\mathrm{n}}^2. \tag{3.35b}$$

Following the same approach in the proof of Lemma 3.1, the optimum relaxed equalizers are derived from (3.34) as:

$$\widehat{g}_{\mathrm{c},k}^{\mathrm{MMSE}} = \mathbf{p}_{\mathrm{c}}^H\widehat{\mathbf{h}}_k\bar{T}_{\mathrm{c},k}^{-1} \quad \text{and} \quad \widehat{g}_k^{\mathrm{MMSE}} = \mathbf{p}_k^H\widehat{\mathbf{h}}_k\bar{T}_k^{-1}. \tag{3.36}$$

Plugging them back into (3.34), the optimum relaxed weights are obtained as:

$$\widehat{u}_{\mathrm{c},k}^{\mathrm{MMSE}} = \big(\widehat{\varepsilon}_{\mathrm{c},k}^{\mathrm{MMSE}}\big)^{-1} \quad \text{and} \quad \widehat{u}_k^{\mathrm{MMSE}} = \big(\widehat{\varepsilon}_k^{\mathrm{MMSE}}\big)^{-1} \tag{3.37}$$

where

$$\widehat{\varepsilon}_{\mathrm{c},k}^{\mathrm{MMSE}} \triangleq 1 - \bar{T}_{\mathrm{c},k}^{-1}|\widehat{\mathbf{h}}_k^H\mathbf{p}_{\mathrm{c}}|^2 \quad \text{and} \quad \widehat{\varepsilon}_k^{\mathrm{MMSE}} \triangleq 1 - \bar{T}_k^{-1}|\widehat{\mathbf{h}}_k^H\mathbf{p}_k|^2. \tag{3.38}$$

It is clear that the optimum relaxed equalizers and weights in (3.36) and (3.37) are functions of $\widehat{\mathbf{H}}$ rather than $\mathbf{H}$. Incorporating this approximation into the AR-AWMMSE relationship in (3.31), the expectations are taken inside the minimizations such that

$$\widehat{\xi}_{\mathrm{c},k}^{\mathrm{MMSE}} \triangleq \min_{\widehat{u}_{\mathrm{c},k},\widehat{g}_{\mathrm{c},k}} \mathbb{E}_{\mathrm{H}|\widehat{\mathrm{H}}} \left\{ \xi_{\mathrm{c},k} \mid \widehat{\mathbf{H}} \right\} = 1 - \widehat{R}_{\mathrm{c},k} \tag{3.39a}$$

$$\widehat{\xi}_{k}^{\mathrm{MMSE}} \triangleq \min_{\widehat{u}_{k},\widehat{g}_{k}} \mathbb{E}_{\mathrm{H}|\widehat{\mathrm{H}}} \left\{ \xi_{k} \mid \widehat{\mathbf{H}} \right\} = 1 - \widehat{R}_{k}. \tag{3.39b}$$

This is referred to as the conservative AR-AWMMSE relationship, where the conservative ARs are given by

$$\widehat{R}_{\mathrm{c},k} \triangleq -\log_2(\widehat{\varepsilon}_{\mathrm{c},k}^{\mathrm{MMSE}}) \quad \text{and} \quad \widehat{R}_{k} \triangleq -\log_2(\widehat{\varepsilon}_{k}^{\mathrm{MMSE}}). \tag{3.40}$$

Next, these approximations are leveraged to formulate a deterministic AWSMSE problem, which can be solved using the AO principle in Algorithm 3.1. Moreover, we prove that the resulting performance is in fact conservative.

3.4.1. Conservative AWSMSE Problem

By employing (3.34), a conservative version of the AWSMSE problem is formulated as

$$\widehat{\mathcal{A}}_{\mathrm{RS}}(P_{\mathrm{t}}) : \begin{cases} \min\limits_{\widehat{\xi}_{\mathrm{c}},\mathbf{P},\widehat{\mathbf{u}},\widehat{\mathbf{g}}} & \widehat{\xi}_{\mathrm{c}} + \sum_{k=1}^{K} \widehat{\xi}_{k} \\ \text{s.t.} & \widehat{\xi}_{\mathrm{c},k} \leq \widehat{\xi}_{\mathrm{c}}, \ \forall k \in \mathcal{K} \\ & \mathrm{tr}\big(\mathbf{P}\mathbf{P}^{H}\big) \leq P_{\mathrm{t}} \end{cases} \tag{3.41}$$

where $\widehat{\mathbf{u}} \triangleq \left\{ \widehat{u}_{\mathrm{c},k}, \widehat{u}_{k} \mid k \in \mathcal{K} \right\}$ and $\widehat{\mathbf{g}} \triangleq \left\{ \widehat{g}_{\mathrm{c},k}, \widehat{g}_{k} \mid k \in \mathcal{K} \right\}$. So far, approximations and formulations have been described as conservative without explaining why. This is shown as follows. It is evident that (3.41) is a restricted version of the AWSMSE problem in (3.32), where the domains of g and u in (3.32) are restricted such that $\mathbf{g}(\mathbf{H}) = \widehat{\mathbf{g}}$ and $\mathbf{u}(\mathbf{H}) = \widehat{\mathbf{u}}$ remain unchanged across all realization of $\mathbf{H}$ given $\widehat{\mathbf{H}}$. Hence by definition, solving (3.41) yields an upper-bound on $\mathcal{A}_{\mathrm{RS}}(P_{\mathrm{t}})$. Equivalently, we obtain a lower-bound on $\mathcal{R}_{\mathrm{RS}}(P_{\mathrm{t}})$ denoted by $\widehat{\mathcal{R}}_{\mathrm{RS}}(P_{\mathrm{t}}) = K + 1 - \widehat{\mathcal{A}}_{\mathrm{RS}}(P_{\mathrm{t}})$. It follows that solving a sequence of problem (3.41) over all $\widehat{\mathbf{H}}$ yields an ESR give by

$$\mathbb{E}_{\widehat{\mathrm{H}}} \left\{ \widehat{\mathcal{R}}_{\mathrm{RS}}(P_{\mathrm{t}}) \right\} \triangleq 1 + K - \mathbb{E}_{\widehat{\mathrm{H}}} \left\{ \widehat{\mathcal{A}}_{\mathrm{RS}}(P_{\mathrm{t}}) \right\} \leq \mathbb{E}_{\widehat{\mathrm{H}}} \left\{ \mathcal{R}_{\mathrm{RS}}(P_{\mathrm{t}}) \right\} \tag{3.42}$$

which is clearly upper-bounded by the ESR obtained by a sequence of (3.32).

Since streams are independently decoded by different users, an achievable ESR is not sufficient to guarantee reliable communication, as achievable rates for each individual stream must be prescribed as well. We show that any sequence of feasible solutions for (3.41)

yields achievable ERs. For any given precoding matrix $\mathbf{P}$, we can write the inequalities

$$\widehat{\xi}_{\mathrm{c},k} \geq \widehat{\xi}_{\mathrm{c},k}^{\mathrm{MMSE}} \geq \bar{\xi}_{\mathrm{c},k}^{\mathrm{MMSE}} \quad \text{and} \quad \widehat{\xi}_k \geq \widehat{\xi}_k^{\mathrm{MMSE}} \geq \bar{\xi}_k^{\mathrm{MMSE}} \tag{3.43}$$

which follow from (3.31), (3.39) and the fact that moving the expectation inside the minimization does not decrease the value. As a direct results, we have

$$\widehat{R}_{\mathrm{c},k} \leq \bar{R}_{\mathrm{c},k} \quad \text{and} \quad \widehat{R}_k \leq \bar{R}_k. \tag{3.44}$$

Combining this with (3.4) and (3.10), it follows that reliable communication is guerenteed at the ERs given by $\mathbb{E}_{\widehat{\mathrm{H}}}\left\{\min_j \widehat{R}_{\mathrm{c},j}\right\}$ for the common message and $\mathbb{E}_{\widehat{\mathrm{H}}}\left\{\widehat{R}_k\right\}$ for the kth private message, for any feasible sequence of $\mathbf{P}$.

3.4.2. Alternating Optimization

The problem in (3.41) can be solved using the AO principle in Algorithm 3.1. In the nth iteration, $(\widehat{\mathbf{g}}, \widehat{\mathbf{u}})$ are updated by plugging $\mathbf{P}^{[n-1]}$ into equations (3.36) and (3.37). On the other hand, $\mathbf{P}$ is updated by solving the problem

$$\widehat{\mathcal{A}}_{\mathrm{RS}}^{[n]}(P_{\mathrm{t}}): \begin{cases} \min\limits_{\bar{\xi}_{\mathrm{c}},\mathbf{P}} & \bar{\xi}_{\mathrm{c}} + \sum_{k=1}^{K}\left(\widehat{u}_k|\widehat{g}_k|^2\bar{T}_k - \widehat{u}_k 2\Re\{\widehat{g}_k\widehat{\mathbf{h}}_k^H\mathbf{p}_k\} + \widehat{u}_k - \log_2(\widehat{u}_k)\right) \\ \text{s.t.} & \widehat{u}_{\mathrm{c},k}|\widehat{g}_{\mathrm{c},k}|^2\bar{T}_{\mathrm{c},k} - 2\widehat{u}_{\mathrm{c},k}\Re\{\widehat{g}_{\mathrm{c},k}\widehat{\mathbf{h}}_k^H\mathbf{p}_{\mathrm{c}}\} + \widehat{u}_{\mathrm{c},k} - \log_2(\widehat{u}_{\mathrm{c},k}) \leq \widehat{\xi}_{\mathrm{c}},\ \forall k \in \mathcal{K} \\ & \|\mathbf{p}_{\mathrm{c}}\|^2 + \sum_{k=1}^{K}\|\mathbf{p}_k\|^2 \leq P_{\mathrm{t}}. \end{cases} \tag{3.45}$$

The above problem is a convex Quadratically Constrained Quadratic Program (QCQP), which is confirmed by checking equations (3.34) and (3.35). Such problem can be solved efficiently to global optimality using interior-point methods [52].

The AO procedure is guaranteed to converge by Proposition 3.2. The optimality of the limit point achieved by AO is not of much interest in the conservative case. One reason is that the best achievable performance is upper-bounded by $\widehat{\mathcal{R}}_{\mathrm{RS}}(P_{\mathrm{t}})$, which is shown to be a loose lower-bound of $\mathcal{R}_{\mathrm{RS}}(P_{\mathrm{t}})$, at least in the high SNR regime.

3.4.3. Limitations

Although the conservative AWMMSE approach appears in literature [46,47], its limitations have not been fully identified. The primary limitation is that the ARs predicted by the BS through solving (3.41) are loose lower-bounds on the actual ARs. Hence, transmitting at the ESR given by $\mathbb{E}_{\widehat{\mathrm{H}}}\left\{\widehat{\mathcal{R}}_{\mathrm{RS}}(P_{\mathrm{t}})\right\}$ causes some performance losses, shown later in the simulation result. Here, the looseness of the conservative ARs is shown analytically. For the sake of analysis, we assume isotropic CSIT errors as in Assumption 3.1.

The conservative ARs in (3.40) can be expressed by

$$\widehat{R}_{\mathrm{c},k} \triangleq \log_2\left(1 + \widehat{\gamma}_{\mathrm{c},k}\right) \quad \text{and} \quad \widehat{R}_k \triangleq \log_2\left(1 + \widehat{\gamma}_k\right). \tag{3.46}$$

where

$$\widehat{\gamma}_{\mathrm{c},k} \triangleq \frac{|\mathbf{p}_{\mathrm{c}}^H \widehat{\mathbf{h}}_k|^2}{\widehat{I}_{\mathrm{c},k} + \widetilde{I}_{\mathrm{c},k}} \quad \text{and} \quad \widehat{\gamma}_k \triangleq \frac{|\mathbf{p}_k^H \widehat{\mathbf{h}}_k|^2}{\widehat{I}_k + \widetilde{I}_k} \tag{3.47}$$

with

$$\widehat{I}_{\mathrm{c},k} \triangleq \sum_{i=1}^{K} |\mathbf{p}_i^H \widehat{\mathbf{h}}_k|^2 + \sigma_{\mathrm{e},k}^2 \|\mathbf{p}_i\|^2 + \sigma_{\mathrm{n}}^2 \quad \text{and} \quad \widehat{I}_k \triangleq \sum_{i \neq k} |\mathbf{p}_i^H \widehat{\mathbf{h}}_k|^2 + \sigma_{\mathrm{e},k}^2 \|\mathbf{p}_i\|^2 + \sigma_{\mathrm{n}}^2 \tag{3.48}$$

and

$$\widetilde{I}_{\mathrm{c},k} = \sigma_{\mathrm{e},k}^2 \|\mathbf{p}_{\mathrm{c}}\|^2 \quad \text{and} \quad \widetilde{I}_k = \sigma_{\mathrm{e},k}^2 \|\mathbf{p}_k\|^2. \tag{3.49}$$

This follows directly from (3.38) and (3.40). By observing (3.47), it can be seen that the useful signal power components in $\widehat{\gamma}_{\mathrm{c},k}$ and $\widehat{\gamma}_k$ only consist of parts incorporating the channel estimate, while the parts of the desired signal power incorporating the CSIT errors, namely $\widetilde{I}_{\mathrm{c},k}$ and $\widetilde{I}_k$, appear as interference. This is explained as follows: by removing the dependencies of equalizers (and weights) on the actual channel, the robust rates (as guaranteed by the BS) are optimized while ignoring the availability of perfect CSIR, and hence treating parts of the desired signal which incorporate $\widetilde{\mathbf{h}}_k$ as noise. This yields the self-interference terms given in (3.49).

Assume a power scaling scaling as: $q_{\mathrm{c}} = O\big(P_{\mathrm{t}}^{a_{\mathrm{c}}}\big)$ and $q_k = O\big(P_{\mathrm{t}}^{a_k}\big)$, where $a_{\mathrm{c}}, a_k \in [0,1]$ are power scaling factors. It follows that the self-interference terms scale as $\widetilde{I}_{\mathrm{c},k} = O\big(P_{\mathrm{t}}^{a_{\mathrm{c}}-\alpha}\big)$ and $\widetilde{I}_k = O\big(P_{\mathrm{t}}^{a_k-\alpha}\big)$. This is detrimental to the ARs as perceived by the BS. For example, a fixed CSIT quality with $\alpha = 0$ yields $\widetilde{I}_{\mathrm{c},k}$ and $\widetilde{I}_k$ that scale as $O\big(P_{\mathrm{t}}^{a_{\mathrm{c}}}\big)$ and $O\big(P_{\mathrm{t}}^{a_k}\big)$ respectively. This results in saturating SINRs in (3.47) for $P_{\mathrm{t}} \to \infty$, as interference scales similar to the desired signal power. The influence of this is seen in Section 3.6.

3.5. Sample Average Approximation

The limitations of the conservative approach shown in the previous section motivate the development of a method that can achieve the performance predicted from the DoF analysis. In this section, we propose a non-conservative AWMSE formulation based on SAA [9]. We start by defining the main building blocks of this method.

For a given estimate $\widehat{\mathbf{H}}$ and index set $\mathcal{M} \triangleq \{1, \ldots, M\}$, we define a sample of M i.i.d realizations of $\mathbf{H}$ drawn from the conditional distribution with density $f_{\mathrm{H}|\widehat{\mathrm{H}}}\big(\mathbf{H} \mid \widehat{\mathbf{H}}\big)$ as

$$\mathbb{H}^{(M)} \triangleq \big\{\mathbf{H}^{(m)} = \widehat{\mathbf{H}} + \widetilde{\mathbf{H}}^{(m)} \mid \widehat{\mathbf{H}},\ m \in \mathcal{M}\big\}. \tag{3.50}$$

This is used to approximate the ARs through their Sample Average Functions (SAFs) defined as:

$$\bar{R}_{\mathrm{c},k}^{(M)} \triangleq \frac{1}{M} \sum_{m=1}^{M} R_{\mathrm{c},k}^{(m)} \quad \text{and} \quad \bar{R}_k^{(M)} \triangleq \frac{1}{M} \sum_{m=1}^{M} R_k^{(m)} \tag{3.51}$$

where

$$R_{\mathrm{c},k}^{(m)} \triangleq R_{\mathrm{c},k}\big(\mathbf{H}^{(m)}\big) \quad \text{and} \quad R_{k}^{(m)} \triangleq R_{k}\big(\mathbf{H}^{(m)}\big) \tag{3.52}$$

are instantaneous rates associated with the mth realization. It should be noted that $\mathbf{P}$ is fixed over the M realizations, which follows from the definition of the ARs. This leads to the SAA of the ASR problem in (3.12) formulated as

$$\mathcal{R}_{\mathrm{RS}}^{(M)}(P_{\mathrm{t}}) : \begin{cases} \max\limits_{\bar{R}_{\mathrm{c}},\mathbf{P}} & \bar{R}_{\mathrm{c}} + \sum_{k=1}^{K} \bar{R}_{k}^{(M)} \\ \text{s.t.} & \bar{R}_{\mathrm{c},k}^{(M)} \geq \bar{R}_{\mathrm{c}}, \ \forall k \in \mathcal{K} \\ & \mathrm{tr}\big(\mathbf{P}\mathbf{P}^{H}\big) \leq P_{\mathrm{t}}. \end{cases} \tag{3.53}$$

Before we proceed, we make the following finite SNR/SINR assumption.

Assumption 3.2. *In the following, it is assumed that* $\mathrm{SNR} < \infty$*, i.e.* $\sigma_{\mathrm{n}}^2 > 0$ *and* $P_{\mathrm{t}} < \infty$*, and that entries of channel state realizations are bounded. Hence,* $\gamma_{\mathrm{c},k}, \gamma_k < \infty$*.*

It follows that all achievable rates are bounded for any P_{t} and all channel realizations. Therefore, from the strong Law of Large Numbers (LLN), we have

$$\lim_{M\to\infty} \bar{R}_{\mathrm{c},k}^{(M)}(\mathbf{P}) = \bar{R}_{\mathrm{c},k}(\mathbf{P}) \quad \text{and} \quad \lim_{M\to\infty} \bar{R}_{k}^{(M)}(\mathbf{P}) = \bar{R}_{k}(\mathbf{P}), \ \text{a.s.} \ \forall \mathbf{P} \in \mathbb{P} \tag{3.54}$$

where the dependency of the rates on a given $\mathbf{P}$ is highlighted, a.s. denotes almost surely, and $\mathbb{P} \triangleq \big\{\mathbf{P} \mid \mathrm{tr}\big(\mathbf{P}\mathbf{P}^{H}\big) \leq P_{\mathrm{t}}\big\}$ is the feasible set of precoders. (3.54) suggests that the solution of problem (3.53) converges to its counterpart of problem (3.12) as $M \to \infty$. We show that this is indeed the case.

First, note that the feasible set $\mathbb{P}$ is compact and the instantaneous rate functions are bounded and continuously differentiable in $\mathbf{P}$. Hence, the convergence in (3.54) is uniform in $\mathbf{P} \in \mathbb{P}$. The ARs are also continuously differentiable with gradients given by $\nabla_{\mathbf{P}} \bar{R}_{\mathrm{c},k}(\mathbf{P}) = \mathbb{E}_{\mathrm{H}|\widehat{\mathrm{H}}}\big\{\nabla_{\mathbf{P}} R_{\mathrm{c},k}(\mathbf{P}) \mid \widehat{\mathbf{H}}\big\}$ and $\nabla_{\mathbf{P}} \bar{R}_{k}(\mathbf{P}) = \mathbb{E}_{\mathrm{H}|\widehat{\mathrm{H}}}\big\{\nabla_{\mathbf{P}} R_{k}(\mathbf{P}) \mid \widehat{\mathbf{H}}\big\}$, which follows from the bounded convergence theorem [55]. Next, we observe that the objective functions of (3.12) and (3.53) can be equivalently reformulated by incorporating the common rates as

$$\bar{R}_{\mathrm{s}} \triangleq \min_{j} \big\{\bar{R}_{\mathrm{c},j}\big\}_{j=1}^{K} + \sum_{k=1}^{k} \bar{R}_{k} \quad \text{and} \quad \bar{R}_{\mathrm{s}}^{(M)} \triangleq \min_{j} \big\{\bar{R}_{\mathrm{c},j}^{(M)}\big\}_{j=1}^{K} + \sum_{k=1}^{k} \bar{R}_{k}^{(M)} \tag{3.55}$$

where $\bar{R}_{\mathrm{s}}$ and $\bar{R}_{\mathrm{s}}^{(M)}$ are the equivalent ASRs. From (3.54), (3.55) and the continuity of the pointwise minimization, we have

$$\lim_{M\to\infty} \bar{R}_{\mathrm{s}}^{(M)}(\mathbf{P}) = \bar{R}_{\mathrm{s}}(\mathbf{P}), \ \text{a.s.} \ \forall \mathbf{P} \in \mathbb{P}. \tag{3.56}$$

From the continuity of the ARs, we observe that $\bar{R}_{\mathrm{s}}$ is continuous in $\mathbf{P}$ although not necessarily differentiable at all points due to the embedded pointwise minimization. Hence, it can be shown that the convergence in (3.56) is also uniform. Combining these observations

with [9, Theorem 5.3], it is concluded that the set of global optimum solutions of the SAA problem in (3.53) converges to that of the stochastic problem in (3.12) a.s. as $M \to \infty$. Extending this result to the set of points that satisfy the first-order optimality conditions (i.e. KKT points) for non-convex non-smooth problems can be found in [56, Section 4.3]. Now, we turn to solving the sampled problem in (3.53) using the WMMSE approach.

3.5.1. Sampled AWSMSE Problem

To formulate a sampled version of problem (3.32), the AWMSEs in (3.30) are first approximated by their SAFs expressed as

$$\bar{\xi}_{\mathrm{c},k}^{(M)} \triangleq \frac{1}{M}\sum_{m=1}^{M} \xi_{\mathrm{c},k}^{(m)} \quad \text{and} \quad \bar{\xi}_{k}^{(M)} \triangleq \frac{1}{M}\sum_{m=1}^{M} \xi_{k}^{(m)} \tag{3.57}$$

where $\xi_{\mathrm{c},k}^{(m)} \triangleq \xi_{\mathrm{c},k}\big(\mathbf{h}_k^{(m)}, g_{\mathrm{c},k}^{(m)}, u_{\mathrm{c},k}^{(m)}\big)$ and $\xi_{k}^{(m)} \triangleq \xi_{k}\big(\mathbf{h}_k^{(m)}, g_{k}^{(m)}, u_{k}^{(m)}\big)$. All variables with superscript (m) are associated with the mth realization in $\mathbb{H}^{(M)}$. It is clear that the sampling in (3.57) includes the equalizers and weights to capture their dependencies on $\mathbf{H}^{(M)}$, where $g_{\mathrm{c},k}^{(m)} \triangleq g_{\mathrm{c},k}\big(\mathbf{h}_k^{(m)}\big)$, $g_{k}^{(m)} \triangleq g_{k}\big(\mathbf{h}_k^{(m)}\big)$, $u_{\mathrm{c},k}^{(m)} \triangleq u_{\mathrm{c},k}\big(\mathbf{h}_k^{(m)}\big)$ and $u_{k}^{(m)} \triangleq u_{k}\big(\mathbf{h}_k^{(m)}\big)$.

For compactness, we define the set of sampled equalizers as: $\mathbf{G} \triangleq \big\{\mathbf{g}_{\mathrm{c},k}, \mathbf{g}_k \mid k \in \mathcal{K}\big\}$, where $\mathbf{g}_{\mathrm{c},k} \triangleq \big\{g_{\mathrm{c},k}^{(m)} \mid m \in \mathcal{M}\big\}$ and $\mathbf{g}_k \triangleq \big\{g_{k}^{(m)} \mid m \in \mathcal{M}\big\}$. In a similar manner, we define: $\mathbf{U} \triangleq \big\{\mathbf{u}_{\mathrm{c},k}, \mathbf{u}_k \mid k \in \mathcal{K}\big\}$, where $\mathbf{u}_{\mathrm{c},k} \triangleq \big\{u_{\mathrm{c},k}^{(m)} \mid m \in \mathcal{M}\big\}$ and $\mathbf{u}_k \triangleq \big\{u_{k}^{(m)} \mid m \in \mathcal{M}\big\}$. The same approach used to prove (3.23) is employed to demonstrate the following relationship

$$\bar{\xi}_{\mathrm{c},k}^{\mathrm{MMSE}(M)} \triangleq \min_{\mathbf{u}_{\mathrm{c},k},\mathbf{g}_{\mathrm{c},k}} \bar{\xi}_{\mathrm{c},k}^{(M)} = 1 - \bar{R}_{\mathrm{c},k}^{(M)} \tag{3.58a}$$

$$\bar{\xi}_{k}^{\mathrm{MMSE}(M)} \triangleq \min_{\mathbf{u}_{k},\mathbf{g}_{k}} \bar{\xi}_{k}^{(M)} = 1 - \bar{R}_{k}^{(M)} \tag{3.58b}$$

where optimality conditions are checked separately for each conditional channel realization. The sets of optimum MMSE equalizers associated with (3.58) are denoted by $\mathbf{g}_{\mathrm{c},k}^{\mathrm{MMSE}} \triangleq \big\{g_{\mathrm{c},k}^{\mathrm{MMSE}(m)} \mid m \in \mathcal{M}\big\}$ and $\mathbf{g}_{k}^{\mathrm{MMSE}} \triangleq \big\{g_{k}^{\mathrm{MMSE}(m)} \mid m \in \mathcal{M}\big\}$. In the same manner, the sets of optimum MMSE weights are given by $\mathbf{u}_{\mathrm{c},k}^{\mathrm{MMSE}} \triangleq \big\{u_{\mathrm{c},k}^{\mathrm{MMSE}(m)} \mid m \in \mathcal{M}\big\}$ and $\mathbf{u}_{k}^{\mathrm{MMSE}} \triangleq \big\{u_{k}^{\mathrm{MMSE}(m)} \mid m \in \mathcal{M}\big\}$. For the K users, the MMSE solution is composed as $\mathbf{G}^{\mathrm{MMSE}} \triangleq \big\{\mathbf{g}_{\mathrm{c},k}^{\mathrm{MMSE}}, \mathbf{g}_{k}^{\mathrm{MMSE}} \mid k \in \mathcal{K}\big\}$ and $\mathbf{U}^{\mathrm{MMSE}} \triangleq \big\{\mathbf{u}_{\mathrm{c},k}^{\mathrm{MMSE}}, \mathbf{u}_{k}^{\mathrm{MMSE}} \mid k \in \mathcal{K}\big\}$.

The sampled AWSMSE problem which is equivalent to (3.53) is given by

$$\mathcal{A}_{\mathrm{RS}}^{(M)}(P_{\mathrm{t}}): \begin{cases} \min\limits_{\bar{\xi}_{\mathrm{c}},\mathbf{P},\mathbf{U},\mathbf{G}} & \bar{\xi}_{\mathrm{c}} + \sum_{k=1}^{K} \bar{\xi}_{k}^{(M)} \\ \text{s.t.} & \bar{\xi}_{\mathrm{c},k}^{(M)} \leq \bar{\xi}_{\mathrm{c}},\ \forall k \in \mathcal{K} \\ & \mathrm{tr}\big(\mathbf{P}\mathbf{P}^{H}\big) \leq P_{\mathrm{t}}. \end{cases} \tag{3.59}$$

The equivalence between (3.59) and (3.53) follows directly from Proposition 3.1. Now the goal is to solve problem (3.59). Since this problem is deterministic, AO can be directly

applied as we see next.

3.5.2. Algorithm, Convergence and Optimality

In nth iteration of the AO algorithm, the equalizers and weights are updated such that $(\mathbf{G},\mathbf{U}) = \big(\mathbf{G}^{\mathrm{MMSE}}(\mathbf{P}^{[n-1]}),\mathbf{U}^{\mathrm{MMSE}}(\mathbf{P}^{[n-1]})\big)$, where the dependencies on $\mathbf{P}^{[n-1]}$ are highlighted, and the MMSE solution is obtained for each realization using (3.19) and (3.25). To facilitate the formulation of the precoder optimization problem in the following step, we introduce the SAFs listed as: $\bar{t}_{\mathrm{c},k}$, $\bar{t}_k$, $\bar{\mathbf{\Psi}}_{\mathrm{c},k}$, $\bar{\mathbf{\Psi}}_k$, $\bar{\mathbf{f}}_{\mathrm{c},k}$, $\bar{\mathbf{f}}_k$, $\bar{u}_{\mathrm{c},k}$, $\bar{u}_k$, $\bar{v}_{\mathrm{c},k}$ and $\bar{v}_k$, which are obtained using the updated $(\mathbf{G},\mathbf{U})$. In particular, $\bar{u}_{\mathrm{c},k}$ and $\bar{u}_k$ are calculated by taking the ensemble averages over the M realizations of $u_{\mathrm{c},k}^{(m)}$ and $u_k^{(m)}$, respectively. The rest are calculated in a similar manner by averaging over their corresponding realizations given by:

$$\begin{aligned}
t_{\mathrm{c},k}^{(m)} &= u_{\mathrm{c},k}^{(m)}\left|g_{\mathrm{c},k}^{(m)}\right|^2 \quad \text{and} \quad t_k^{(m)} = u_k^{(m)}\left|g_k^{(m)}\right|^2 \\
\mathbf{\Psi}_{\mathrm{c},k}^{(m)} &= t_{\mathrm{c},k}^{(m)}\mathbf{h}_k^{(m)}\mathbf{h}_k^{(m)H} \quad \text{and} \quad \mathbf{\Psi}_k^{(m)} = t_k^{(m)}\mathbf{h}_k^{(m)}\mathbf{h}_k^{(m)H} \\
\mathbf{f}_{\mathrm{c},k}^{(m)} &= u_{\mathrm{c},k}^{(m)}\mathbf{h}_k^{(m)}g_{\mathrm{c},k}^{(m)H} \quad \text{and} \quad \mathbf{f}_k^{(m)} = u_k^{(m)}\mathbf{h}_k^{(m)}g_k^{(m)H} \\
v_{\mathrm{c},k}^{(m)} &= \log_2\left(u_{\mathrm{c},k}^{(m)}\right) \quad \text{and} \quad v_k^{(m)} = \log_2\left(u_k^{(m)}\right).
\end{aligned} \tag{3.60}$$

Following the previous step, the problem of updating $\mathbf{P}$ is formulated by plugging $(\mathbf{G},\mathbf{U}) = \big(\mathbf{G}^{\mathrm{MMSE}}(\mathbf{P}^{[n-1]}),\mathbf{U}^{\mathrm{MMSE}}(\mathbf{P}^{[n-1]})\big)$ into (3.59). This yields the following problem

$$\mathcal{A}_{\mathrm{RS}}^{\mathrm{MMSE}[n]}\big(P_{\mathrm{t}},\mathbf{G},\mathbf{U}\big): \begin{cases} \min\limits_{\bar{\xi}_{\mathrm{c}},\mathbf{P}} & \bar{\xi}_{\mathrm{c}} + \sum_{k=1}^{K}\Big(\sum_{i=1}^{K}\mathbf{p}_i^H\bar{\mathbf{\Psi}}_k\mathbf{p}_i + \sigma_{\mathrm{n}}^2\bar{t}_k - 2\Re\big\{\bar{\mathbf{f}}_k^H\mathbf{p}_k\big\} + \bar{u}_k - \bar{v}_k\Big) \\ \text{s.t.} & \mathbf{p}_{\mathrm{c}}^H\bar{\mathbf{\Psi}}_{\mathrm{c},k}\mathbf{p}_{\mathrm{c}} + \sum_{i=1}^{K}\mathbf{p}_i^H\bar{\mathbf{\Psi}}_{\mathrm{c},k}\mathbf{p}_i + \sigma_{\mathrm{n}}^2\bar{t}_{\mathrm{c},k} - 2\Re\big\{\bar{\mathbf{f}}_{\mathrm{c},k}^H\mathbf{p}_{\mathrm{c}}\big\} \\ & \qquad +\bar{u}_{\mathrm{c},k} - \bar{v}_{\mathrm{c},k} \le \bar{\xi}_{\mathrm{c}},\ \forall k\in\mathcal{K} \\ & \|\mathbf{p}_{\mathrm{c}}\|^2 + \sum_{k=1}^{K}\|\mathbf{p}_k\|^2 \le P_{\mathrm{t}}. \end{cases} \tag{3.61}$$

The AWMSEs expressions in (3.61) are obtained by substituting the SAFs obtained from (3.60) into (3.57). This is carried out while considering the updated equalizers and weights from the previous step, embedded into the realizations in (3.60). Similar to problem (3.45), problem (3.61) is also a convex QCQP which can be solved using interior-point methods. The steps to the described procedure are summarized in Algorithm 3.2.

It turns out that since Algorithm (3.2) does not employ conservative approximations, conclusions about the optimality of the solution with respect to the original ASR problem in (3.12) can be made as we see next.

Proposition 3.3. *For a given $\mathbb{H}^{(M)}$, the iterates generated by Algorithm 3.2 converge to the set of KKT points of the corresponding sampled ASR problem in* (3.53).

Proof. The AO procedure described above is an instance of the Successive Convex Approximation (SCA) method in [54, Section 2.1]. In particular, solving (3.61) is equivalent

Algorithm 3.2 Alternating Optimization

Initialize: $n \leftarrow 0$, $\mathcal{A}_{\mathrm{RS}}^{\mathrm{MMSE}[n]} \leftarrow 0$, $\mathbf{P}$
repeat
  $n \leftarrow n+1$, $\mathbf{P}^{[n-1]} \leftarrow \mathbf{P}$
  $\mathbf{G} \leftarrow \mathbf{G}^{\mathrm{MMSE}}(\mathbf{P}^{[n-1]})$, $\mathbf{U} \leftarrow \mathbf{U}^{\mathrm{MMSE}}(\mathbf{P}^{[n-1]})$
  update $\left\{\bar{\mathbf{\Psi}}_{\mathrm{c},k}, \bar{\mathbf{\Psi}}_k, \bar{\mathbf{F}}_{\mathrm{c},k}, \bar{\mathbf{F}}_k, \bar{t}_{\mathrm{c},k}, \bar{t}_k, \bar{u}_{\mathrm{c},k}, \bar{u}_k, \bar{v}_{\mathrm{c},k}, \bar{v}_k\right\}_{k=1}^{K}$
  $\mathbf{P} \leftarrow \arg \mathcal{A}_{\mathrm{RS}}^{\mathrm{MMSE}[n]}$
until $\left|\mathcal{A}_{\mathrm{RS}}^{\mathrm{MMSE}[n]} - \mathcal{A}_{\mathrm{RS}}^{\mathrm{MMSE}[n-1]}\right| < \epsilon_R$

to solving a convex approximation of (3.53) around $\mathbf{P}^{[n-1]}$, obtained from the previous iteration. Moreover, (3.61) satisfies the conditions in [54, Assumption 1]. Hence, it follows from [54, Theorem 1] that any limit point of the AO procedure is a KKT point of problem (3.53). Now, since the iterates lie in a compact set, characterized by $\mathrm{tr}\left(\mathbf{P}\mathbf{P}^H\right) \leq P_{\mathrm{t}}$, then the convergence to the set of KKT points is guaranteed (see [53, Corollary 1]). □

As mentioned in Section 3.3.3, it is not possible to guarantee the global optimality of the solution as the problem in non-convex. Moreover, different initializations of $\mathbf{P}$ may lead to different limit points. This is examined through simulations in the following section where it is shown that appropriate initialization yields good convergence and rate performances.

3.6. Numerical Results and Analysis

In this section, the proposed algorithms are evaluated through simulations. First, we present the simulation parameters before discussing various results.

3.6.1. Simulation Parameters

We consider a system with $K = N_{\mathrm{t}}$ users, where K is varied and highlighted throughout the section. The channel $\mathbf{H}$ has i.i.d. complex Gaussian entries with unit variance, i.e. drawn from $\mathcal{CN}(0,1)$. The noise variance is fixed as $\sigma_{\mathrm{n}}^2 = 1$, from which the long-term SNR is given as P_{t}. Entries of $\widetilde{\mathbf{H}}$ are also i.i.d complex Gaussian drawn from $\mathcal{CN}\left(0, \sigma_{\mathrm{e}}^2\right)$, where $\sigma_{\mathrm{e}}^2 = N_t^{-1}\sigma_{\mathrm{e},k}^2$, $\forall k \in \mathcal{K}$. The error variance is given as $\sigma_{\mathrm{e}}^2 = \beta P_{\mathrm{t}}^{-\alpha}$, where $\beta \geq 0$ and $\alpha \in [0,1]$ are varied to represent different CSIT accuracies and SNR scalings. It follows that the channel estimate $\widehat{\mathbf{H}} = \mathbf{H} - \widetilde{\mathbf{H}}$ also has entries with distribution $\mathcal{CN}\left(0, 1-\sigma_{\mathrm{e}}^2\right)$.

The channel realization $\mathbf{H}$ should not be confused with a conditional realization $\mathbf{H}^{(m)}$. While the former represents the actual fading state experienced by the users and unknown to the BS, the latter is part of a sample $\mathbb{H}^{(M)}$ available at the BS and used to calculate the SAFs. The size of the sample is set to $M = 1000$ throughout the simulations, unless otherwise stated. For a given estimate $\widehat{\mathbf{H}}$, the mth conditional realization is obtained as $\mathbf{H}^{(m)} = \widehat{\mathbf{H}} + \widetilde{\mathbf{H}}^{(m)}$, where $\widetilde{\mathbf{H}}^{(m)}$ is drawn from the error distribution.

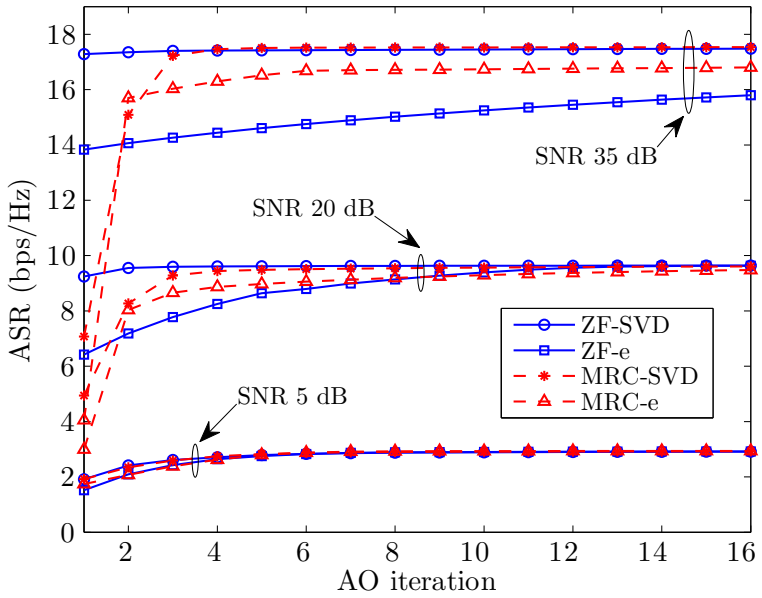


Figure 3.1.: ASR convergence of Algorithm 3.2 using four different initializations of $\mathbf{P}$ for one randomly generated channel estimate. Other parameters are given by: $K, N_t = 2$, $\sigma_{\mathrm{e}}^2 = P_{\mathrm{t}}^{-0.6}$, and 5, 20 and 35 dB SNRs.

Remark 3.3. *All convex optimization problems requiring interior-point methods throughout this thesis are solved using the CVX toolbox [57].*

3.6.2. Convergence

First, we examine the convergence of Algorithm 3.2 under four different initializations of $\mathbf{P}$. Recalling from Section 2.4.1, precoding vectors can be expressed in terms of power allocation and beamforming direction as $\mathbf{p}_{\mathrm{c}} = \sqrt{q_{\mathrm{c}}}\widehat{\mathbf{p}}_{\mathrm{c}}$ and $\mathbf{p}_k = \sqrt{q_k}\widehat{\mathbf{p}}_k$.

The first initialization is denoted by ZF-e, in which ZF-BF over $\widehat{\mathbf{H}}$ is used for private precoders such that

$$\mathbf{P}_{\mathrm{p}} = \mathbf{P}_{\mathrm{p}}^{\mathrm{ZF}}\left(\widehat{\mathbf{H}}, [P_{\mathrm{t}}^{\alpha}/K, \ldots, P_{\mathrm{t}}^{\alpha}/K]\right) \tag{3.62}$$

while the common precoder has $q_{\mathrm{c}} = P_{\mathrm{t}} - P_{\mathrm{t}}^{\alpha}$ and $\widehat{\mathbf{p}}_{\mathrm{c}} = \mathbf{e}_1$. The second initialization, labeled as ZF-SVD, is a modification of ZF-e, where $\widehat{\mathbf{p}}_c$ is slightly optimized by choosing it as the dominant left singular vector of $\widehat{\mathbf{H}}$. The third and fourth initializations, labeled as MRC-e and MRC-SVD respectively, maintain the same common precoders as ZF-e and ZF-SVD respectively. However, Maximum Ratio Combining (MRC) given by $\widehat{\mathbf{p}}_k = \widehat{\mathbf{h}}_k/\|\widehat{\mathbf{h}}_k\|$ for all $k \in \mathcal{K}$, is employed instead of ZF-BF.

The ASR convergence for $\sigma_{\mathrm{e}}^2 = P_{\mathrm{t}}^{-0.6}$ and SNRs 5, 20 and 35 dB is shown in Figure 3.1. It is evident that the algorithm converges to a limit point regardless of the initialization. However, the speed of convergence is influenced by the initial state, which may also influence the limit point as the problem is non-convex. The initialization effect becomes more visible

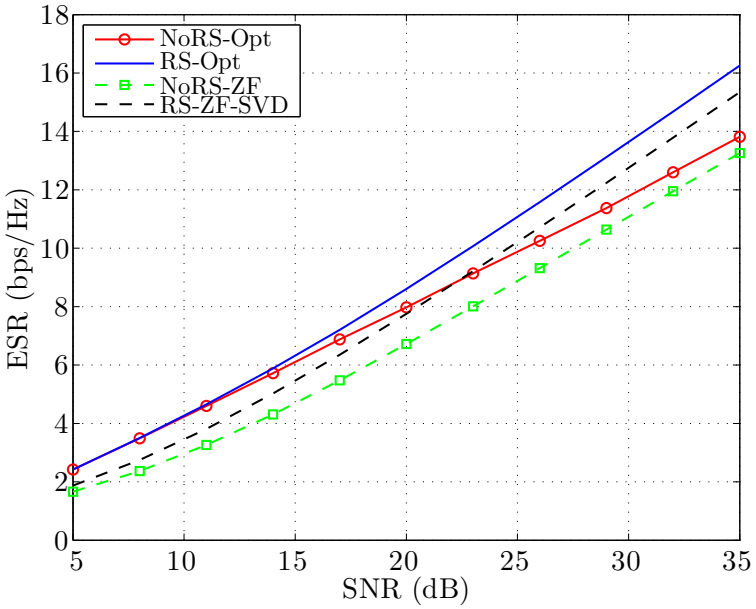


Figure 3.2.: Comparison between the ESRs of NoRS and RS transmission schemes obtained by averaging over 100 random channels. $K, N_t = 2$ and $\sigma_{\mathrm{e}}^2 = P_{\mathrm{t}}^{-0.6}$.

as SNR grows large. For example, initializing the common precoder using SVD enhances the convergence at high SNRs. In the following results, MRC-SVD is adopted as it provides good overall performance over various channel realizations and a wide range of SNRs.

3.6.3. Ergodic Sum-Rate Performance

Here we evaluate the achievable ESRs obtained by averaging the ASRs over 100 channel realizations. For each channel realization, Algorithm 3.2 is used to obtain the RS ASR. For NoRS, a modification of Algorithm 3.2 is used where the common precoder and common rates are switched off .

RS vs. NoRS

The performance of the optimized RS scheme (RS-Opt) is compared to the optimized NoRS scheme (NoRS-Opt). ZF-BF with Water-Filling (WF), where power allocation is carried out assuming that the estimate $\widehat{\mathbf{H}}$ is perfect, is considered as a baseline for NoRS transmission. This scheme is termed as NoRS-ZF. On the other hand, we consider a modified version of the ZF-SVD initialization in the previous subsection as a baseline for RS transmission. In particular, the power splitting between the common message and the private messages is maintained, while WF is used to allocate the power among the private messages. This scheme is termed RS-ZF-SVD.

The 2-user ESRs obtained using the different optimization schemes for $\sigma_{\mathrm{e}}^2 = P_{\mathrm{t}}^{-0.6}$ and $P_{\mathrm{t}}^{-0.9}$ are shown in Figure 3.2 and Figure 3.3, respectively. Comparing RS-Opt and NoRS-

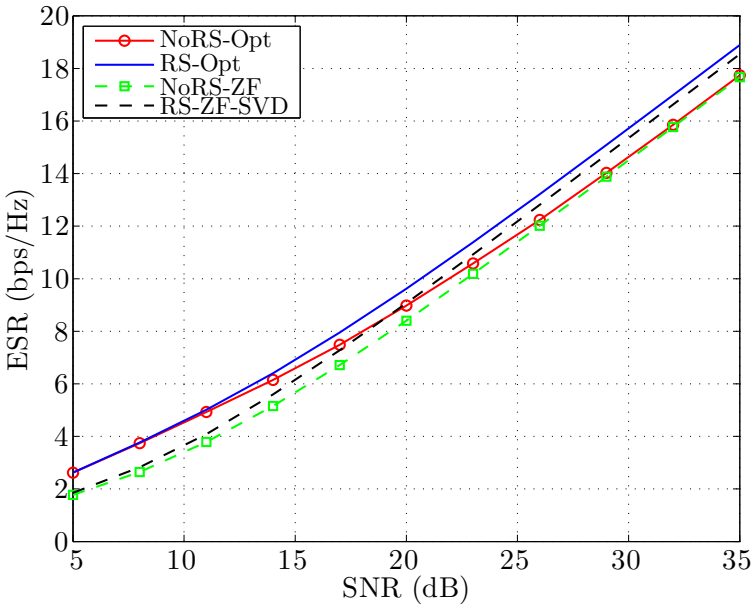


Figure 3.3.: Comparison between the ESRs of NoRS and RS transmission schemes obtained by averaging over 100 random channels. $K, N_t = 2$ and $\sigma_{\mathrm{e}}^2 = P_{\mathrm{t}}^{-0.9}$.

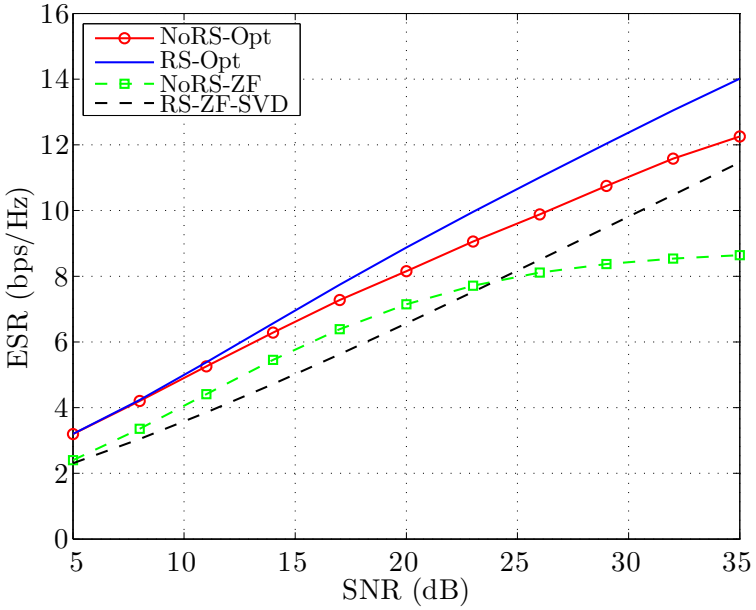


Figure 3.4.: Comparison between the ESRs of NoRS and RS transmission schemes obtained by averaging over 100 random channels. $K, N_t = 2$ and $\sigma_{\mathrm{e}}^2 = 0.063$.

Opt, the two schemes perform similarly at low SNRs. In particular, both reduce to single-user transmission by switching off the weaker user. As SNR grows, multiuser transmission starts to take over. However, reducing multiuser interference to the level of noise (or eliminating it) is not possible due to imperfect CSIT. Therefore, the overall performance does not necessarily benefit from additional power inaccurately directed towards a given user, as it may cause more damage by interfering with other users. At this stage, RS-Opt starts to part from NoRS-Opt by switching on the common message. The contribution of the common message primarily manifests at high SNRs, where the gap between RS-Opt and NoRS-Opt exceeds 4 dB for $\alpha = 0.6$. For $\alpha = 0.9$, which is almost perfect from a DoF perspective, RS is not as instrumental as it is for $\alpha = 0.6$. However, rate gains can still be observed at high SNRs. Regarding the baseline schemes, their inferiority compared to optimized schemes is evident over the entire SNR range. In the low SNR regime, baselines fall behind due to the strict application of ZF-BF, which is not ideal in this case. However, the RS-ZF-SVD scheme slightly compensates for this through the common message. However, it fails to match the optimized transmission in RS-Opt and NoRS-Opt. In the high SNR regime, each baseline scheme achieves the same DoF (slope of the curve) as its optimized counterpart with a rate gap, which is particularly noticeable for $\alpha = 0.6$.

The performance under CSIT errors that do not scale with SNR ($\alpha = 0$) is shown in Figure 3.4, where $\sigma_{\mathrm{e}}^2 = \left(10^{\frac{20}{10}}\right)^{-0.6} = 0.063$, which corresponds to the CSIT quality obtained at 20 dB SNR w.r.t Figure 3.2. First, it can be seen that the rate of NoRS-ZF saturates. This is due to the naive employment of $\widehat{\mathbf{H}}$ to design the ZF-BF vectors and allocate power as if it was a perfect estimate. At high SNRs, residual interference dominates and caps the performance. RS-ZF-SVD employs the DoF-motivated power splitting (i.e. $q_{\mathrm{c}} = P_{\mathrm{t}}$ and $q_k = 0$) and hence reduces to multicast transmission (sending only the common message) maintaining a DoF of 1. On the other hand, NoRS-Opt and RS-Opt schemes achieve significantly better performances compared to their corresponding baselines, over the entire range of SNRs. Moreover, although RS-Opt is not expected to achieve DoF gains over NoRS-Opt as $P_{\mathrm{t}} \to \infty$, the former still manages to deliver superior rate performance at medium and high SNRs.

Increased Number of Users

Theorem 3.1 implies that the relative DoF gain achieved by RS over NoRS decreases as K increases. Figure 3.5 shows the ESRs achieved by NoRS-Opt and RS-Opt for $K = 2$, 4, 6 and 8 users, assuming $\alpha = 0.6$. Figure 3.5 also shows the NoRS-Opt performance for a CSIT error exponent given as: $\alpha_2 = \frac{1+(K-1)\alpha}{K}$. This corresponds to a conventional system that achieves the same sum DoF as the RS system, but at the cost of higher CSIT quality requirements since $\alpha_2 > \alpha$, $\forall \alpha \in (0, 1)$. It is evident that the two perform closely, which highlights another benefit of RS, i.e. relaxed CSIT requirements compared to NoRS.

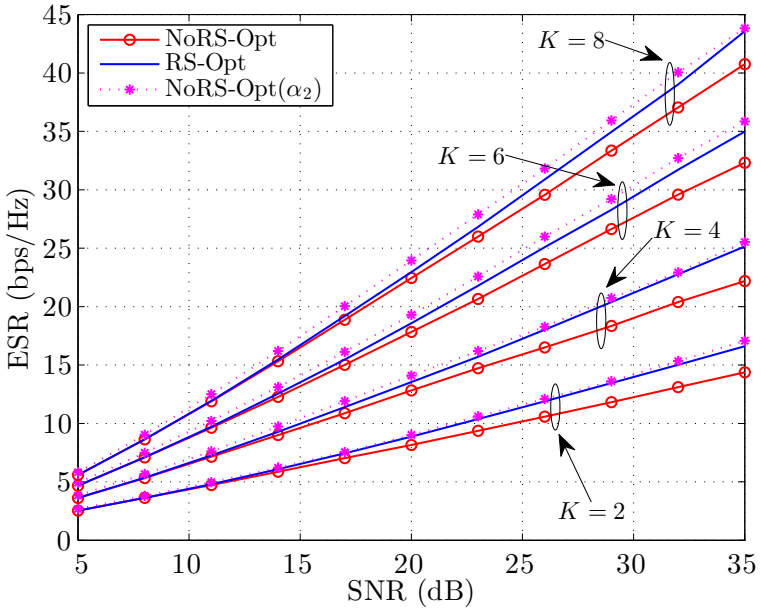


Figure 3.5.: Rate performance of NoRS-Opt and RS-Opt with $\sigma_{\mathrm{e}}^2 = P_{\mathrm{t}}^{-\alpha}$, and NoRS-Opt with $\sigma_{\mathrm{e}}^2 = P_{\mathrm{t}}^{-\alpha_2}$, where $\alpha_2 = \frac{1+(K-1)\alpha}{K}$, for $\alpha = 0.6$, and $K, N_t = 2$, 4,6 and 8.

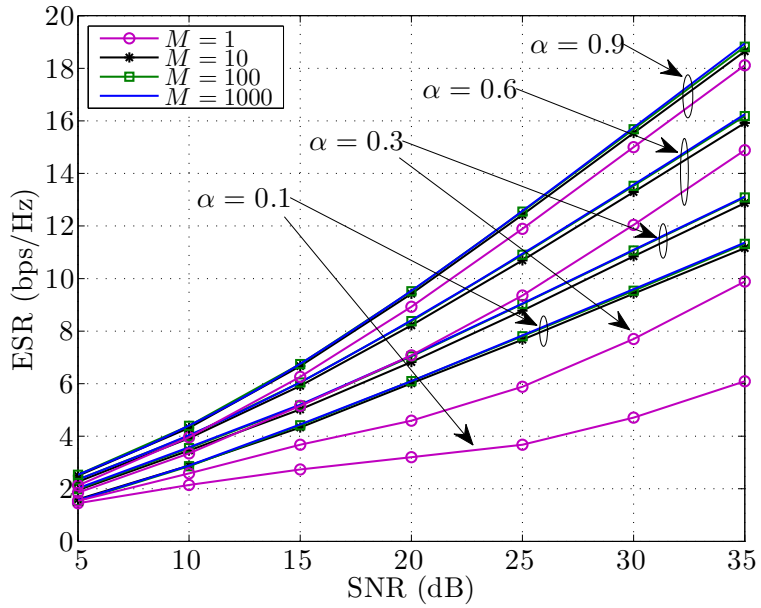


Figure 3.6.: The effect of changing the sample size M on the rate performance of the RS SAA algorithm for $K, N_t = 2$, $\sigma_{\mathrm{e}}^2 = P_{\mathrm{t}}^{-\alpha}$, and $\alpha = 0.1$, 0.3, 0.6 and 0.9.

SAA Sample Size

Next, we look at the influence of changing the sample size M on the SAA algorithm. The 2-user ESRs achieved by solving Algorithm 3.2 for $M = 1, 10, 100$ and 1000 are given in Figure 3.6. Different scaling CSIT qualities are considered where $\sigma_{\mathrm{e}}^2 = P_{\mathrm{t}}^{-\alpha}$, and $\alpha = 0.1, 0.3, 0.6$

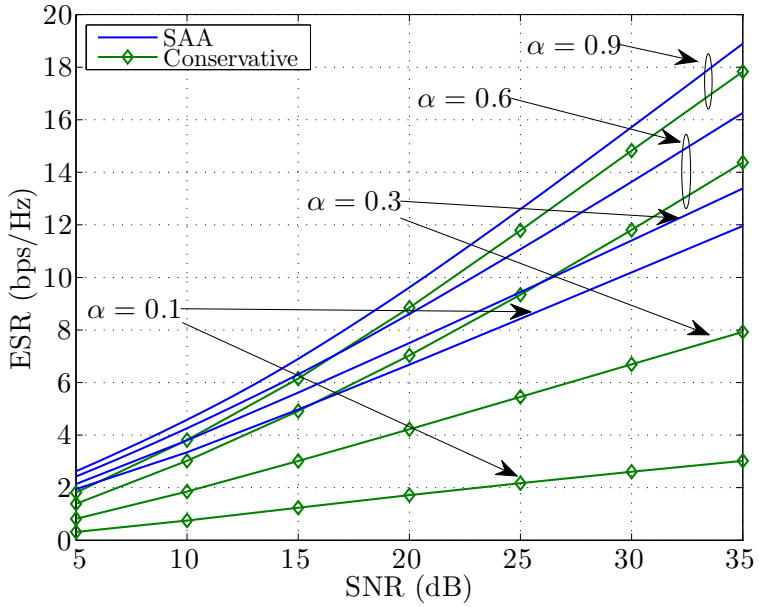


Figure 3.7.: Rate performance of the SAA and conservative approaches for $K, N_t = 2$, $\sigma_{\mathrm{e}}^2 = P_{\mathrm{t}}^{-\alpha}$, and $\alpha = 0.1$, 0.3, 0.6 and 0.9.

and 0.9. Using a sample of one realization significantly degrades the performance, as the resulting design lacks statistical knowledge of the CSIT uncertainty. On the other hand, the SAA algorithm performs very well with a sample size as small as $M = 10$ under the specified settings. Higher sample sizes of $M = 100$ and 1000 perform almost identically. Note that the achievable ESRs in this part are obtained by numerically evaluating the ERs, as taking the expectations of the sampled ASR objective functions for low Ms is not reflective of the achievable performance.

Conservative Algorithm

In this part, we look at the performance of the conservative design in Section **ref**. For the Gaussian CSIT error described earlier in this section, we have $\mathbf{R}_{\mathrm{e},k} = \sigma_{\mathrm{e}}^2\mathbf{I}$, $\forall k \in \mathcal{K}$. The conservative ESR is obtained by averaging the conservative ASRs obtained by solving (3.41). The ESRs predicted by the BS using the SAA approach and the conservative approach are given in Figure 3.7 for $K = 2$ and $\sigma_{\mathrm{e}}^2 = P_{\mathrm{t}}^{-\alpha}$, with $\alpha = 0.1, 0.3, 0.6$ and 0.9. The SAA scheme achieves significant gains over the conservative scheme, which increase with decreased CSIT qualities. This is explained by the discussion in Section 3.4.3. In particular, self-interference terms become dominant and the ASR lower-bound becomes looser. This effect is extremely detrimental for low CSIT qualities, as the achievable ESRs are highly undermined by the BS.

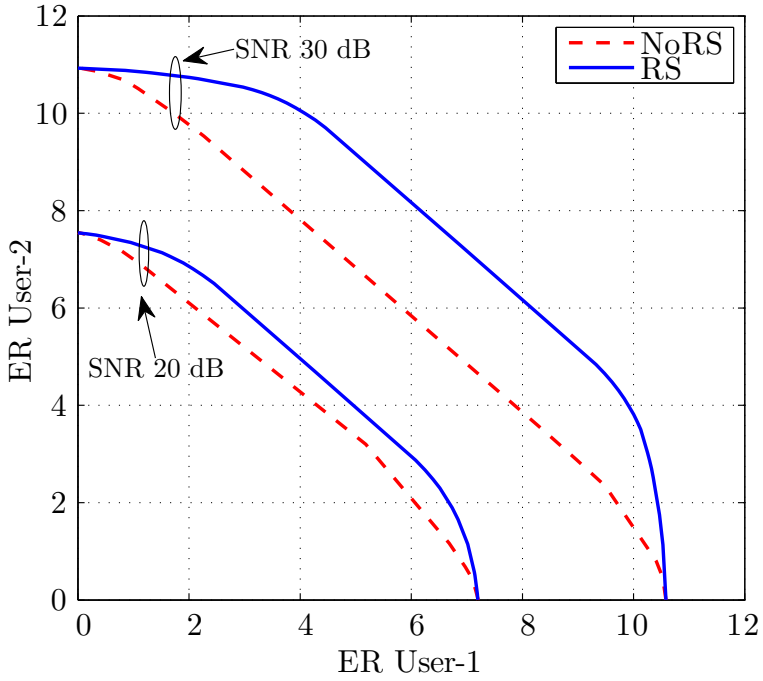


Figure 3.8.: Two-user ER regions for $\sigma_{\mathrm{e}}^2 = P_{\mathrm{t}}^{-0.6}$, and SNR = 20 and 30 dB.

3.6.4. Ergodic Rate Region

In the final part, we investigate the NoRS and RS achievable ER regions in a two-user scenario. ER regions are obtained by solving Weighted ESR (WESR) problems, where different boundary points are realized by varying the weights. For NoRS, this is achieved by incorporating the weights into the ASR problem in (3.6). For a fixed pair of weights, the problem is solved for several channel realizations. A boundary point of the ER region is obtained by averaging the resulting AR realizations for a given pair of weights. For the RS problem, applying the weights directly to (3.12) results in a region with the common message allocated to one user the entire time. A second region is obtained by allocating the common message to the other user, and the full region is given by the convex-hull enclosing the two regions by time-sharing of the extremity points.

Alternatively, all boundary points for the RS region can be obtained by assuming that $\bar{R}_{\mathrm{c}}$ is shared between users such that $\bar{C}_k$ is the kth user's portion of the common AR with $\sum_{k=1}^{K} \bar{C}_k = \bar{R}_{\mathrm{c}}$. The WASR problem is formulated as

$$\mathcal{R}_{\mathrm{WRS}}(P_{\mathrm{t}}) : \begin{cases} \max\limits_{\mathbf{P},\{\bar{C}_k\}_{k=1}^{K}} & \sum_{k=1}^{K} w_k(\bar{R}_k + \bar{C}_k) \\ \text{s.t.} & \bar{R}_{\mathrm{c},k} \geq \sum_{j=1}^{K} \bar{C}_j, \ \forall k \in \mathcal{K} \\ & \bar{C}_k \geq 0, \ \forall k \in \mathcal{K} \\ & \mathrm{tr}\left(\mathbf{P}\mathbf{P}^H\right) \leq P_{\mathrm{t}} \end{cases} \tag{3.63}$$

where the weights $\{w_k \mid k \in \mathcal{K}\}$ are fixed parameters. (3.63) is solved by formulating its equivalent AWSMSE problem, and modifying Algorithm 3.2 accordingly. To obtain the two-user regions shown in Figure 3.8, the corresponding problems are solved for 43 different pairs of weights. The first weight is fixed as $w_1 = 1$, and the second weight changes as $w_2 \in \{10^{-3}, 10^{-1}, 10^{-0.95}, \ldots, 10^{0.95}, 10, 10^3\}$. What is meaningful for each pair is the ratio between the weights rather than their absolute levels [7]. Each point on the ER region is characterized by the tuple $\mathbb{E}_{\widehat{\mathrm{H}}}\left\{(\bar{R}_1, \bar{R}_2)\right\}$ for NoRS, and $\mathbb{E}_{\widehat{\mathrm{H}}}\left\{(\bar{R}_1 + \bar{C}_1, \bar{R}_2 + \bar{C}_2)\right\}$ for RS, where each pair of ARs is obtained from solving the corresponding WASR problem.

The ER regions shown in Figure 3.8 are obtained for SNRs: 20 and 30 dB, and $\sigma_{\mathrm{e}}^2 = P_{\mathrm{t}}^{-0.6}$. As expected, the gap between RS and NoRS grows with SNR. The performance for 30 dB SNR is particularly interesting, as it is evident that RS enlarges the whole ER region significantly. For example, while guaranteeing an ER of 10 bps/Hz for user-1, an ER of almost 4 bps/Hz can be achieved by user-2. On the other hand, guaranteeing the same user-1 ER using NoRS restricts the ER of user-2 to just over 1.5 bps/Hz. This observation is of special interest for designs exploiting different points of the rate regions, e.g. max-min fairness and QoS constrained designs as seen in the following chapter.

3.7. Summary and Conclusion

In this chapter, ESR maximization in MU-MISO systems with partial CSIT was achieved by optimizing the linear precoding scheme of the RS transmission strategy. The precoding scheme is optimized such that a conditional ASR metric is maximized based on the available channel estimate. This was shown to be optimum in the ergodic sense, where the transmission spans a large number of channel states. The stochastic ASR problem was solved using the conservative WMMSE approach in [46, 47]. The limitation of this approach were identified and analyses. Then, an alternative non-conservative WMMSE algorithm based on the SAA method was proposed. The effectiveness of the proposed algorithm and the benefits of adopting the RS strategy were demonstrated through simulations, where the gains anticipated by the DoF analysis were achieved, in addition to an optimized performance across the entire range of SNRs. Finally, the two-user ER region was numerically obtained by solving a sequence of WASR problems. RS proved to enlarge the entire achievable ER region, drawing attention to the potentials of employing RS in other scenarios.

4. Fairness and Quality of Service

After seeing the benefits of combining the RS strategy with precoder optimization techniques, we extend this to a different scenario in this chapter. In particular, we consider a MU-MISO system with bounded CSIT uncertainty. This allows for non-ergodic transmission with a guaranteed worst-case performance (zero-outage). However, instead of targeting the overall system performance through sum-rate maximization, we consider the problem of achieving max-min fairness amongst users. Moreover, we consider the inverse problem of satisfying minimum Quality of Service (QoS) constraints.

The same design *methodology* used in the previous chapter is revisited here. In particular, we propose two WMMSE-based robust designs: a conservative design that ignores perfect CSIR, and a non-conservative design that samples the uncertainty region. However, due to the different nature of the CSIT uncertainty and worst-case performance considered here, the optimization algorithms developed in the previous chapter cannot be reused. Next, the problems considered in this chapter are introduced.

4.1. Introduction

While maximizing the sum-rate guarantees the best overall performance, it does not account for fairness amongst users. This may be less of an issue under symmetric ergodic fading where all users would eventually experience all fading states over the long transmission period, hence achieving symmetric performances. However, if long-term channel gains and CSIT qualities vary amongst users, then we may end up with an asymmetric long-term performance. This problem is even more pronounced when considering non-ergodic transmission under one joint channel state. Such channel state is drawn randomly and may favour one user over the others, yielding an unfair performance.

In this chapter, we depart from some of the main features of the previous chapter, namely: ergodic transmission and sum-rate maximization under a transmit power constraint. Instead, we consider non-ergodic transmission under one random joint channel state. Moreover, we consider rather different design objectivs, namely

- Maximizing the minimum rate among users subject to a total power constraint.
- Transmit power minimization subject to prescribed rate constraints.

The former is known in literature as the MMF problem or the rate-balancing problem, while the latter is known as the QoS problem or the power problem. The two problems

are known to be very closely related. For example, there is an inverse relation between the two when all users request the same rate in the second problem. Hence, we focus on the former throughout this chapter. Once the RS design is well established, it is extended to the QoS problem in Section 4.8.

4.2. Achieving Max-Min Fairness

Under a fixed channel state and in the presence of perfect CSIT, the problem of designing a linear precoder to achieve max-min fairness is given by

$$\begin{aligned} \max_{R,\mathbf{P}_{\mathrm{p}}} \quad & R \\ \text{s.t.} \quad & R_k \geq R, \ \forall k \in \mathcal{K} \\ & \mathrm{tr}\big(\mathbf{P}_{\mathrm{p}}\mathbf{P}_{\mathrm{p}}^H\big) \leq P_{\mathrm{t}} \end{aligned} \tag{4.1}$$

where R is a slack variables representing the minimum rate amongst users. It can be checked that at least one of the rate constraints is satisfied with equality at optimality.

Problem (4.1) assumes an equivalent formulation in terms of the SINRs or MMSEs, which follows from the one-to-one monotonic relationship between R_k, γ_k, and ε_k, observed from equations (2.21) and (3.21). Such formulations are much more tractable, leading to optimum solutions by means of conic programming [58, 59], or iterative and alternating optimization methods that exploit the downlink-uplink duality [60–62]. However, such tractability and optimality does not normally hold under imperfect CSIT formulations.

4.2.1. Bounded CSIT Errors

Before we proceed to introduce the robust version of the MMF problem in (4.1), we introduce the bounded uncertainty model adopted in this chapter. We recall that the kth user's channel vector writes as the sum of the channel estimate and an error vector such that $\mathbf{h}_k = \widehat{\mathbf{h}}_k + \widetilde{\mathbf{h}}_k$. It is assumed that $\widetilde{\mathbf{h}}_k$ is confined within an origin-centered ball with radius δ_k. Hence, the kth user's CSIT uncertainty region is expressed by

$$\mathbb{H}_k \triangleq \left\{ \mathbf{h}_k \mid \mathbf{h}_k = \widehat{\mathbf{h}}_k + \widetilde{\mathbf{h}}_k, \|\widetilde{\mathbf{h}}_k\| \leq \delta_k \right\}. \tag{4.2}$$

As far as the BS is aware, the actual channel vector $\mathbf{h}_k$ experienced by the receiver could be any value in $\mathbb{H}_k$, which is centered around the estimate $\widehat{\mathbf{h}}_k$.

It follows that the uncertainty region of the channel state $\mathbf{H}$ given an estimated $\widehat{\mathbf{H}}$ is defined as $\mathbb{H} \triangleq \mathbb{H}_1 \times \mathbb{H}_2 \times \ldots \times \mathbb{H}_K$. This uncertainty model is highly relevant in limited feedback systems, where each receiver estimates its channel vector through downlink training and then sends back a quantized version to the BS [15, 18, 19, 63]. The resulting quantization errors fall within a bounded region which may either be a closed ball, or is fully contained within one. It should also be highlighted that such bounded error model is

also applied to control the outage probability under unbounded CSIT errors [64–66].

4.2.2. Robustness in the Worst-Case Sense

In this chapter, we consider non-ergodic transmission over one random joint fading state, i.e. fixed $\{\mathbf{H}, \widehat{\mathbf{H}}\}$. As seen in the previous chapter, CSIT uncertainty prohibits the BS from considering the instantaneous rates as design metrics, or even guarantee reliable communication at such rates. Moreover, the average approach employed in the previous chapter is not applicable under non-ergodic transmission, as the instantaneous rates experienced by the receivers may be lower than the instantaneous ARs in (3.3) and (3.7)[1]. Alternatively, the bounded nature of CSIT uncertainty is exploited to redefine robustness in the worst-case sense [67]. In particular, for a given estimate $\widehat{\mathbf{H}}$ known to the BS, reliable communication is guaranteed under zero-outage probability with resect to CSIT uncertainty.

For given precoder $\mathbf{P}$ and channel estimate $\widehat{\mathbf{H}}$, let $R_{\mathbf{c},k}(\mathbf{h}_k)$ and $R_k(\mathbf{h}_k)$ be the kth user's achievable rates[2] under $\mathbf{h}_k \in \mathbb{H}_k$, where $\mathbb{H}_k$ is centered at $\widehat{\mathbf{h}}_k$. Since the rate functions are continuous and $\mathbb{H}_k$ is compact, then rates defined over such uncertainty regions are bounded and attain their extreme values. Therefore, from the BS's point of view, $R_{\mathrm{c},k}(\mathbf{h}_k)$ and $R_k(\mathbf{h}_k)$ are confined within bounded uncertainty regions, and reliable communication is guaranteed at the worst-case rates defined as follows.

Definition 4.1. *The kth user's worst-case achievable rates are defined as*

$$\bar{R}_{\mathrm{c},k} \triangleq \min_{\mathbf{h}_k \in \mathbb{H}_k} R_{\mathrm{c},k}(\mathbf{h}_k) \quad \textit{and} \quad \bar{R}_k \triangleq \min_{\mathbf{h}_k \in \mathbb{H}_k} R_k(\mathbf{h}_k). \tag{4.3}$$

It follows that transmitting the common message at the common rate given by

$$\bar{R}_{\mathrm{c}} \triangleq \min_{l} \{\bar{R}_{\mathrm{c},l}\}_{l=1}^{K} \tag{4.4}$$

guarantees that it is successfully decoded by all users.

Remark 4.1. *We use $\bar{R}_{\mathrm{c}}$, $\bar{R}_{\mathrm{c},k}$ and $\bar{R}_k$ to denote worst-case rates in this chapter. These should not be confused with the ARs in Chapter 3, defined using the same notation.*

4.2.3. Robust Max-Min Fairness

By employing Definition 4.1, the robust version of the MMF problem in (4.1) is given by

$$\mathcal{R}(P_{\mathrm{t}}): \begin{cases} \max\limits_{\bar{R}, \mathbf{P}_{\mathrm{p}}} & \bar{R} \\ \text{s.t.} & \bar{R}_k \geq \bar{R}, \ \forall k \in \mathcal{K} \\ & \mathrm{tr}\big(\mathbf{P}_{\mathrm{p}} \mathbf{P}_{\mathrm{p}}^H\big) \leq P_{\mathrm{t}} \end{cases} \tag{4.5}$$

[1] We recall that this approach can be used to guarantee reliable communication in the ergodic setting where each receiver experiences a long-term average of ARs as explained Chapter 3.

[2] While common rates are not used until the next section, we include their worst-case definition here.

where $\bar{R}$ is the minimum rate slack variable. Problem (4.5) is a semi-infinite program [68,69], with a finite number of optimization variables and an infinite number of constraints. This is shown by writing the set of rate constraints in (4.5) as

$$R_k(\mathbf{h}_k) \geq \bar{R},\ \forall \mathbf{h}_k \in \mathbb{H}_k, k \in \mathcal{K}. \tag{4.6}$$

In contrast to (4.1) where the minimum rate across users is maximized, the robust problem in (4.5) seeks to maximize the minimum worst-case rate, which is taken as the minimum across infinite sets of rates defined over the continuous uncertainty regions.

As in (4.1), problem (4.5) also assumes equivalent SINR and MMSE formulations. However, unlike the perfect CSIT case, different formulations here generally lead to different solutions. This can be regarded to the fact that the nice properties that allow solving problem in (4.1) to global optimality do not generally hold under the CSIT uncertainties in (4.5). Therefore, different approximations are usually employed for each formulation to obtain suboptimal solutions as seen in [70–76]. Such approximations vary in the degree of conservatism and computational complexity. One of the approximations, based on the $\mathcal{S}$-lemma [77] and Semidefinite Relaxation (SDR) [78], was shown to solve the problem optimally under certain conditions which have been partially identified in [73, 79].

All existing works consider the conventional NoRS transmission exhibited in the formulation of problem (4.5). Optimum max-min fair designs ultimately achieve balanced rates, as the contrary implies that power is not allocated across users in the *fairest* manner. Achieving symmetric rates requires a simultaneous increase in users' powers as P_{t} increases. Under CSIT errors that do not decay with increased SNR, e.g. fixed number of feedback bits, this is known to limit the rate performance [19]. In particular, rates saturate at high SNRs where multiuser interference becomes dominant, and cannot be completely dealt with due to CSIT imperfection. While robust designs enhance the performance by *tweaking* the precoder design and transmit at reliable rates, the inherent interference cannot be eliminated, and rates still saturate as observed in [74]. However, such limitation can be transcended through RS as demonstrated in the remainder of this chapter.

4.3. Achieving Max-Min Fairness through Rate-Splitting

Here, we formulate a robust MMF optimization problem based on RS. The RS strategy is first generalized such that all users can (potentially) benefit from the common message.

The sum-DoF gains achieved through RS as shown in Section 2.5.3 are indifferent to which user's message is being split. This observation holds for sum-rate maximization in Chapter 3. In particular, the RS sum-rate question is posed as whether splitting is required or not, and in case the answer is in the affirmative, how much of the total information should be relayed by the common message, regardless of which user message(s) is split.

Under a MMF design, allocating the entire common message to one user could be an unfair strategy. However, the fact that the common message is decoded by all users could

be exploited by sharing the rate of such message. This is shown next.

4.3.1. A Generalized Rate-Splitting Strategy

We recall that the BS wishes to communicate K independent messages $W_1, \ldots, W_K$ to receivers $1, \ldots, K$ respectively. The RS scheme is extended such that the message of each user is split into a common part and a private part, i.e. $W_k = \{W_{k0}, W_{k1}\}$ with W_{k0} and W_{k1} denoting the common and private parts, respectively. A super message (known as the common message) is composed by packing the common parts such that $W_{\mathrm{c}} = \{W_{10}, \ldots, W_{K0}\}$. The resulting $K+1$ messages are encoded into independent data streams such that

$$W_{\mathrm{c}}, W_{11}, \ldots, W_{K1} \mapsto \big\{s_{\mathrm{c}}(t)\big\}_{t=1}^{T}, \big\{s_1(t)\big\}_{t=1}^{T}, \ldots, \big\{s_K(t)\big\}_{t=1}^{T} \tag{4.7}$$

which are precoded and transmitted as described in the previous chapters. Decoding at each receiver is carried out as described before. The difference here is that the common message would (potentially) have a portion intended to each receiver.

The fraction of the common message allocated to the kth user is given by $|W_{k0}|/|W_{\mathrm{c}}|$, where $|W|$ denotes the length of a message W. Hence, as guaranteed by the BS, the kth user gets a portion of the worst-case common rate defined as

$$\bar{C}_k \triangleq \frac{|W_{k0}|}{|W_{\mathrm{c}}|}\bar{R}_{\mathrm{c}}. \tag{4.8}$$

This can be thought of as the rate at which W_{k0} is communicated. An obvious condition is $\sum_{k=1}^{K}\bar{C}_k = \bar{R}_{\mathrm{c}}$, as the sum of all portions cannot exceed the rate of the common message. Therefore, the kth user's worst-case total achievable rate is given as $\bar{R}_k + \bar{C}_k$, corresponding to the rate at which the original message W_k is communicated.

In the light of the generalized RS strategy, robust max-min fairness is achieved by solving

$$\mathcal{R}_{\mathrm{RS}}(P_{\mathrm{t}}): \begin{cases} \max\limits_{\bar{R}_{\mathrm{t}}, \bar{\mathbf{c}}, \mathbf{P}} & \bar{R}_{\mathrm{t}} \\ \text{s.t.} & \bar{R}_k + \bar{C}_k \geq \bar{R}_{\mathrm{t}},\ \forall k \in \mathcal{K} \\ & \bar{R}_{\mathrm{c},k} \geq \sum_{l=1}^{K}\bar{C}_l,\ \forall k \in \mathcal{K} \\ & \bar{C}_k \geq 0,\ \forall k \in \mathcal{K} \\ & \mathrm{tr}\big(\mathbf{P}\mathbf{P}^H\big) \leq P_{\mathrm{t}} \end{cases} \tag{4.9}$$

where $\bar{R}_{\mathrm{t}}$ is a slack variable corresponding to the minimum user rate, and $\bar{\mathbf{c}} \triangleq [\bar{C}_1, \ldots, \bar{C}_K]^T$. The constraint $\bar{C}_k \geq 0$ guarantees non-negative portions, $\bar{R}_{\mathrm{c},k} \geq \sum_{l=1}^{K}\bar{C}_l$ guarantees the decodability of the common message, and $\bar{R}_k + \bar{C}_k \geq \bar{R}_{\mathrm{t}}$ guarantees fairness across users. The fact that each user's rate consists of two parts is clearly reflected in the formulation.

It is evident that solving (4.5) is equivalent to solving (4.9) over a restricted domain characterized by setting $\bar{\mathbf{c}} = \mathbf{0}$, which in turn forces $\|\mathbf{p}_{\mathrm{c}}\|^2$ to zero at optimality. As a result,

we can write $\mathcal{R}_{\text{RS}}(P_{\text{t}}) \geq \mathcal{R}(P_{\text{t}})$. While this tells us that at optimality, RS cannot perform worse than NoRS, it does not answer the following question: how much improvement does RS bring to the performance? if any at all. Analytically, we answer this question (in part) through DoF analysis in the following section.

4.4. Performance Limits

In this section, we explore the performance limits of the conventional (NoRS) design obtained by solving (4.5) and the RS design obtained by solving (4.9). The performance is characterized in terms of the DoF. In particular, we derive the MMF-DoF achieved by optimally designed NoRS and RS schemes in the presence of CSIT uncertainty that scales with SNR. We start by highlighting the CSIT scaling laws.

4.4.1. CSIT Scaling

Incorporating the concept of CSIT scaling in Section 2.2.2 into the bounded error model yields uncertainty regions that shrink with increased SNR. In particular, we write

$$\delta_k^2 = O(P_{\text{t}}^{-\alpha_k}) \tag{4.10}$$

where $\alpha_k \in [0,1]$ is the kth user's scaling factor defined as $\alpha_k \triangleq \lim_{P_{\text{t}}\to\infty} -\frac{\log(\delta_k^2)}{\log(P_{\text{t}})}$. Note that (4.10) follows from the scaling of the average in (2.8), which implies that the maximum $\|\widetilde{\mathbf{h}}_k\|^2$ in that average must scale as $O(P_{\text{t}}^{-\alpha_k})$.

The scaling law in (4.10) is explained as follows. Assume that we have a limited feedback system with a scaling number of feedback bits [19]. As SNR increases, the resolution of channel quantization is increased by using larger codebooks, resulting in more densely packed quantized channels (or channel estimates). Equivalently, quantization errors become smaller, and CSIT uncertainty regions shrink with increased SNR as captured by (4.10).

One difference compared to the previous chapter is that we emphasise the case where CSIT qualities may vary across users and $\alpha_1, \ldots, \alpha_K$ are not equal in general. This may correspond to different receiver capabilities or uplink channel qualities. We recall the ascending order of qualities in Assumption 2.1, i.e. $\alpha_1 \leq \alpha_2 \leq \ldots \leq \alpha_K$.

4.4.2. DoF Analysis

To facilitate analytic derivations of the DoF, we start with the following assumptions regarding the channel state and its estimate.

Assumption 4.1. *We assume that:*

1. *Entries of the channel state* $\mathbf{H}$ *(i.e. the fading coefficients) have absolute values bounded away from zero and infinity, to avoid degenerate situation.*
2. *The channel estimate matrix* $\widehat{\mathbf{H}} \triangleq \left[\widehat{\mathbf{h}}_1, \ldots, \widehat{\mathbf{h}}_K\right]$ *is of full column rank.*

The first assumption is not considered as a major restriction even when entries are drawn from unbounded distributions. In particular, omitted neighborhoods can be reduced to an arbitrary small probability measure with a vanishing impact on the DoF [20]. The second assumption can be guaranteed in a feedback system by prohibiting the scheduling of users with similar (or linearly dependent) quantized channel vectors in the same time-frequency slot. However, it should be noted that while this is required for the proof of the following result, is not necessary for the optimization solutions in the following sections, as problems (4.9) and (4.5) can still be solved under linearly dependent channel estimates.

Theorem 4.1. For the NoRS problem in (4.5), the optimum MMF-DoF is given by

$$\bar{d}^* \triangleq \lim_{P_\mathrm{t}\to\infty} \frac{\mathcal{R}(P_\mathrm{t})}{\log(P_\mathrm{t})} = \frac{\alpha_1+\alpha_2}{2} \tag{4.11}$$

while the RS problem in (4.9) yields an optimum MMF-DoF of

$$\bar{d}^*_{\mathrm{RS}} \triangleq \lim_{P_\mathrm{t}\to\infty} \frac{\mathcal{R}_{\mathrm{RS}}(P_\mathrm{t})}{\log(P_\mathrm{t})} = \min_{J\in\{2,\ldots,K\}} \frac{1+\sum_{k=1}^{J-1}\alpha_k}{J}. \tag{4.12}$$

Similar to Theorem 3.1, the results are obtained through two steps. First, we show that the max-min DoFs are upper-bounded by (4.11) and (4.12), then we show that the upper-bounds are achievable via feasible precoding schemes. However, the proof is more involved in this case due to the considered MMF-DoF and the variation in CSIT qualities across users. The full proof is relegated to Appendix B.1.

It should be noted that while the sum-DoF of the MISO-BC with imperfect CSIT has received considerable attention in literature, the MMF-DoF (also known as the symmetric-DoF [27, 80]) has been less treated in general. Hence, while the sum-DoF result in Theorem 3.1 is known, the MMF-DoF in Theorem 4.1 does not appear in prior work.

4.4.3. Insight

In the following, we take a closer look at Theorem 4.1 to gain some insight into the influence of CSIT uncertainty on the MMF performance, and the role of RS. To facilitate the analysis, we define the worst-case DoF achieved by a given precoding scheme as:

$$\bar{d}_\mathrm{c} \triangleq \lim_{P_\mathrm{t}\to\infty} \frac{\bar{R}_\mathrm{c}(P_\mathrm{t})}{\log_2(P_\mathrm{t})} \quad \text{and} \quad \bar{d}_k \triangleq \lim_{P_\mathrm{t}\to\infty} \frac{\bar{R}_k(P_\mathrm{t})}{\log_2(P_\mathrm{t})} \tag{4.13}$$

where dependencies on the power level are highlighted in (4.13). The portion of $\bar{d}_\mathrm{c}$ allocated to the kth user is given by $\bar{c}_k \triangleq \lim_{P_\mathrm{t}\to\infty} \frac{\bar{C}_k(P_\mathrm{t})}{\log_2(P_\mathrm{t})}$, where $\sum_{k=1}^{K} \bar{c}_k = \bar{d}_\mathrm{c}$. In the following, the worst-case DoF is simply referred to as the DoF for brevity. All definitions extend to the NoRS case where the common part is discarded.

The optimum MMF-DoF is given by

$$\bar{d}^* = \max_{\left\{\mathbf{P}_{\mathrm{p}}(P_{\mathrm{t}})\right\}_{P_{\mathrm{t}}}} \min_{k\in\mathcal{K}} \bar{d}_k \quad \text{and} \quad \bar{d}^*_{\mathrm{RS}} = \max_{\left\{\mathbf{P}(P_{\mathrm{t}})\right\}_{P_{\mathrm{t}}}} \min_{k\in\mathcal{K}} \bar{c}_k + \bar{d}_k \tag{4.14}$$

where maximization is over the precoding scheme, which maps each power level to a precoding matrix. Before we proceed, we should highlight that the MMF-DoF maximization problems in (4.14) do not assume unique solutions in general. In particular, there exists infinitely many precoding schemes that are optimum in a DoF sense. The set of DoF-optimum solutions contains the rate-optimum precoding scheme, obtained by solving a sequence of problem (4.5) for NoRS or problem (4.9) for RS, as rate optimality implies DoF optimality. However, the converse does not hold in general, e.g. the ZF-BF strategy is rate-suboptimal in general yet DoF-optimum under perfect CSIT as seen in Section 2.4. This was observed in the achievability of the result in Theorem 3.1, and appears in the proof of Theorem 4.1, where it is shown that the optimum DoF performance is achieved via suboptimal precoders in the rate sense.

NoRS MMF-DoF

It can be seen from (4.11) that the NoRS optimum MMF-DoF is determined by the worst two CSIT scaling factors. This is explained as follows.

First, consider switching off all receivers except for user-1 and user-2. It can be shown that a simultaneous and proportional increase in powers allocated to the two users combined with a proper design of precoders achieves $\bar{d}_1 = \alpha_1$ and $\bar{d}_2 = \alpha_2$. While this is rigorously shown in Lemma B.2 in Appendix B, it can be intuitively observed by using ZF-BF, which results in residual interference experience by user-1 and user-2 that scales as $O(P_{\mathrm{t}}^{1-\alpha_1})$ and $O(P_{\mathrm{t}}^{1-\alpha_2})$ respectively. By compromising user-2's DoF through power control, i.e. reducing the corresponding power scaling, the interference experienced by user-1 is reduced. It can be shown that the fairest DoF achieved through power control is given by $\frac{\alpha_1+\alpha_2}{2}$.

Now, moving back to the K-user case by introducing users (with possibly better CSIT qualities) does not improve the MMF-DoF. In particular, even if the new users can achieve higher DoF (ultimately 1), the MMF-DoF is limited by that of user-1 and user-2, i.e. $\frac{\alpha_1+\alpha_2}{2}$. Hence, the best they could do is achieve an equal (or higher) DoF without influencing $\bar{d}_1$ and $\bar{d}_2$, which can be shown to be possible given their better CSIT qualities.

RS MMF-DoF

Considering the same 2-user scenario under RS, it can be shown that allocating powers that scale as $O\big(P_{\mathrm{t}}^{\alpha_1}\big)$ to both private streams, and a common power that scales as $O\big(P_{\mathrm{t}}\big)$, achieves the DoF given by: $\bar{d}_1, \bar{d}_2 = \alpha_1$ and $\bar{d}_{\mathrm{c}} = 1 - \alpha_1$. The main idea is similar to the sum-DoF achievability scheme in Section 2.5 while assuming that $\alpha = \alpha_1$, and more details are given in Lemma B.2. Splitting $\bar{d}_{\mathrm{c}}$ evenly, each user achieves a total DoF of $\frac{1+\alpha_1}{2}$. It is

evident that this is strictly greater than $\frac{\alpha_1+\alpha_2}{2}$ achieved with NoRS given that $\alpha_2 < 1$.

Increasing K may decrease this max-min DoF, as $\bar{d}_{\mathrm{c}}$ may be divided among a larger set of users. This is reflected in (4.12) through the minimization over subsets of users, which is further explained in the full proof in Appendix B. Observe that $\bar{d}^*_{\mathrm{RS}} > \bar{d}^*$ holds for all $\alpha_1, \ldots, \alpha_K \in [0, 1)$, i.e. RS provides a strict DoF improvement over NoRS.

It can be seen that the existence of two or more users with $\alpha_k = 0$ yields $\bar{d}^* = 0$. This explains the saturating MMF rate performance observed in [74]. On the contrary, $\bar{d}^*_{\mathrm{RS}}$ is lower-bounded by $\bar{d}^*_{\mathrm{RS}} \geq 1/K$ regardless of the CSIT scaling, and hence achieves an ever-growing MMF rate.

4.5. Conservative Approach

Having demonstrated the MMF-DoF gain of employing RS, we turn to designing a robust precoding scheme that can realize such high SNR gains, yet perform well over the entire range of SNRs. We start by invoking the Rate-WMMSE relationship in Lemma 3.1. This time, we define a worst-case version of this relationship given by

$$\bar{R}_{\mathrm{c},k} = 1 - \max_{\mathbf{h}_k \in \mathbb{H}_k} \min_{u_{\mathrm{c},k}, g_{\mathrm{c},k}} \xi_{\mathrm{c},k}\big(\mathbf{h}_k, g_{\mathrm{c},k}, u_{\mathrm{c},k}\big) \tag{4.15a}$$

$$\bar{R}_k = 1 - \max_{\mathbf{h}_k \in \mathbb{H}_k} \min_{u_k, g_k} \xi_k\big(\mathbf{h}_k, g_k, u_k\big). \tag{4.15b}$$

It is evident that the optimum equalizers and weights in (4.15) depend on the channel state $\mathbf{h}_k$, as we have previously seen. Hence, a WMSE formulation of problem (4.9) would yield a problem with infinitely many variables and constraints.

The techniques used in the previous chapter cannot be applied to the case here, as the nature of the error is different and the problem in (4.12) is semi-infinite. However, the same philosophy in general can be employed. In particular, we can use a conservative approximation which relies on relaxing the dependencies of the equalizers and weights on the actual fading state as we see in this section. Moreover, we develop a non-conservative method based on sampling the uncertainty region in the following section.

4.5.1. Conservative Worst-Case Approximation

The conservative approximation is carried out by swapping the order of the maximization (worst-case channel) and minimization (optimization) in (4.15) such that

$$\widehat{R}_{\mathrm{c},k} = 1 - \min_{\widehat{u}_{\mathrm{c},k}, \widehat{g}_{\mathrm{c},k}} \max_{\mathbf{h}_k \in \mathbb{H}_k} \xi_{\mathrm{c},k}\big(\mathbf{h}_k, \widehat{g}_{\mathrm{c},k}, \widehat{u}_{\mathrm{c},k}\big) \tag{4.16a}$$

$$\widehat{R}_k = 1 - \min_{\widehat{u}_k, \widehat{g}_k} \max_{\mathbf{h}_k \in \mathbb{H}_k} \xi_k\big(\mathbf{h}_k, \widehat{g}_k, \widehat{u}_k\big) \tag{4.16b}$$

where $(\widehat{g}_{\mathrm{c},k}, \widehat{g}_k)$ and $(\widehat{u}_{\mathrm{c},k}, \widehat{u}_k)$ denote the *abstracted* equalizers and weights. As we see later, the optimum equalizers and weights in (4.16) depend only on $\widehat{\mathbf{h}}_k$, in reminiscence

to the conservative approximation in Section (3.4). Taking (4.16a) for example, the same equalizer-weight pair $(\widehat{u}_{\mathrm{c},k}, \widehat{g}_{\mathrm{c},k})$ is employed for all $\mathbf{h}_k \in \mathbb{H}_k$, hence loosing its dependency on the actual realizations of the channel.

It can be seen that for any given precoder $\mathbf{P}$, we have the following relationship

$$\widehat{R}_{\mathrm{c},k} \leq \bar{R}_{\mathrm{c},k} \quad \text{and} \quad \widehat{R}_k \leq \bar{R}_k. \tag{4.17}$$

This follows from the fact that $\max_x \min_y f(x,y) \leq \min_y \max_x f(x,y)$ for any function $f(x,y)$, which is shown in [45, footnote 1]. Applying this to (4.16) and (4.15), the relationship in (4.17) is obtained. Hence, by employing such approximation, we arrive at a conservative WMSE formulation given by

$$\widehat{\mathcal{R}}_{\mathrm{RS}}(P_{\mathrm{t}}): \begin{cases} \max\limits_{\widehat{R}_{\mathrm{t}}, \widehat{\mathbf{c}}, \mathbf{P}, \widehat{\mathbf{g}}, \widehat{\mathbf{u}}} & \widehat{R}_{\mathrm{t}} \\ \text{s.t.} & 1 - \xi_k\big(\mathbf{h}_k, \widehat{g}_k, \widehat{u}_k\big) + \widehat{C}_k \geq \widehat{R}_{\mathrm{t}}, \forall \mathbf{h}_k \in \mathbb{H}_k, k \in \mathcal{K} \\ & 1 - \xi_{\mathrm{c},k}\big(\mathbf{h}_k, \widehat{g}_{\mathrm{c},k}, \widehat{u}_{\mathrm{c},k}\big) \geq \sum_{l=1}^{K} \widehat{C}_l, \forall \mathbf{h}_k \in \mathbb{H}_k, k \in \mathcal{K} \\ & \widehat{C}_k \geq 0, \ \forall k \in \mathcal{K} \\ & \mathrm{tr}\big(\mathbf{P}\mathbf{P}^H\big) \leq P_{\mathrm{t}} \end{cases} \tag{4.18}$$

where $\widehat{\mathbf{g}} \triangleq \{\widehat{g}_{\mathrm{c},k}, \widehat{g}_k \mid k \in \mathcal{K}\}$ and $\widehat{\mathbf{u}} \triangleq \{\widehat{u}_{\mathrm{c},k}, \widehat{u}_k \mid k \in \mathcal{K}\}$. The conservatism of (4.18) follows from (4.17), as the worst-case rates that this problem yields are always upper-bounded by the actual achievable worst-case rates.

Problem (4.18) is semi-infinite with infinitely many WMSE constraint. Fortunately, (4.18) has an equivalent formulation with a finite number of constraints. Such formulation is obtained using the following result.

Lemma 4.1. *The infinite set of constraints given by*

$$1 - \xi_{\mathrm{c},k}\big(\mathbf{h}_k, \widehat{g}_{\mathrm{c},k}, \widehat{u}_{\mathrm{c},k}\big) \geq \sum_{l=1}^{K} \widehat{C}_l, \ \forall \mathbf{h}_k \in \mathbb{H}_k \tag{4.19}$$

are equivalently expressed by

$$\widehat{u}_{\mathrm{c},k}\big(\tau_{\mathrm{c},k} + |\widehat{g}_{\mathrm{c},k}|^2 \sigma_{\mathrm{n}}^2\big) - \log_2(\widehat{u}_{\mathrm{c},k}) \leq 1 - \widehat{R}_{\mathrm{c}} \tag{4.20a}$$

$$\begin{bmatrix} \tau_{\mathrm{c},k} - \lambda_{\mathrm{c},k} & \boldsymbol{\psi}_{\mathrm{c},k}^H & \mathbf{0}^T \\ \boldsymbol{\psi}_{\mathrm{c},k} & \mathbf{I} & -\delta_k \mathbf{P}^H \widehat{g}_{\mathrm{c},k}^H \\ \mathbf{0} & -\delta_k \widehat{g}_{\mathrm{c},k} \mathbf{P} & \lambda_{\mathrm{c},k} \mathbf{I} \end{bmatrix} \succeq 0 \tag{4.20b}$$

where $\boldsymbol{\psi}_{\mathrm{c},k}^H \triangleq \widehat{g}_{\mathrm{c},k} \widehat{\mathbf{h}}_k^H \mathbf{P} - \mathbf{e}_1^T$. *Similarly, the constraint given by*

$$1 - \xi_k\big(\mathbf{h}_k, \widehat{g}_k, \widehat{u}_k\big) + \widehat{C}_k \geq \widehat{R}_{\mathrm{t}}, \ \forall \mathbf{h}_k \in \mathbb{H}_k \tag{4.21}$$

is equivalently expressed by the finite constraints given by

$$\widehat{u}_k\big(\tau_k + |\widehat{g}_k|^2\sigma_{\mathrm{n}}^2\big) - \log_2(\widehat{u}_k) \leq 1 + \widehat{C}_k - \widehat{R}_{\mathrm{t}} \tag{4.22a}$$

$$\begin{bmatrix} \tau_k - \lambda_k & \boldsymbol{\psi}_k^H & \mathbf{0}^T \\ \boldsymbol{\psi}_k & \mathbf{I} & -\delta_k \mathbf{P}_{\mathrm{p}}^H \widehat{g}_k^H \\ \mathbf{0} & -\delta_k \widehat{g}_k \mathbf{P}_{\mathrm{p}} & \lambda_k \mathbf{I} \end{bmatrix} \succeq 0 \tag{4.22b}$$

where $\boldsymbol{\psi}_k^H \triangleq \widehat{g}_k \widehat{\mathbf{h}}_k^H \mathbf{P}_{\mathrm{p}} - \mathbf{e}_k^T$.

This is a known transformation which has been used in the robust MMSE optimization literature, e.g. [44, 45, 71, 72] just to name a few. The idea of the transformation is first to write the MSEs embedded in (4.19) and (4.21) as

$$\varepsilon_{\mathrm{c},k}\big(\mathbf{h}_k, \widehat{g}_{\mathrm{c},k}\big) = \| \widehat{g}_{\mathrm{c},k} \mathbf{h}_k^H \mathbf{P} - \mathbf{e}_1^T \|^2 + |\widehat{g}_{\mathrm{c},k}|^2 \sigma_{\mathrm{n}}^2 \tag{4.23}$$

$$\varepsilon_k\big(\mathbf{h}_k, \widehat{g}_k\big) = \| \widehat{g}_k \mathbf{h}_k^H \mathbf{P}_{\mathrm{p}} - \mathbf{e}_k^T \|^2 + |\widehat{g}_k|^2 \sigma_{\mathrm{n}}^2. \tag{4.24}$$

This is followed by applying the Schur Complement [81] to obtain Linear Matrix Inequality (LMI) forms. Finally, the result in [82, Proposition 2], which is based on the $\mathcal{S}$-lemma, is used. More details are found in the related literature.

4.5.2. Alternating Optimization Algorithm

By substituting (4.20) and (4.22) into (4.18), we arrive at a deterministic WMSE formulation which can be solving using the AO principle in Section 3.3.3.

First, $\widehat{\mathbf{g}}$ is optimized by solving the problems:

$$\min_{\widehat{g}_{\mathrm{c},k}} \max_{\mathbf{h}_k \in \mathbb{H}_k} \varepsilon_{\mathrm{c},k}\big(\mathbf{h}_k, \widehat{g}_{\mathrm{c},k}\big) \quad \text{and} \quad \min_{\widehat{g}_k} \max_{\mathbf{h}_k \in \mathbb{H}_k} \varepsilon_k\big(\mathbf{h}_k, \widehat{g}_k\big) \tag{4.25}$$

for all $k \in \mathcal{K}$. Note that updating $\widehat{\mathbf{g}}$ reduces to minimizing the worst-case MSEs, as weights are fixed in this step. Invoking the transformation in Lemma 4.1, the problems in (4.25) are rewritten as

$$\min_{\widehat{g}_{\mathrm{c},k}} \tau_{\mathrm{c},k} + |\widehat{g}_{\mathrm{c},k}|^2 \sigma_{\mathrm{n}}^2 \ \text{ s.t. (4.20b)} \quad \text{and} \quad \min_{\widehat{g}_k} \tau_k + |\widehat{g}_k|^2 \sigma_{\mathrm{n}}^2 \ \text{ s.t. (4.20b)}. \tag{4.26}$$

The problems in (4.26) have convex quadratic objective functions and semidefinite constraints, and can be posed as Semidefinite Programs (SDPs) by applying the Schur Complement to $|\widehat{g}_{\mathrm{c},k}|^2$ and $|\widehat{g}_k|^2$. Such problems can be solved efficiently using interior-point methods [52]. The resulting conservative MMSEs are denoted by $\widehat{\varepsilon}_k^{\mathrm{MMSE}}$ and $\widehat{\varepsilon}_{\mathrm{c},k}^{\mathrm{MMSE}}$.

Next, the weights are updated as $\widehat{u}_k = 1/\widehat{\varepsilon}_k^{\mathrm{MMSE}}$ and $\widehat{u}_{\mathrm{c},k} = 1/\widehat{\varepsilon}_{\mathrm{c},k}^{\mathrm{MMSE}}$. The optimality of this step follows by checking the first order optimality conditions of (4.16).

Finally, $(\widehat{R}_{\mathrm{t}}, \widehat{\mathbf{c}}, \mathbf{P})$ are updated by solving the problem given by

$$\widehat{\mathcal{R}}_{\mathrm{RS}}^{\mathrm{MMSE}}(P_{\mathrm{t}}, \widehat{\mathbf{g}}, \widehat{\mathbf{u}}): \begin{cases} \underset{\widehat{R}_{\mathrm{t}}, \widehat{\mathbf{c}}, \mathbf{P}}{\max} & \widehat{R}_{\mathrm{t}} \\ \text{s.t.} & (4.22) \text{ and } (4.22\text{b}), \forall k \in \mathcal{K} \\ & (4.20) \text{ and } (4.20\text{b}), \forall k \in \mathcal{K} \\ & \widehat{C}_k \geq 0, \ \forall k \in \mathcal{K} \\ & \mathrm{tr}\big(\mathbf{P}\mathbf{P}^H\big) \leq P_{\mathrm{t}} \end{cases} \tag{4.27}$$

formulated by fixing the equalizers and weights, obtained from the previous step, in problem (4.18). Problem (4.27) is also a SDP, which can be solved using interior-point methods. This procedure is repeated in an iterative manner as described in Algorithm 4.1.

Algorithm 4.1 is guaranteed to converge since the objective function of (4.18) is bounded above for a given power constraint, and increases monotonically at each iteration. The relationship in (4.17) guarantees that the solution obtained from Algorithm 4.1 is feasible for the original problem, i.e. (4.9). However, global optimality (w.r.t the conservative problem) cannot be guaranteed due to non-convexity.

Algorithm 4.1 Conservative WMMSE Alternating Optimization

1: **Initialize**: $n \leftarrow 0$, $\widehat{R}_{\mathrm{t}}^{[n]} \leftarrow 0$, $\mathbf{P}$
2: **repeat**
3: $\quad n \leftarrow n + 1$
4: $\quad \big(\widehat{g}_{\mathrm{c},k}, \widehat{g}_k\big) \leftarrow \big(\arg \widehat{\varepsilon}_{\mathrm{c},k}^{\mathrm{MMSE}}, \arg \widehat{\varepsilon}_k^{\mathrm{MMSE}}\big), \ \forall k \in \mathcal{K}$
5: $\quad \big(\widehat{u}_{\mathrm{c},k}, \widehat{u}_k\big) \leftarrow \big(1/\widehat{\varepsilon}_{\mathrm{c},k}^{\mathrm{MMSE}}, 1/\widehat{\varepsilon}_k^{\mathrm{MMSE}}\big), \ \forall k \in \mathcal{K}$
6: $\quad \big(\widehat{R}_{\mathrm{t}}^{[n]}, \widehat{\mathbf{c}}, \mathbf{P}\big) \leftarrow \arg \widehat{\mathcal{R}}_{\mathrm{RS}}^{\mathrm{MMSE}}(P_{\mathrm{t}}, \widehat{\mathbf{g}}, \widehat{\mathbf{u}})$
7: **until** $\big|\widehat{R}_{\mathrm{t}}^{[n]} - \widehat{R}_{\mathrm{t}}^{[n-1]}\big| < \epsilon_R$

4.5.3. Limitations

Similar to the conservative approach in Section 3.4, the one here also comes with limitations. In general, the cause of such limitation is the same in both cases: relaxing the dependencies of the equalizers and weights on perfect CSI. As in Section 3.4, this leads to an *ignorant* optimization of precoders, that completely ignores the presence of perfect CSIR.

It is well understood that the worst-case approach adopted in this chapter does not give any consideration to the statistical distribution of the CSIT errors. However, to facilitate the calculation of upper-bounds on the conservative worst-case rates in (4.16), let us assume that $\widetilde{\mathbf{h}}_k$ is drawn from an arbitrary distribution defined over an origin-centered ball with radius δ_k. For such distribution, let us compare the conservative worst-case Rate-WMMSE relationship in (4.16) with the conservative AR-AWMMSE in (3.40). We observe that conservative ARs in (3.40) are in fact upper-bounds on the conservative worst-case rates in (4.16). This follows from the fact that minimizing the AWMSE in (3.40) is upper-bounded

by the minimization of the maximum WMSE in (4.16). Hence, the limitations derived for the conservative approach in Section 3.4.3 apply here.

For example, non-scaling CSIT uncertainty yields self-interference that scales as the desired signal component. This explains the saturating performance in [45], which contradicts the DoF result there. This is also observed in the simulation results presented later in this chapter, where the conservative RS performance saturates, contradicting Theorem 4.1.

4.6. The Cutting-Set Method

In this section, we propose an algorithm that avoids conservative approximations by directly optimizing the worst-case achievable rates. This algorithm employs an iterative sampling-based procedure, known as the cutting-set method [10], which switches between two steps in each iteration: 1) optimization and 2) pessimization (worst-case analysis). In the optimization step, a sampled version of the semi-infinite problem (with a finite number of constraints) is solved over finite subsets of the uncertainty regions. In the pessimization step, worst-case analysis is carried out where channels that violate the constraints are determined and appended to the uncertainty subsets.

It was shown in [10] that the cutting-set algorithm converges to the global optimum solution of the original semi-infinite problem given that in each iteration, both the optimization and pessimization steps are globally solved[3]. To guarantee the tractability of the optimization step, the authors in [10] assume the convexity of the problem. A common approach when applying the cutting-set method to non-convex problems is to *convexify* the problem through conservative approximations, for example see [70, 83–85]. However, pessimization in these works is inexact, i.e. also approximated, and hence convergence to a feasible solution of the conservative problems is not even guaranteed according to [10].

To avoid conservative approximations, and hence a conservative performance, we directly apply the cutting-set method to the non-convex problem in (4.9). We show that as long as a KKT solution of the non-convex problem in (4.28) is obtained in each optimization step, the cutting-set algorithm converges to the set of KKT solutions of the original problem. Moreover, we guarantee such convergence through exact pessimization.

4.6.1. Cutting-Set Algorithm

For the ith iteration, let $\overline{\mathbb{H}}_k \triangleq \left\{\mathbf{h}_k^{(1)}, \ldots, \mathbf{h}_k^{(i_k)}\right\}$ and $\overline{\mathbb{H}}_{\mathrm{c},k} \triangleq \left\{\mathbf{h}_{\mathrm{c},k}^{(1)}, \ldots, \mathbf{h}_{\mathrm{c},k}^{(i_{\mathrm{c},k})}\right\}$ be the kth user's discretized channel uncertainty regions for the private rate and the common rate

[3] Certain conditions regarding the constraint functions and feasible set should also be satisfied. This is clarified in the proof of Proposition 4.1.

respectively. The corresponding sampled version of problem (4.9) writes as

$$\mathcal{R}_{\mathrm{RS}}^{(i)}(P_{\mathrm{t}}):\begin{cases}\max\limits_{\bar{R}_{\mathrm{t}},\bar{\mathbf{c}},\mathbf{P}} \bar{R}_{\mathrm{t}} \\ \text{s.t. } R_k\big(\mathbf{h}_k^{(j)}\big)+\bar{C}_k \geq \bar{R}_{\mathrm{t}}, \forall j \in \overline{\mathcal{I}}_k^{(i)}, k \in \mathcal{K} \\ \quad R_{\mathrm{c},k}\big(\mathbf{h}_{\mathrm{c},k}^{(j_{\mathrm{c}})}\big) \geq \sum_{l=1}^{K} \bar{C}_l, \forall j_{\mathrm{c}} \in \overline{\mathcal{I}}_{\mathrm{c},k}^{(i)}, k \in \mathcal{K} \\ \quad \bar{C}_k \geq 0,\ \forall k \in \mathcal{K} \\ \quad \mathrm{tr}\big(\mathbf{P}\mathbf{P}^H\big) \leq P_{\mathrm{t}}\end{cases} \tag{4.28}$$

where $\overline{\mathcal{I}}_k^{(i)} \triangleq \{1,\ldots,i_k\}$ and $\overline{\mathcal{I}}_{\mathrm{c},k}^{(i)} \triangleq \{1,\ldots,i_{\mathrm{c},k}\}$ are index sets. Note that $|\overline{\mathcal{I}}_k^{(i)}|, |\overline{\mathcal{I}}_{\mathrm{c},k}^{(i)}| \leq i$, as a maximum of one channel vector per set is added in each iteration.

Although R_k and $R_{\mathrm{c},k}$ are functions of the same channel vector ($\mathbf{h}_k$), the corresponding uncertainty region is sampled twice. This is due to the fact that the private and common messages are independently decoded, and hence worst-case constraints should be satisfied for each individually. This will become clearer when we discuss the pessimization step.

The optimization step involves solving the discretized problem in (4.28) yielding the point $\big(\bar{R}_{\mathrm{t}}^{(i)}, \bar{\mathbf{c}}^{(i)}, \mathbf{P}^{(i)}\big)$. Pessimization is then carried out to update the sampled uncertainty regions, by determining the channel vectors under which the rate constraints are most violated. The worst-case channels corresponding to the kth user's rates are obtained as

$$\mathbf{h}_k^* = \arg\min_{\mathbf{h}_k \in \mathbb{H}_k} R_k^{(i)}(\mathbf{h}_k) \quad \text{and} \quad \mathbf{h}_{\mathrm{c},k}^* = \arg\min_{\mathbf{h}_{\mathrm{c},k} \in \mathbb{H}_k} R_{\mathrm{c},k}^{(i)}(\mathbf{h}_{\mathrm{c},k}) \tag{4.29}$$

where $R_k^{(i)}$ and $R_{\mathrm{c},k}^{(i)}$ denote the rates when $\mathbf{P}^{(i)}$ is employed. The rate constraints in (4.28) are examined under the worst-case channels from (4.29). For example, if we have

$$R_k^{(i)}\big(\mathbf{h}_k^*\big) + \bar{C}_k^{(i)} < \bar{R}_{\mathrm{t}}^{(i)} \tag{4.30}$$

then $\mathbf{h}_k^*$ is appended to $\overline{\mathbb{H}}_k$. In a similar manner, if we have

$$R_{\mathrm{c},k}^{(i)}\big(\mathbf{h}_{\mathrm{c},k}^*\big) < \sum_{l=1}^{K} \bar{C}_l^{(i)} \tag{4.31}$$

then $\mathbf{h}_{\mathrm{c},k}^*$ is appended to $\overline{\mathbb{H}}_{\mathrm{c},k}$.

The cutting-set procedure is summarized in Algorithm 4.2. Defining the rate violations after the ith pessimization as

$$V_k^{(i)} \triangleq \bar{R}_{\mathrm{t}}^{(i)} - R_k^{(i)}\big(\mathbf{h}_k^*\big) - \bar{C}_k^{(i)} \quad \text{and} \quad V_{\mathrm{c},k}^{(i)} \triangleq \sum_{l \in \mathcal{K}} \bar{C}_l^{(i)} - R_{\mathrm{c},k}^{(i)}\big(\mathbf{h}_{\mathrm{c},k}^*\big) \tag{4.32}$$

the stopping criteria in Algorithm 4.2 is specified as a maximum violation, i.e.

$$\max_k \left\{\max\left\{V_k^{(i)}, V_{\mathrm{c},k}^{(i)}\right\}\right\}_{k \in \mathcal{K}} \leq \epsilon_V \tag{4.33}$$

Algorithm 4.2 Cutting-Set method.

Initialize: $i \leftarrow 0$, $i_k, i_{\mathrm{c},k} \leftarrow 1$ and $\overline{\mathbb{H}}_k, \overline{\mathbb{H}}_{\mathrm{c},k} \leftarrow \{\widehat{\mathbf{h}}_k\}, \forall k \in \mathcal{K}$
repeat
  $i \leftarrow i + 1$
  Optimization:
  $\left(\bar{R}_{\mathrm{t}}^{(i)}, \bar{\mathbf{c}}^{(i)}, \mathbf{P}^{(i)}\right) \leftarrow \arg \mathcal{R}^{(i)}(P_{\mathrm{t}})$
  Pessimization:
  For all $k \in \mathcal{K}$, **do**
  Obtain $\mathbf{h}_k^*, \mathbf{h}_{\mathrm{c},k}^*$ by solving (4.29)
  if $R_k^{(i)}(\mathbf{h}_k^*) + \bar{C}_k^{(i)} < \bar{R}_{\mathrm{t}}^{(i)}$ **then**
    $i_k \leftarrow i_k + 1$ and $\mathbf{h}_k^{(i_k)} \leftarrow \mathbf{h}_k^*$
    $\overline{\mathbb{H}}_k \leftarrow \{\overline{\mathbb{H}}_k, \mathbf{h}_k^{(i_k)}\}$
  end if
  if $R_{\mathrm{c},k}^{(i)}(\mathbf{h}_k^*) < \sum_{l=1}^{K} \bar{C}_l^{(i)}$ **then**
    $i_{\mathrm{c},k} \leftarrow i_{\mathrm{c},k} + 1$ and $\mathbf{h}_{\mathrm{c},k}^{(i_{\mathrm{c},k})} \leftarrow \mathbf{h}_{\mathrm{c},k}^*$
    $\overline{\mathbb{H}}_{\mathrm{c},k} \leftarrow \{\overline{\mathbb{H}}_{\mathrm{c},k}, \mathbf{h}_{\mathrm{c},k}^{(i_{\mathrm{c},k})}\}$
  end if
until stopping criteria is met

where $\epsilon_V > 0$ is some arbitrary tolerance constant.

Proposition 4.1. *Given that the optimization step in Algorithm 4.2 yields a KKT solution of the corresponding sampled problem in* (4.28)*, and the pessimization step is exact, i.e. the global minimizers in* (4.29) *are obtained, then the iterates generated by Algorithm 4.2 converge to the set of KKT solutions of the semi-infinite problem in* (4.9)*.*

The proof of Proposition 4.1 is relegated to Appendix B.3. We recall that if the KKT solution to the optimization step is a global optimum solution, then the iterates converge to a global optimum solution of problem (4.9) [10, Section 3.1].

In the following, the optimization and pessimization steps are thoroughly addressed. For ease of notation, the superscript (i) is dropped from the variables where is its understood that optimization and pessimization are part of a given iteration.

4.6.2. Optimization

The Rate-WMMSE relationship in Section 3.3.2 is revisited to transform the sampled problem (4.28) into an equivalent WMSE problem formulated as

$$\bar{\mathcal{R}}_{\mathrm{RS}}^{(i)}(P_{\mathrm{t}}):\begin{cases}\max\limits_{\bar{R}_{\mathrm{t}},\bar{\mathbf{c}},\mathbf{P},\mathbf{g},\mathbf{u}} \bar{R}_{\mathrm{t}} \\ \text{s.t. } 1-\xi_k\big(\mathbf{h}_k^{(j)},g_k^{(j)},u_k^{(j)}\big)+\bar{C}_k\geq\bar{R}_{\mathrm{t}},\forall j\in\overline{\mathcal{I}}_k^{(i)},k\in\mathcal{K} \\ \quad 1-\xi_{\mathrm{c},k}\big(\mathbf{h}_{\mathrm{c},k}^{(j_{\mathrm{c}})},g_{\mathrm{c},k}^{(j_{\mathrm{c}})},u_{\mathrm{c},k}^{(j_{\mathrm{c}})}\big)\geq\sum_{l=1}^{K}\bar{C}_l,\forall j_{\mathrm{c}}\in\overline{\mathcal{I}}_{\mathrm{c},k}^{(i)},k\in\mathcal{K} \\ \quad \bar{C}_k\geq 0,\ \forall k\in\mathcal{K} \\ \quad \mathrm{tr}\big(\mathbf{P}\mathbf{P}^H\big)\leq P_{\mathrm{t}}.\end{cases} \tag{4.34}$$

Note that $\mathbf{g}\triangleq\{\mathbf{g}_{\mathrm{c},k},\mathbf{g}_k\mid k\in\mathcal{K}\}$ is the sampled set of equalizers, with $\mathbf{g}_{\mathrm{c},k}\triangleq\left\{g_{\mathrm{c},k}^{(j_{\mathrm{c}})}\mid j_{\mathrm{c}}\in\overline{\mathcal{I}}_{\mathrm{c},k}^{(i)}\right\}$ and $\mathbf{g}_k\triangleq\left\{g_k^{(j)}\mid j\in\overline{\mathcal{I}}_k^{(i)}\right\}$, while $\mathbf{u}\triangleq\{\mathbf{u}_{\mathrm{c},k},\mathbf{u}_k\mid k\in\mathcal{K}\}$ is the sampled set of weights where $\mathbf{u}_{\mathrm{c},k}\triangleq\left\{u_{\mathrm{c},k}^{(j_{\mathrm{c}})}\mid j_{\mathrm{c}}\in\overline{\mathcal{I}}_{\mathrm{c},k}^{(i)}\right\}$ and $\mathbf{u}_k\triangleq\left\{u_k^{(j)}\mid j\in\overline{\mathcal{I}}_k^{(i)}\right\}$. Contrary to the conservative approach, the sampling of the equalizers and weights in (4.34) captures their dependencies on the actual channel, which reflects the availability of perfect CSIR.

The AO principle Section 3.3.3 is yet again employed, here to solve the sampled WMSE problem in (4.34). In a given iteration, $\mathbf{g}$ is first optimized by applying the MMSE solution in Lemma 3.1 to each equalizer in the sampled set, i.e. $g_{\mathrm{c},k}^{(j_{\mathrm{c}})}=g_{\mathrm{c},k}^{\mathrm{MMSE}(j_{\mathrm{c}})}$ and $g_k^{(j)}=g_k^{\mathrm{MMSE}(j)}$ for all $k\in\mathcal{K}$. Next, $\mathbf{u}$ is optimized in a similar manner using the solution in Lemma 3.1, i.e. $u_{\mathrm{c},k}^{(j_{\mathrm{c}})}=u_{\mathrm{c},k}^{\mathrm{MMSE}(j_{\mathrm{c}})}$ and $u_k^{(j)}=u_k^{\mathrm{MMSE}(j)}$ for all $k\in\mathcal{K}$. Finally, $(\bar{R}_{\mathrm{t}},\bar{\mathbf{c}},\mathbf{P})$ are updated by solving problem (4.34) for fixed $\mathbf{g}$ and $\mathbf{u}$. The resulting problem is convex with a linear objective function and quadratic and linear constraints, and can be efficiently solved using interior-point methods [52]. Following the same argument in Section 4.5.2, the AO algorithm described here is guaranteed to converge. The KKT optimality of the generated solution is established in the following result.

Proposition 4.2. *The iterates generated by the AO procedure described above converge to the set of KKT points of the ith sampled rate problem in* (4.28).

The proof is similar to the proof of Proposition 3.3, which is based on the work in [54]. Due to non-convexity, the KKT point obtained by the AO algorithm may be suboptimal. However, the effectiveness of this algorithm is demonstrated through simulation results.

4.6.3. Pessimization

For the outputs of the optimization step, the pessimization step determines whether the rate constraints are violated under an updated set of worst-case channels. This involves solving the problems in (4.29), which can be formulated in terms of the MMSEs due to

their monotonic relationship with the rates. The worst-case MMSEs are defined as

$$\bar{\varepsilon}_k^{\mathrm{MMSE}} \triangleq \max_{\mathbf{h}_k \in \mathbb{H}_k} \varepsilon_k^{\mathrm{MMSE}}(\mathbf{h}_k) \tag{4.35a}$$

$$\bar{\varepsilon}_{\mathrm{c},k}^{\mathrm{MMSE}} \triangleq \max_{\mathbf{h}_{\mathrm{c},k} \in \mathbb{H}_k} \varepsilon_{\mathrm{c},k}^{\mathrm{MMSE}}(\mathbf{h}_{\mathrm{c},k}). \tag{4.35b}$$

The private rate constraint is violated if we have $\bar{\varepsilon}_k^{\mathrm{MMSE}} > 2^{-(\bar{R}_{\mathrm{t}} - \bar{C}_k)}$, and the common rate constraint is violated given that $\bar{\varepsilon}_{\mathrm{c},k}^{\mathrm{MMSE}} > 2^{-\bar{R}_{\mathrm{c}}}$.

Solving the problems in (4.35) involves maximizing non-linear fractional functions over compact convex sets. Such problems can be solved using Dinkelbach's method [86], where the fractional program is transformed into a parametric auxiliary problem solved iteratively. Such transformation is based on the following result.

Lemma 4.2. *$\mathbf{h}_k^*$ and $\mathbf{h}_{\mathrm{c},k}^*$ are the global maximizers of problems* (4.35a) *and* (4.35b) *respectively if and only if they are the global maximizers of the parametric problems*

$$\mathcal{D}_k(\lambda_k) : \max_{\mathbf{h}_k \in \mathbb{H}_k} I_k - \lambda_k T_k \quad \textit{and} \quad \mathcal{D}_{\mathrm{c},k}(\lambda_{\mathrm{c},k}) : \max_{\mathbf{h}_{\mathrm{c},k} \in \mathbb{H}_k} I_{\mathrm{c},k} - \lambda_{\mathrm{c},k} T_{\mathrm{c},k} \tag{4.36}$$

respectively, when the parameters are given by

$$\lambda_k = \lambda_k^* \triangleq \bar{\varepsilon}_k^{\mathrm{MMSE}} \quad \textit{and} \quad \lambda_{\mathrm{c},k} = \lambda_{\mathrm{c},k}^* \triangleq \bar{\varepsilon}_{\mathrm{c},k}^{\mathrm{MMSE}}. \tag{4.37}$$

The above results follow directly from the theorem in [86], where a detailed proof is given. By substituting (4.37) into (4.36), we observe that

$$\mathcal{D}_k(\lambda_k^*) = 0 \quad \text{and} \quad \mathcal{D}_{\mathrm{c},k}(\lambda_{\mathrm{c},k}^*) = 0. \tag{4.38}$$

Therefore, solving the pessimization problems in (4.35) is equivalent to finding the zeros of the parametric auxiliary problems in (4.36). It turns out that the auxiliary problems have rather nice properties in this regard. Taking $\mathcal{D}_k(\lambda_k)$ for example, it is:

1. Continuous and strictly decreasing in λ_k.
2. $\mathcal{D}_k(\lambda_k) \to -\infty$ as $\lambda_k \to \infty$.
3. $\mathcal{D}_k(\lambda_k) \to \infty$ as $\lambda_k \to -\infty$.
4. $\mathcal{D}_k(\lambda_k) = 0$ has a unique solution λ_k^*.

The zero solution can be obtained using an iterative algorithm [86].

Note that it is commonly assumed that the fractional program is concave-convex, i.e. with a concave numerator and a convex denominator, which yields auxiliary problems that are convex. Nevertheless, it follows from the proof in [86] that this assumption is not necessary as long as the auxiliary problem can be solved globally for a given parameter. Next, we describe how to pessimization is carried out through Dinkelbach's algorithm assuming

Algorithm 4.3 Pessimization through Dinkelbach's Algorithm

Initialize: $\lambda_k^{(1)} \leftarrow 2^{-(\bar{R}_{\mathrm{t}}-\bar{C}_k)}$ and $m \leftarrow 0$
repeat
  $m \leftarrow m+1$
  obtain $\mathcal{D}_k\big(\lambda_k^{(m)}\big)$
  $\mathbf{h}_k^{(m)} \leftarrow \arg \mathcal{D}_k\big(\lambda_k^{(m)}\big)$
  $\lambda_k^{(m+1)} \leftarrow \varepsilon_k^{\mathrm{MMSE}}\big(\mathbf{h}_k^{(m)}\big)$
until $\mathcal{D}_k\big(\lambda_k^{(m)}\big) \leq \epsilon_{\mathcal{D}}$, or $m = m_{\max}$
if $m > 1$ **then**
  $\mathbf{h}_k^* \leftarrow \mathbf{h}_k^{(m)}$
  Output: $\mathbf{h}_k^*$
else
  Output: $\{\}$
end if

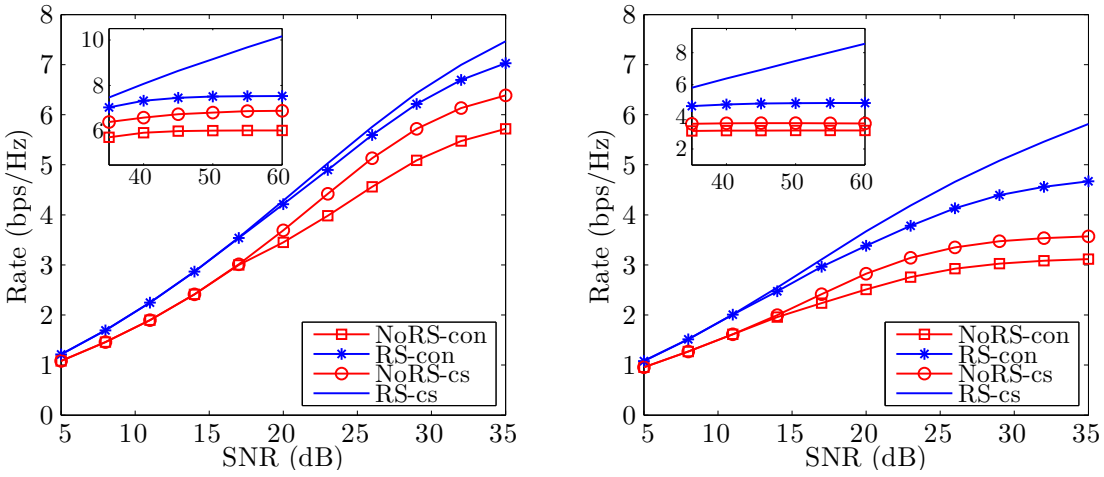
that the problems in (4.36) can be solved to global optimality for any given parameters. Before we proceed, we reemphasize that pessimization is carried out separately for $\varepsilon_k^{\mathrm{MMSE}}$ and $\varepsilon_{\mathrm{c},k}^{\mathrm{MMSE}}$ for all $k \in \mathcal{K}$, as the worst-case analysis is independent in each case. In the following, we describe the pessimization procedure for $\varepsilon_k^{\mathrm{MMSE}}$. This can be extended to all other private and common MMSEs.

First, the parameter is initialized as $\lambda_k^{(1)} = 2^{-(\bar{R}_{\mathrm{t}}-\bar{C}_k)}$, and $\mathcal{D}_k\big(\lambda_k^{(1)}\big)$ is obtained by solving the corresponding optimization problem. From the properties of $\mathcal{D}_k\big(\cdot\big)$, having $\mathcal{D}_k\big(\lambda_k^{(1)}\big) \leq 0$ implies that $\lambda_k^{(1)} \geq \bar{\varepsilon}_k^{\mathrm{MMSE}}$. In this case, the rate constraint is not violated and there is no need to proceed. Otherwise, the worst-case channel is updated as $\mathbf{h}_k^{(1)} = \arg \mathcal{D}_k\big(\lambda_k^{(1)}\big)$, and the parameter to be used in the next iteration is obtained as $\lambda_k^{(2)} = \varepsilon_k^{\mathrm{MMSE}}\big(\mathbf{h}_k^{(1)}\big)$. This is repeated until convergence. This procedure is summarized in Algorithm 4.3, where $\epsilon_{\mathcal{D}} > 0$ determines the accuracy of the solution, and $\{\}$ corresponds to an empty set, returned if the rate constraint is not violated. It follows directly from Lemma 4.2 that if the rate constraint is violated, i.e. pessimization has a non-empty solution, then $\lambda_k^{(m)}$ and $\mathbf{h}_k^{(m)}$ converge to $\bar{\varepsilon}_k^{\mathrm{MMSE}}$ and $\mathbf{h}_k^*$ respectively. For $\varepsilon_{\mathrm{c},k}^{\mathrm{MMSE}}$, the parameter is initialized as $\lambda_{\mathrm{c},k}^{(1)} = 2^{-\bar{R}_{\mathrm{c}}}$ and the same steps are followed yielding $\mathbf{h}_{\mathrm{c},k}^*$ if the common rate is violated, and $\{\}$ otherwise.

Now, it remains to show that the problems in (4.36) can be solved to global optimality. This is achieved as shown in Appendix B.4.

4.7. Numerical Results and Analysis

In this section, the performance of the proposed algorithms are assessed through simulations. A three-user system with $K = N_t = 3$ is considered throughout the simulations, unless stated otherwise. The noise variance is fixed as $\sigma_{\mathrm{n}}^2 = 1$. A given channel matrix $\mathbf{H}$ has i.i.d. entries drawn from the distribution $\mathcal{CN}(0,1)$. The corresponding estimate is obtained as $\widehat{\mathbf{H}} = \mathbf{H} - \widetilde{\mathbf{H}}$, where each error vector is drawn from a uniform distribution over



(a) $\delta = 0.05$ (b) $\delta = 0.15$

Figure 4.1.: Rate performance for $K, N_{\mathrm{t}} = 3$, and $\delta_1, \delta_2, \delta_3 = \delta$.

the corresponding uncertainty region with $\delta_k^2 = \beta_k P_{\mathrm{t}}^{-\alpha_k}$, where β_k is a constant.

We consider the conservative (con) and the cutting-set (cs) methods for both the NoRS and RS strategies, yielding four different designs: NoRS-con, NoRS-cs, RS-con and RS-cs. The NoRS-con and NoRS-cs results are obtained from Algorithm 4.1 and Algorithm 4.2 respectively, by discarding the common message. It should be noted that the NoRS-con design is equivalent to the MSE-based design in [71].

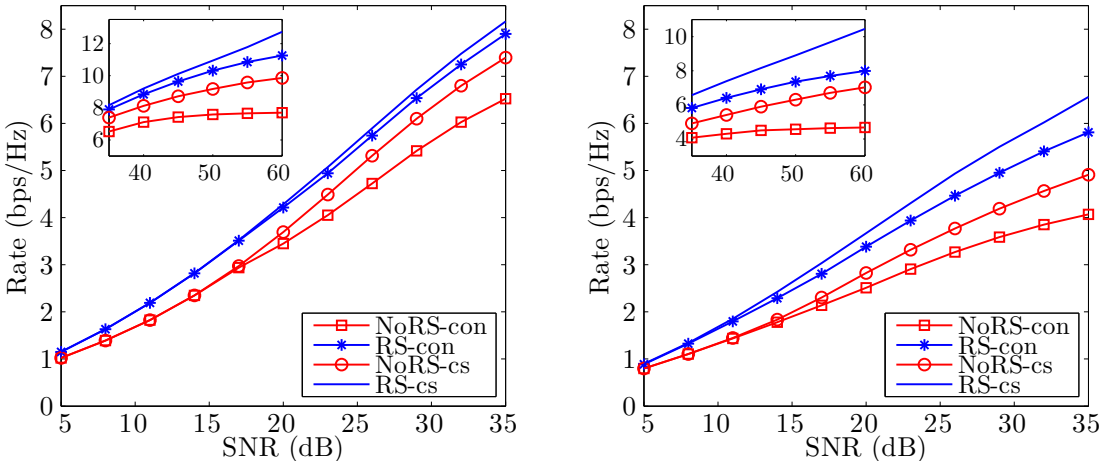
4.7.1. Max-Min Fair Rate Performance

First, we examine the robust max-min rate performance for the four aforementioned designs. The worst-case rates are averaged over 100 realizations of $\widehat{\mathbf{H}}$, where each estimate is obtained from an independent realization of $\mathbf{H}$. This is given for non-scaling CSIT error, scaling CSIT errors, and larger systems for a wide range of SNRs as we see next.

Non-scaling CSIT

Results for non-scaling CSIT errors (i.e. $\alpha_1, \alpha_2, \alpha_3 = 0$) are shown in Figure 4.1 with $\delta_1, \delta_2, \delta_3 = \delta$, where $\delta = 0.05$ and 0.15 for Figure 4.1a and Figure 4.1b, respectively.

For a given strategy (NoRS or RS), the cs design outperforms the con design, specifically in the intermediate and high SNR regimes. This gap grows larger with increased SNR and CSIT uncertainty, due to the increased influence of self-interference resulting from the conservative approximation. As expected from Theorem 4.1, NoRS schemes saturate as SNR grows large ($\bar{d}^* = 0$). This trend is also followed by the RS-con design, which at first glance seems to contradict the result in Theorem 4.1, yet can be explained in the light of the analysis in Section 4.5.3. On the other hand, the RS-cs design coincides with the result in (4.12) and achieves an ever growing rate performance. In particular, RS-cs achieves



(a) $\delta = 0.05$ (b) $\delta = 0.15$

Figure 4.2.: Rate performance for $K, N_{\mathrm{t}} = 3$, $\delta_1 = \delta$, and $\delta_2, \delta_3 = \delta\sqrt{10P_{\mathrm{t}}^{-0.5}}$.

MMF-DoFs of 0.31 and 0.33 for $\delta = 0.05$ and 0.15 respectively, obtained by scaling the high-SNR slopes of the rate curves in Figure 4.1. This is fairly close to the theoretically anticipated value, i.e. $\bar{d}^*_{\mathrm{RS}} = 0.333$.

In general, RS schemes give significant performance gains over their NoRS counterparts for the entire SNR range, with rate gains exceeding 20% and 60% for $\delta = 0.05$ and $\delta = 0.15$ respectively at high SNRs.

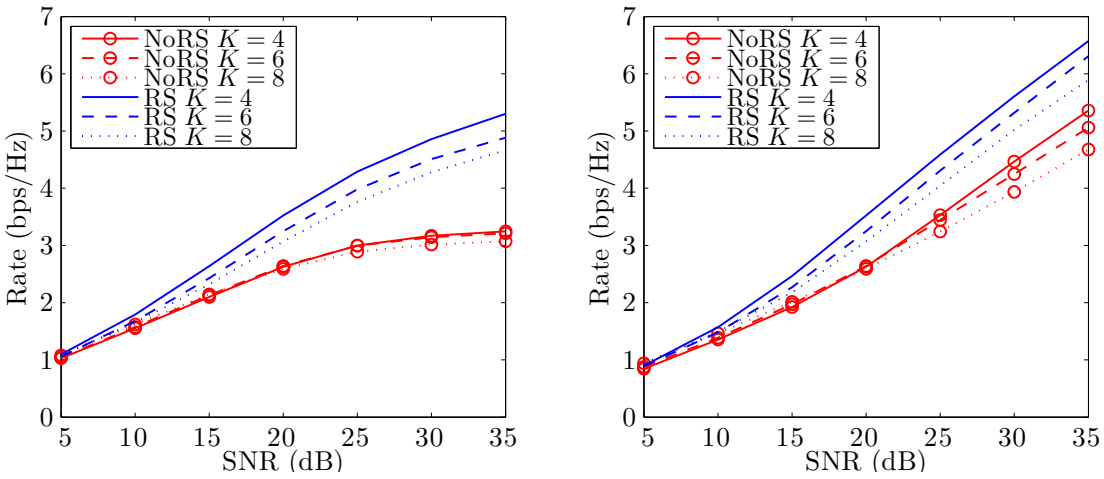
Scaling CSIT

Results for scaling CSIT errors are given in Figure 4.2. The CSIT quality of user-1 remains fixed with $\delta_1 = \delta$, while errors for user-2 and user-3 decay with SNR such that $\alpha_2, \alpha_3 = 0.5$ and $\beta_2, \beta_3 = 10\delta^2$, yielding $\delta_2, \delta_3 = \delta\sqrt{10P_{\mathrm{t}}^{-0.5}}$. Therefore, we have $\delta_2, \delta_3 < \delta_1$ for SNRs less than 20 dB, $\delta_2, \delta_3 = \delta_1$ for 20 dB SNR, and $\delta_2, \delta_3 > \delta_1$ for SNRs greater than 20 dB.

The general observations made for Figure 4.1 still hold, with the cs method providing improved performance over the con method, and RS schemes outperforming NoRS schemes. From a DoF perspective, the cs schemes perform almost as predicted in Theorem 4.1. While the theoretically anticipated MMF-DoFs are given by $\bar{d}^* = 0.25$ and $\bar{d}^*_{\mathrm{RS}} = 0.5$. NoRS-cs achieves DoFs of 0.26 and 0.24, where RS-cs achieves DoFs of 0.53 and 0.47, for $\delta = 0.05$ and 0.15 respectively. On the other hand, NoRS-con and RS-con fail to achieve the corresponding DoFs due to self-interference.

Larger Systems

After demonstrating the superiority of the cutting-set method, we examine the RS gains in larger systems with $K, N_{\mathrm{t}} = 4, 6$ and 8. The performances of NoRS-cs and RS-cs are given



(a) Fixed CSIT ($\alpha = 0$)

(b) Scaling CSIT ($\alpha = 0.5$)

Figure 4.3.: Cutting-set rate performance for $K, N_\mathrm{t} = 4, 6$ and 8, fixed CSIT: $\delta_1, \ldots, \delta_K = 0.15$, and scaling CSIT: $\delta_1, \ldots, \delta_K = 0.15\sqrt{10P_\mathrm{t}^{-0.5}}$.

in Figure 4.3a for non-scaling CSIT with $\delta_1, \ldots, \delta_K = \delta$, and Figure 4.3b for scaling CSIT with $\delta_1, \ldots, \delta_K = \delta\sqrt{10P_\mathrm{t}^{-\alpha}}$, where $\delta = 0.15$ and $\alpha = 0.5$.

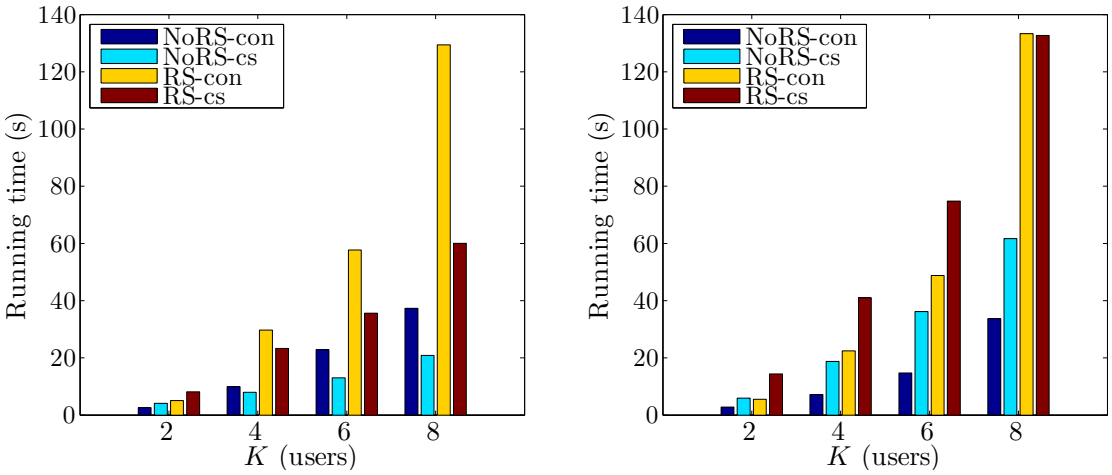
For a given scheme, the performance generally degrades as the number of users increases. This can be regarded to the increased MU interference in NoRS, in addition to the fact that the common message is shared among more users in RS. However, the performance gains associated with the RS scheme are still significant.

4.7.2. Complexity Comparison

Algorithm 4.1 solves a number of SDPs in each iteration, while in each iterations of Algorithm 4.2, a sequence of problems with quadratic constraints is solved in the optimization step, and a sequence of SDPs in the pessimization step. Contrary to the algorithms proposed in the previous chapter, the two algorithms here have different structures and seem to exhibit different running times. This calls for a complexity comparison.

We consider $K = N_\mathrm{t}$ for simplicity, hence reducing the complexity scaling orders to one parameter. In each iteration of Algorithm 4.1, $2K$ equalizers are updated by solving SDPs with a worst-case complexity of $O(K^{3.5})$ each[4], while precoders are updated by solving a SDP with a worst-case complexity of $O(K^8)$. On the other hand, each cutting-set iteration of Algorithm 4.2 consists of an optimization step and a pessimization step, which are iterative in their own rights. Each optimization-iteration involves updating the precoders by solving a convex problem with a number of quadratic constraints that grows with the outer (cutting-set) iteration. For the ith cutting-set iteration, the number of WMSE constraints

[4]Worse-case computational costs of solving standard problems using interior-point methods are given according to [87, Lecture 6]. The term that accounts for the solution's accuracy is omitted, e.g. [71].



(a) $\delta = 0.05$ (b) $\delta = 0.15$

Figure 4.4.: Average run-time of NoRS and RS, conservative and cutting-set methods for $K, N_{\mathrm{t}} = 2, 4, 6$ and 8, SNR = 20 dB, and $\delta_1, \ldots, \delta_K = \delta$.

cannot exceed $2iK$, and updating the precoders in each inner (optimization) iteration can be formulated as a Second Order Cone Program (SOCP) with a worst-case complexity of $O(i^{1.5}K^{7.5})$. On the other hand, each pessimization-iteration involves solving $2K$ SDPs with a cost of $O(K^{6.5})$ each.

Due to the iterative nature of the two algorithms, in addition to the nested structure of Algorithm 4.2, a rigorous analytic complexity comparison is not possible. Alternatively, we evaluate their average running times using MATLAB on a computer equipped with an Intel Core i7-3770 @3.4GHz processor and 8.00 GB of RAM. Figure 4.4 shows the average running times (over 100 realizations) of the different schemes versus the number of users/antennas at 20 dB SNR. For a given method (con or cs), RS has longer running times than NoRS due to the higher number of variables involved. The con method (NoRS and RS) is hardly influenced by the level of CSIT uncertainty, exhibiting a slight increase in running times for larger δ due to the increased involvement of the common message, which influences the convergence of the AO algorithm. On the other hand, the cs method (NoRS and RS) is more influenced by the degree of uncertainty, exhibiting a faster increase in running times with K for higher δ. This appears to be due to the higher number of pessimization steps required to sample larger uncertainty regions, resulting in an increased number of cutting-set iterations and a growing complexity of the optimization step.

It should be highlighted that in feedback systems, channel quantization codebooks are predetermined and known to the BS. Hence, corresponding precoders can be calculated beforehand, and relatively long running times do not prohibit the real-time application of such algorithms under limited BS processing capabilities.

4.8. QoS Problem

In this part, we extend the RS strategy to address the inverse power problem, i.e. minimizing the total transmit power under a minimum rate constraint (the QoS problem). Conventionally, this problem is formulated as

$$\mathcal{P}(\bar{R}): \begin{cases} \min\limits_{\mathbf{P}_\mathrm{p}} & \mathrm{tr}\big(\mathbf{P}_\mathrm{p}\mathbf{P}_\mathrm{p}^H\big) \\ \text{s.t.} & \bar{R}_k \geq \bar{R},\ \forall k \in \mathcal{K} \end{cases} \tag{4.39}$$

where $\bar{R}$ is the worst-case rate that should be guaranteed to all users. While the formulation in (4.39) assumes the same QoS demand by all users, it can be easily extended to the scenario where different QoS levels are required.

The RS version of this problem with a minimum rate target $\bar{R}_\mathrm{t}$ writes as

$$\mathcal{P}_\mathrm{RS}(\bar{R}_\mathrm{t}): \begin{cases} \min\limits_{\bar{\mathbf{c}},\mathbf{P}} & \mathrm{tr}\big(\mathbf{P}\mathbf{P}^H\big) \\ \text{s.t.} & \bar{R}_k + \bar{C}_k \geq \bar{R}_\mathrm{t},\ \forall k \in \mathcal{K} \\ & \bar{R}_{\mathrm{c},k} \geq \sum_{l=1}^{K} \bar{C}_l,\ \forall k \in \mathcal{K} \\ & \bar{C}_k \geq 0,\ \forall k \in \mathcal{K}. \end{cases} \tag{4.40}$$

For the power problem addressed in this section we only consider non-scaling CSIT, i.e. $\delta_1^2, \ldots, \delta_K^2 = O(1)$ and $\alpha_1, \ldots, \alpha_K = 0$. This is particularly relevant in this scenario where we assume no BS power constraint, and the CSIT quality is not expected to scale with the transmit power variation during the optimization procedure as channel estimation and feedback is carried out prior to the precoder design.

4.8.1. The Feasibility Issue

As we saw in the previous analysis, under non-scaling CSIT quality, the NoRS MMF rates saturates at high SNRs, where multiuser interference becomes dominant and cannot be completely dealt with. Conversely, this creates a feasibility issue for the power minimization problem, since rates beyond the saturation level cannot be achieved. This has been noted in the related literature [72, 74].

We observe that the rate and the power problems are monotonically non-decreasing in their arguments, and are related such that

$$\mathcal{R}\big(\mathcal{P}(\bar{R})\big) = \bar{R} \quad \text{and} \quad \mathcal{R}_\mathrm{RS}\big(\mathcal{P}_\mathrm{RS}(\bar{R}_\mathrm{t})\big) = \bar{R}_\mathrm{t} \tag{4.41}$$

This can be shown by contradiction and power scaling [6, 59]. From the monotonicity of $\mathcal{R}(P_\mathrm{t})$ and Theorem 4.1, it follows that under non-scaling CSIT qualities, $\mathcal{R}\big(P_\mathrm{t}\big)$ converges to a finite maximum value as $P_\mathrm{t} \to \infty$. The monotonicity of $\mathcal{P}(\bar{R})$ dictates that this value is the maximum feasible rate. On the other hand, $\mathcal{R}_\mathrm{RS}(P_\mathrm{t})$ does not converge. Therefore, any

finite rate is feasible for $\mathcal{P}_{\mathrm{RS}}(\bar{R}_{\mathrm{t}})$, which is always guaranteed by the cutting-set method. This can also be explained by noting the QoS multicast problem [6], which is always feasible, is in fact a subproblem of (4.40). Hence, a proper RS QoS design is expected to eliminate the feasibility issue arising in NoRS QoS designs.

4.8.2. Numerical Results

The power problems in (4.39) and (4.40) are solved using the conservative and cutting-set methods described in the previous sections, yielding the four designs: NoRS-con, NoRS-cs, RS-con and RS-cs. While modifying Algorithm 4.1 and Algorithm 4.2 to address (4.40) and (4.39) is straightforward, it should be noted that an arbitrary initialization of $\mathbf{P}$ may easily yield an infeasible point which fails to satisfy the rate constraints. In this case, the AO algorithm fails to produce a feasible solution, making the initialization a crucial step.

Before we proceed to discuss the initialization step, we state the simulation parameters. The minimum rate constraint is set to 3.3219 bps/Hz, which corresponds to a worst-case user SINR of 9 dB for the NoRS case [72,74]. The four designs are tested under 100 channel realizations, for $\delta_1, \delta_2, \delta_3 = \delta$, where $\delta \in \{0.01, 0.05, 0.1, 0.15\}$.

Initialization

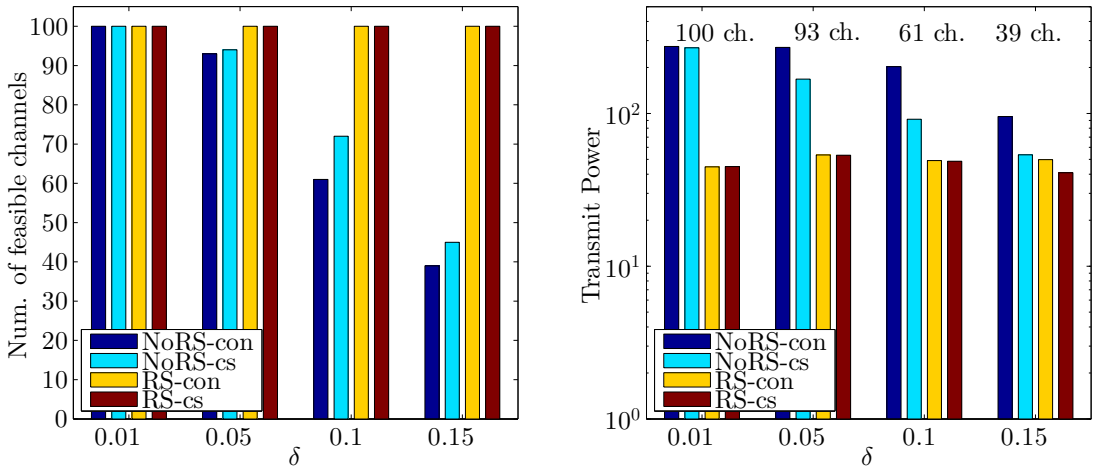
Note that the rate optimization problems are easily initialized by picking any precoder that satisfies $\mathrm{tr}\big(\mathbf{P}\mathbf{P}^H\big) \leq P_{\mathrm{t}}$. This is exploited to obtain a feasible $\mathbf{P}$ for the power problems.

First, let us consider RS-con with a rate constraint $\widehat{R}_{\mathrm{t}}$. P_{t} is initialized and the rate optimization procedure in Algorithm 4.1 is performed until we obtain $\widehat{R}_{\mathrm{t}}^{(n)} \geq \widehat{R}_{\mathrm{t}}$. The corresponding $\mathbf{P}$ is feasible for the power problem since it satisfies the rate constraint. If $\widehat{R}_{\mathrm{t}}^{(n)}$ converges before satisfying the rate constraint, P_{t} is increased until a feasible point is found. A feasible $\mathbf{P}$ can be obtained using very few iterations if P_{t} is adjusted properly.

For RS-cs and NoRS-cs, a similar procedure is followed at the beginning of each optimization step, while noting that some rate constraints may not be feasible for NoRS-cs, and hence P_{t} should not be increased indefinitely. NoRS-con boils down to the SDP solution in [72], which does not require initialization.

Feasibility

Figure 4.5a shows the number of realizations for which the different designs yield a feasible solution. For the NoRS schemes, the number of feasible channels decreases as the CSIT uncertainty increases. NoRS-cs outperforms NoRS-con in this regards due to the latter's employment of conservative approximations. The RS schemes yield feasible solutions for all realizations, with an improvement exceeding 100% compared to NoRS schemes at $\delta = 0.15$.



(a) Num. of feasible realizations

(b) Transmit power

Figure 4.5.: Power minimization under a QoS constraint of 3.3219 bps/Hz in a system with $K, N_{\mathrm{t}} = 3$, $\sigma_{\mathrm{n}}^2 = 1$, and $\delta_1, \delta_2, \delta_3 = \delta$.

Transmit Power

Figure 4.5b shows the total transmit powers averaged over realizations which are feasible for all designs, i.e. the intersection of the three feasible sets for a given δ. It can be seen that RS schemes are more efficient in terms of total transmit power compared to NoRS designs. Intuitively, we expect this contrast to increase with δ (by reversing the observations in Figure 4.1). This holds if infeasible realizations are assigned infinitely large transmit powers. However, since more realizations are omitted for increased δ, the powers obtained in Figure 4.5a for a larger δ are in fact averaged over very well conditioned channels.

4.9. Summary and Conclusion

In this chapter, the classical robust optimization problem of achieving max-min fairness in a MU-MISO system with bounded CSIT errors was addressed through the RS transmission strategy. We proved that the proposed RS strategy outperforms the conventional NoRS strategy in the high SNR regime through DoF analysis. We proposed two robust WMMSE-based algorithms to realize such gains: a conservative algorithm based on the approximations in [45], and a non-conservative algorithm based on the cutting-set method in [10]. Moreover, the convergence of the cutting-set algorithm to the set of KKT solutions of the non-convex semi-infinite rate optimization problem was established.

The performances of the proposed algorithms and the gains of RS were demonstrated through simulations. The approach was also extended to solve the QoS problem. It was shown that RS eliminates the feasibility issue arising in QoS designs with NoRS. Moreover, RS reduces the amount of transmission power required to satisfy the QoS constraints.

5. Multigroup Multicasting

So far, we have considered scenarios where the BS communicates a distinct private message to each receiver. In this chapter, we depart from this setup and consider the case where users form groups, with each group requesting the same *content*, or message. This is known in literature as multigroup multicasting, as each message is transmitted in a multicasting manner to its intended group of users. We consider the system under perfect CSIT, as its performance limits have not been fully identified.

This scenario draws resemblance to the MU-MISO setups with partial CSIT considered in the previous chapters. In particular, each group here can be thought of as a receiver in the MU-MISO setup, while each user in a given group resembles a state within a finite uncertainty region. Such resemblance motivates the application of the RS strategy to multigroup multicast beamforming.

5.1. Introduction

The broadcast nature of the wireless channel allows all receivers in the range of a BS to *listen* to its transmitted signal. This is normally perceived as a *curse* in typical scenarios where users demand distinct messages, due to the resulting multiuser interference as we saw in the previous chapters. However, the misfortunes of some people are indeed advantages to others. In particular, considering a scenario where all users are interested in the same message (or content), the broadcast nature of the wireless channel turns into a *blessing*, as the signal can be simply transmitted in an isotropic manner. This scenario is known in literature as physical-layer multicasting, or simply multicasting.

Although isotropic transmission is simple and sufficient in such scenarios, it was shown in [6, 88] that the availability of multiple transmit antennas and perfect CSIT can be employed to achieve nontrivial gains through multicast beamforming. Since all users are interested in decoding the same message, the power-constrained multicast beamformer is designed such that the minimum rate (or SINR) amongst all users is maximized. The problem of finding the optimum beamforming direction is non-convex, and was shown to be NP-hard in general [6]. However, an upper-bound on the performance can be efficiently obtained using SDR, i.e. by relaxing a rank-one constraint on the covariance matrix of the transmitted signal and solving the resulting SDP. This upper-bound coincided with the information theoretic capacity of the multicast channel [89]. Moreover, beamforming vectors can be extracted from the (possibly non-rank-one) capacity achieving covariance matrix using Gaussian randomization, which gives a good approximation [6].

The work in [6] was later generalized to the case where the BS serves multiple cochannel multicast groups, also known as multigroup multicasting [90]. In particular, users served by the BS form multicast groups depending on the content they are interested in, and the BS transmits a distinct message to each group. It can be seen that single-group multicasting is a special case of this scenario, i.e. when there is only one group. On the other hand, multigroup multicasting reduces to standard multiuser transmission with private messages when each group consists of only one user. Since multigroup multicasting generally involves multiple distinct messages, the problem of interference returns. The interference in such setups is discussed and analysed in detail after introducing the system model.

5.2. Multigroup Multicast Beamforming

Here we describe the multigroup multicasting system model. The general MISO-BC signal model described in Section 2.1.1 applies, where a BS with N_{t} transmitting antennas communicates with K single antenna receivers. However, the K users here are grouped into the M multicasting groups $\mathcal{G}_1, \ldots, \mathcal{G}_M$, where all users belonging to the same group are interested in the same message. Let $\mathcal{M} \triangleq \{1, \ldots, M\}$ be the group-index set. We assume that $\bigcup_{m \in \mathcal{M}} \mathcal{G}_m = \mathcal{K}$, and $\mathcal{G}_m \cap \mathcal{G}_j = \emptyset$, for all $m, j \in \mathcal{M}$ and $m \neq j$. In words, each user belongs to one and only one multicasting group. Denoting the size of the mth group by $G_m = |\mathcal{G}_m|$, we assume without loss of generality that group sizes are in an ascending order:

$$G_1 \leq G_2 \leq \ldots \leq G_M. \tag{5.1}$$

Let $W_1, \ldots, W_M$ be the messages intended to the groups $\mathcal{G}_1, \ldots, \mathcal{G}_M$ respectively. We consider conventional beamforming transmission. Messages are first encoded into independent data streams, where the vector of coded data symbols in a given channel use is given by $\mathbf{s}_{\mathrm{p}} \triangleq [s_1, \ldots, s_M]^T$. Data streams are then mapped to the transmit antennas through a linear precoding matrix $\mathbf{P}_{\mathrm{p}} \triangleq [\mathbf{p}_1, \ldots, \mathbf{p}_M]$, where the transmit signal is given by

$$\mathbf{x} = \mathbf{P}_{\mathrm{p}} \mathbf{s}_{\mathrm{p}} = \sum_{m=1}^{M} \mathbf{p}_m s_m. \tag{5.2}$$

Note that while the model in Section 2.4.1 had one precoding vector (and one symbol stream) for each user, here we have one precoding vector for each group. We recall that under linear precoding, the power constraint writes as $\sum_{m=1}^{M} \|\mathbf{p}_m\|^2 \leq P_{\mathrm{t}}$.

The signal received by the kth user is expressed by

$$y_k = \overbrace{\mathbf{h}_k^H \mathbf{p}_{\mu(k)} s_{\mu(k)}}^{\text{desired signal}} + \overbrace{\mathbf{h}_k^H \sum_{m \neq \mu(k)} \mathbf{p}_m s_m}^{\text{interference}} + n_k \tag{5.3}$$

where $\mu : \mathcal{K} \mapsto \mathcal{M}$ maps a user-index to the corresponding group-index, i.e.

$$\mu(k) = m \;\; \text{such that} \;\; k \in \mathcal{G}_m. \tag{5.4}$$

In the following, $\mu(k)$ is referred to as μ for brevity where the argument of the function is clear from the context. Note that while the interference in (2.18) corresponds to inter-user interference, the one in (5.3) is inter-group interference.

The kth user's average receive power writes as

$$T_k = \overbrace{|\mathbf{h}_k^H \mathbf{p}_\mu|^2}^{S_k} + \overbrace{\sum_{m \neq \mu} |\mathbf{h}_k^H \mathbf{p}_m|^2 + \sigma_{\mathrm{n}}^2}^{I_k} \tag{5.5}$$

from which the SINR experienced by the kth user is defined as $\gamma_k \triangleq S_k I_k^{-1}$. Under Gaussian signalling, the kth achievable user-rate is given by $R_k = \log_2(1 + \gamma_k)$. In multigroup multicasting, users belonging to the same group decode the same data stream. Hence, to guarantee that all users in the mth group are able to recover W_m successfully, it should be transmitted at the group-rate defined as

$$r_m \triangleq \min_{i \in \mathcal{G}_m} R_i. \tag{5.6}$$

The main problems considered in the multicasting literature are those of classical multiuser beamforming, namely the QoS constrained power minimization problem and the power constrained MMF problem [90–93]. Such multicasting problems were also considered under various constraints and setups, including: per-antenna power constraints [94], large-scale arrays [95] and coordinated multicell transmission [96]. The sum-rate maximization problem was also considered in [97]. In this chapter, the focus on the MMF problem. However, the developed schemes can be easily extended to the QoS and sum-rate problems.

Assumption 5.1. *In this chapter, we assume perfect CSIT. As a result, non-ergodic transmission is considered as instantaneous rates are well defined and known to the BS.*

5.2.1. Achieving Max-Min Fairness

In the light of the conventional multi-stream beamforming model for multigroup multicasting, the MMF problem is formulated as

$$\mathcal{R}(P_{\mathrm{t}}) : \begin{cases} \max\limits_{\mathbf{P}_{\mathrm{p}}} & \min\limits_{m \in \mathcal{M}} \min\limits_{i \in \mathcal{G}_m} R_i \\ \text{s.t.} & \sum\limits_{m=1}^{M} \|\mathbf{p}_m\|^2 \leq P_{\mathrm{t}}. \end{cases} \tag{5.7}$$

The inner minimization in (5.7)'s objective function accounts for the multicast nature within each group, i.e. to guarantee that each message is successfully decoded by all users

in the corresponding group. The outer minimization on the other hand accounts for the fairness across groups.

The above problem was shown to be NP-hard in general [90]. It is common practice to formulate an equivalent problem in terms of the SINRs[1] and solve it approximately. The method based on SDR and randomization proposed for the single-group multicasting in [6] was extended to the multigroup scenario in [90]. An alternative solution based on convex approximation was later proposed in [93], exhibiting marginally improved performances under certain conditions, and more importantly, lower complexity.

5.3. Performance Limits

While considerable attention has been given to solving the complicated multigroup multicasting optimization problems, little has been done regarding the analytic derivation and characterization of the performance. This can be regarded to the difficulty (or rather impossibility) of obtaining closed-form expressions for the optimum performance metrics, as the original problems can only be solved approximately through numerical methods. Hence, performance analysis has been carried out through extensive simulations.

In this section, we characterize the performance of (5.7) through DoF analysis. The relevance of DoF analysis here is explained as follows. We recall that achieving max-min fairness requires a simultaneous increase in powers allocated to all streams as P_{t} increases. In scenarios where the number of transmit antennas N_{t} is sufficient to place each beam in the null space of all its unintended groups, we could safely assume that each multicasting group receives an interference free stream. However, if such condition is violated, it is not clear what happens. As we see in the following, this can be unveiled through DoF analysis[2].

5.3.1. DoF Analysis

Let us start by defining the DoF achieved by a precoding scheme denoted by $\left\{\mathbf{P}_{\mathrm{p}}(P_{\mathrm{t}})\right\}_{P_{\mathrm{t}}}$. The kth user's DoF is defined as $D_k \triangleq \lim_{P_{\mathrm{t}}\to\infty} \frac{R_k(P_{\mathrm{t}})}{\log_2(P_{\mathrm{t}})}$. It follows that the mth group-DoF is given by

$$d_m \triangleq \lim_{P_{\mathrm{t}}\to\infty} \frac{r_m(P_{\mathrm{t}})}{\log_2(P_{\mathrm{t}})} = \min_{i\in\mathcal{G}_m} D_i \tag{5.8}$$

which is achieved by all users in the group. Hence, the performance of the precoding scheme is characterized by the tuple $[d_1, \ldots, d_M]$. The MMF-DoF achieved by the precoding scheme is defined as $d \triangleq \min_{m\in\mathcal{M}} d_m$, corresponding to the DoF that can be simultaneously achieved by all groups. Since each user is equipped with a single antenna, we observe that

$$d \leq d_m \leq D_i \leq 1, \ \forall i \in \mathcal{G}_m, m \in \mathcal{M}. \tag{5.9}$$

[1] We recall that there exists a one-to-one monotonic relationship between R_i and γ_i.

[2] We recall that the DoF is roughly interpreted as the number of interference-free streams that can be simultaneously communicated in a single channel use. Hence, such analysis is highly relevant here.

Hence, if $d = 1$ is achievable, then it is also optimum. Next, we make an assumption regarding the channel vectors.

Assumption 5.2. *The channel vectors* $\mathbf{h}_1, \ldots, \mathbf{h}_K$ *are independently drawn from their corresponding distributions, assumed to be continuous. Hence, for any* $N_\mathrm{t} \times K_\mathrm{sub}$ *matrix in which the* K_sub *column vectors constitute any subset of the* K *channel vectors, it holds with probability one that the rank is* $\min\{N_\mathrm{t}, K_\mathrm{sub}\}$.

Let us define $\mathbf{H}_m$ as the matrix with columns constituting channel vectors of all users in $\mathcal{G}_m$, and $\bar{\mathbf{H}}_m \triangleq [\mathbf{H}_1, \ldots, \mathbf{H}_{m-1}, \mathbf{H}_{m+1}, \ldots, \mathbf{H}_M]$ as the complementary set of channel vectors. Moreover, $\bar{\mathbf{H}}_m \setminus \mathbf{H}_j$ denotes the matrix formed by excluding $\mathbf{H}_j$ from $\bar{\mathbf{H}}_m$, where $j \in \mathcal{M} \backslash m$. By Assumption 5.2, the null space dimensions of such matrices can be identified, which plays a significant role in deriving the DoF results in this chapter.

Now, we identify the condition under which the full DoF of $[d_1, \ldots, d_M] = [1, \ldots, 1]$, and hence $d = 1$, is achieved. By Assumption 5.2, it can be seen that

$$\dim\Big\{\mathrm{null}\big(\bar{\mathbf{H}}_m^H\big)\Big\} = \max\big\{N_\mathrm{t} + G_m - K, 0\big\} \tag{5.10}$$

where $\dim\{\cdot\}$ and $\mathrm{null}\{\cdot\}$ correspond to the dimension and the null space respectively. It follows that that having $\dim\Big\{\mathrm{null}\big(\bar{\mathbf{H}}_m^H\big)\Big\} \geq 1$ requires $N_\mathrm{t} \geq N_m$ where

$$N_m \triangleq 1 + \sum_{i \neq m} G_i = 1 + K - G_m. \tag{5.11}$$

When this condition is satisfied, we can design the mth beamforming vector such that it does not cause interference to any other group, i.e. $\mathbf{p}_m \in \mathrm{null}\big(\bar{\mathbf{H}}_m^H\big)$. From (5.1), it follows that $N_1 \geq N_2 \geq \ldots \geq N_M$. This reflects the fact that larger groups are more *demanding* in terms of spatial dimensions required to null interference caused to them. Hence, $\mathbf{p}_1$ requires the largest number of transmitting antennas to meet this demand. Therefore, it follows that having $N_\mathrm{t} \geq N_m$ for all $m \in \mathcal{M}$ is equivalent to

$$N_\mathrm{t} \geq N_1 \tag{5.12}$$

which is a sufficient condition to fully eliminate inter-group interference.

Lemma 5.1. *The condition in* (5.12) *is necessary and sufficient to achieve interference-free beamforming to all groups, and hence a full MMF-DoF of* $d = 1$.

The sufficiency follows from the previous discussion. The necessity come from the fact that if the condition is violated, at least one group will experience inter-group interference, which in turn limits the achievable DoF. This is shown in more detail in the proof of the following theorem in Appendix C.1. The question that comes to mind at this point is: what happens to the MMF-DoF when condition (5.12) is violated? This is answered next.

Theorem 5.1. *The optimum MMF-DoF achieved by solving* (5.7) *is given by*

$$d^* \triangleq \lim_{P_\mathrm{t}\to\infty} \frac{\mathcal{R}(P_\mathrm{t})}{\log_2(P_\mathrm{t})} = \begin{cases} 1, & N_\mathrm{t} \geq N_1 \\ 0.5, & N_M \leq N_\mathrm{t} < N_1 \\ 0, & N_\mathrm{t} < N_M \end{cases} \tag{5.13}$$

Similar to Theorem 3.1 and Theorem 4.1, the above result is shown through two steps: achievability and converse. The former is discussed in the following, where we also draw some insights into the result. The latter is relegated to Appendix C.1.

5.3.2. Achievability and Insight

We start by presenting an instrumental observation which is based on Lemma 5.1.

Remark 5.1. *Assume that beamforming is carried out over a subset of groups $\mathcal{M}_\mathrm{ms} \subseteq \mathcal{M}$ where all remaining groups are switched off, the condition*

$$N_\mathrm{t} \geq 1 + K - \sum_{m\in\mathcal{M}\setminus\mathcal{M}_\mathrm{ms}} G_m - \min_{j\in\mathcal{M}_\mathrm{ms}} G_j \tag{5.14}$$

is necessary and sufficient to achieve interference free beamforming (full DoF).

This follows directly from Lemma 5.1 after excluding all groups in $\mathcal{M} \setminus \mathcal{M}_\mathrm{ms}$.

Returning to the achievability, $d = 1$ under $N_\mathrm{t} \geq N_1$ follows from the discussion that precedes Theorem 5.1. We focus on achieving $d = 0.5$. It is sufficient to show this under

$$N_\mathrm{t} = N_M = 1 + G_1 + G_2 + \ldots + G_{M-1} \tag{5.15}$$

as further increasing the number of transmitting antennas cannot decrease the DoF. From (5.15), it is easy to conclude the following:

1. Interference from the Mth stream to all other groups can be nulled.
2. By excluding the largest group (i.e. $\mathcal{G}_M$), interference free transmission amongst the remaining $M-1$ groups is possible.

The first point follows from (5.11). The second point follows by excluding the Mth group from the system and Remark 5.1. Hence, the precoding vectors can be designed such that

$$\mathbf{p}_m \in \begin{cases} \mathrm{null}\big(\bar{\mathbf{H}}_m^H \setminus \mathbf{H}_M^H\big), & \forall m \in \mathcal{M} \setminus M \\ \mathrm{null}\big(\bar{\mathbf{H}}_M^H\big), & m = M \end{cases} \tag{5.16}$$

from which the kth user's SINR writes as

$$\gamma_k = \begin{cases} \frac{|\mathbf{h}_k^H \mathbf{p}_\mu|^2}{\sigma_{\mathrm{n}}^2}, & \forall k \in \mathcal{K} \setminus \mathcal{G}_M \\ \frac{|\mathbf{h}_k^H \mathbf{p}_M|^2}{\sum_{j \neq M} |\mathbf{h}_k^H \mathbf{p}_j|^2 + \sigma_{\mathrm{n}}^2}, & k \in \mathcal{G}_M. \end{cases} \tag{5.17}$$

It can be seen that users in groups $1, \ldots, M-1$ see no interference, while users in group M see interference from all other groups. By setting the power scaling such that

$$q_m = \|\mathbf{p}_m\|^2 = \begin{cases} O\big(P_{\mathrm{t}}^{0.5}\big), & \forall m \in \mathcal{M} \setminus M \\ O\big(P_{\mathrm{t}}\big), & m = M \end{cases} \tag{5.18}$$

we have $\gamma_k = O\big(P_{\mathrm{t}}^{0.5}\big)$ for all $k \in \mathcal{K}$. Hence, $D_1, \ldots, D_K = 0.5$ from which the tuple $[d_1, \ldots, d_M] = [0.5, \ldots, 0.5]$ is achieved, and hence $d = 0.5$.

The idea of the scheme is to align all interference through beamforming such that it lies in the signal subspace of the largest group. The power scaling in (5.18) is then used to partition this signal subspace into two halves, one for interference and the other for the desired signal. Note that after applying the precoders in (5.16), taking the scalar precoded channels and only one user in each group, the resulting system can be interpreted as a partially connected interference channel with no internal conflicts [80, 98]. This can be used to show that the MMF-DoF of 0.5 cannot be exceeded. An alternative way of showing this upper-bound is given in Appendix C.1.

When $N_{\mathrm{t}} < N_M$, the interference from group M cannot be eliminated anymore. This creates mutual interference between groups, and the MMF-DoF collapses to zero as we show in Appendix C.1. We refer to this scenario as an *overloaded* system. We conclude this section by reemphasising the impact a collapsing DoF has on the rate performance. As we saw in the previous chapter, a MMF-DoF of zero implies that the MMF rate stops growing as SNR grows large, reaching a saturated performance. Although the DoF analysis is carried out as SNR goes to infinity, its results are highly visible in finite SNR regimes as we saw previously, and as we show later when we present results from simulations.

5.4. Single-Stream/Multi-Stream Transmission

The saturating performance in overloaded systems can be avoided through single-stream transmission. In particular, the M messages are packed into one super common message given by $W_{\mathrm{c}} = \{W_1, \ldots, W_M\}$. This is encoded into a single data stream, then broadcasted in a single-group multicasting manner such that it is decoded by all users in the system. Since this interference-free transmission achieves a total DoF of 1, each group is guaranteed a non-saturating rate with a group-DoF of $1/M$.

Relying solely on the single-stream strategy may jeopardize (partial) gains achieved through multi-stream beamforming. In particular, while it may not be possible to guarantee interference-free beamforming to all groups, the number of transmitting antennas may

still be able to support beamforming to a subset of groups. Ultimately, the objective is to realize the potential benefits of both single-stream and multi-stream strategies through one general transmission strategy. In this section, we make progress towards this through proposing a mixed scheme that combines the two strategies. This scheme paves the way to, and motivates, the unifying RS scheme proposed in the following section.

5.4.1. The Mixed Scheme

Here, we propose a mixed scheme that applies the multi-stream beamforming of Section 5.2 to a subset of the groups given by $\mathcal{M}_{\text{ms}} \subseteq \mathcal{M}$, and the single-stream transmission describe above to the remaining groups, i.e. $\mathcal{M}_{\text{ss}} = \mathcal{M} \setminus \mathcal{M}_{\text{ms}}$.

Messages intended to the groups in $\mathcal{M}_{\text{ms}}$ are encoded into independent data streams. On the other hand, messages intended to the groups in $\mathcal{M}_{\text{ss}}$ are first packed into a common message, which is then encoded into a common data stream[3]. All data streams are linearly precoded from which the transmitted signal is given by

$$\mathbf{x} = \mathbf{p}_{\text{c}} s_{\text{c}} + \sum_{m \in \mathcal{M}_{\text{ms}}} \mathbf{p}_m s_m. \tag{5.19}$$

At each receiver, the common message is decoded first by treating all interference from multi-stream transmission as noise. Then, if the receiver belongs to one of the groups in $\mathcal{M}_{\text{ms}}$, the common stream is removed using SIC and the *designated*[4] stream is decoded.

The SINR of the common stream at the kth receiver is given by

$$\gamma_{\text{c},k} = \frac{|\mathbf{h}_k^H \mathbf{p}_{\text{c}}|^2}{\sum_{j \in \mathcal{M}_{\text{ms}}} |\mathbf{h}_k^H \mathbf{p}_j|^2 + \sigma_{\text{n}}^2} \tag{5.20}$$

The corresponding achievable rate is given by $R_{\text{c},k} = \log_2(1 + \gamma_{\text{c},k})$, from which the rate of the common stream is given by $R_{\text{c}} = \min_{k \in \mathcal{K}} R_{\text{c},k}$. We recall that the common stream is decoded by all users in the system, and hence the minimization is taken over $\mathcal{K}$. After SIC, the SINR of the designated stream at the kth receiver is given by

$$\gamma_k = \frac{|\mathbf{h}_k^H \mathbf{p}_\mu|^2}{\sum_{j \in \mathcal{M}_{\text{ms}} \setminus \mu} |\mathbf{h}_k^H \mathbf{p}_j|^2 + \sigma_{\text{n}}^2}, \ \forall k \in \mathcal{K} \setminus \bigcup_{m \in \mathcal{M}_{\text{ss}}} \mathcal{G}_m \tag{5.21}$$

which only applies to users belonging to groups in $\mathcal{M}_{\text{ms}}$, or equivalently $k \in \mathcal{K} \setminus \bigcup_{m \in \mathcal{M}_{\text{ss}}} \mathcal{G}_m$. The corresponding rate achieved by the user is given by $R_k = \log_2(1 + \gamma_k)$.

Although all users decode the common stream, its rate is only shared by the groups in $\mathcal{M}_{\text{ss}}$. Since the objective is to achieve fairness, the common rate is divided equally amongst these groups. On the other hand, groups in $\mathcal{M}_{\text{ms}}$ receive their messages through

[3] Common in the sense that is is decoded by all groups as we see next, but clearly carries information intended to a subset of them

[4] This replaces the term *private* used in the previous chapters.

the designated streams. It follows that the group rate is given by

$$r_m = \begin{cases} \min_{i \in \mathcal{G}_m} R_i, & \forall m \in \mathcal{M}_{\text{ms}} \\ R_{\text{c}}/|\mathcal{M}_{\text{ss}}|, & \forall m \in \mathcal{M}_{\text{ss}}. \end{cases} \tag{5.22}$$

Hence, the corresponding optimization problem is formulated as

$$\mathcal{R}_{\text{mix}}(P_{\text{t}}) : \begin{cases} \max\limits_{\mathbf{P}, \mathcal{M}_{\text{ms}}, \mathcal{M}_{\text{ss}}} & \min\left\{R_{\text{c}}/|\mathcal{M}_{\text{ss}}|, \min\limits_{m \in \mathcal{M}_{\text{ms}}} \min\limits_{i \in \mathcal{G}_m} R_i\right\} \\ \text{s.t.} & \mathcal{M}_{\text{ms}} \subseteq \mathcal{M},\ \mathcal{M}_{\text{ss}} = \mathcal{M} \setminus \mathcal{M}_{\text{ms}} \\ & R_{\text{c},k} \geq R_{\text{c}}, \forall k \in \mathcal{K} \\ & \|\mathbf{p}_{\text{c}}\|^2 + \sum_{m \in \mathcal{M}_{\text{ms}}} \|\mathbf{p}_m\|^2 \leq P_{\text{t}} \end{cases} \tag{5.23}$$

where the assignment of groups to $\mathcal{M}_{\text{ms}}$ and $\mathcal{M}_{\text{ss}}$ is part of the optimization problem.

5.4.2. DoF Analysis

Analysing the performance obtained by solving (5.23) requires knowledge of the optimum group assignment, i.e. $\mathcal{M}_{\text{ms}}$ and $\mathcal{M}_{\text{ss}}$. To facilitate such analysis, we derive a simple assignment strategy and analysis the corresponding MMF-DoF performance. As we see in the next section, this assignment turns out to be optimum in a DoF sense. Before we proceed, we recall that designated streams see no interference from the common stream. Hence their DoF is given by (5.8), while the DoF of the common stream is given by

$$d_{\text{c}} \triangleq \lim_{P_{\text{t}} \to \infty} \frac{R_{\text{c}}(P_{\text{t}})}{\log_2(P_{\text{t}})} \tag{5.24}$$

For a given precoding scheme, the DoF achieved by each group in $\mathcal{M}_{\text{ss}}$ is given by $d_{\text{c}}/|\mathcal{M}_{\text{ss}}|$.

Note that groups in $\mathcal{M}_{\text{ms}}$ exploit the multiplexing gain of the channel, while groups in $\mathcal{M}_{\text{ss}}$ share a single stream with a bounded DoF, i.e. $d_{\text{c}} \leq 1$ due to single-antenna receivers. Therefore, we seek to maximize the number of groups in $\mathcal{M}_{\text{ms}}$. Let $M_{\text{ms}}(N_{\text{t}})$ denote the maximum number of groups for which interference-free multi-stream beamforming is guaranteed for a given N_{t}. This is obtained as

$$M_{\text{ms}}(N_{\text{t}}) : \begin{cases} \max & |\mathcal{M}_{\text{ms}}| \\ \text{s.t.} & \mathcal{M}_{\text{ms}} \subseteq \mathcal{M} \text{ and (5.14)}. \end{cases} \tag{5.25}$$

In the following, (N_{t}) is dropped for notational brevity. It turns out that M_{ms} can be

expressed in terms of N_{t}, M and $G_1, \ldots, G_M$ as

$$M_{\mathrm{ms}} = \begin{cases} 1, & 1 \leq N_{\mathrm{t}} < 1 + G_2 \\ 2, & 1 + G_2 \leq N_{\mathrm{t}} < 1 + G_2 + G_3 \\ \vdots & \vdots \\ M-2, & 1 + \sum_{j=2}^{M-2} G_j \leq N_{\mathrm{t}} < 1 + \sum_{j=2}^{M-1} G_j \\ M-1, & 1 + \sum_{j=2}^{M-1} G_j \leq N_{\mathrm{t}} < 1 + \sum_{j=2}^{M} G_j \\ M, & 1 + \sum_{j=2}^{M} G_j \leq N_{\mathrm{t}} \end{cases} \tag{5.26}$$

This is shown as follows: when $N_{\mathrm{t}} < 1 + G_2$, it is not possible to perform interference-free transmission between any pair of groups, i.e. there is no $m_1, m_2 \in \mathcal{M}$ such that $\mathbf{p}_{m_1} \in \mathrm{null}\big(\mathbf{H}_{m_2}^H\big)$ and $\mathbf{p}_{m_2} \in \mathrm{null}\big(\mathbf{H}_{m_1}^H\big)$. Hence, $M_{\mathrm{ms}} = 1$, implying that interference-free transmission is only achieved by excluding $M-1$ groups. $N_{\mathrm{t}} = 1 + G_2$ however is sufficient for interference-free transmission among groups 1 and 2, which follows from Remark 5.1. For larger N_{t}, we add more groups starting by the least demanding, from which (5.26) is obtained. It can shown that for any given N_{t}, M_{ms} in (5.26) cannot be exceeded by packing larger groups as they are more demanding in terms of spatial dimensions.

Now that M_{ms} is characterized, the simple group assignment strategy is given by

$$\mathcal{M}_{\mathrm{ms}} = \{1, \ldots, M_{\mathrm{ms}}\} \quad \text{and} \quad \mathcal{M}_{\mathrm{ss}} = \{M_{\mathrm{ms}} + 1, \ldots, M\}. \tag{5.27}$$

The MMF-DoF achieved by fixing this assignment strategy is given in the next result.

Proposition 5.1. *The mixed scheme achieves the MMF-DoF given by*

$$d_{\mathrm{mix}}^* \triangleq \lim_{P_{\mathrm{t}} \to \infty} \frac{\mathcal{R}_{\mathrm{mix}}(P_{\mathrm{t}})}{\log_2(P_{\mathrm{t}})} \geq \frac{1}{1 + M - M_{\mathrm{ms}}} \tag{5.28}$$

Proof. Under the assignment in (5.27), the designated precoders are designed as

$$\mathbf{p}_m \in \mathrm{null}\big(\bar{\mathbf{H}}_m^H \setminus [\mathbf{H}_{1+M_{\mathrm{ms}}}, \ldots, \mathbf{H}_M]^H\big), \ \forall m \in \mathcal{M}_{\mathrm{ms}}. \tag{5.29}$$

That is, interference-free multi-stream transmission is carried out within the groups in $\mathcal{M}_{\mathrm{ms}}$. On the other hand, $\mathbf{p}_{\mathrm{c}}$ is randomly selected from the space spanned by $[\mathbf{H}_1, \ldots, \mathbf{H}_M]^H$. Now, let us assign powers that scale as $q_{\mathrm{c}} = O\big(P_{\mathrm{t}}\big)$ for the common stream, and $q_m = O\big(P_{\mathrm{t}}^{\alpha}\big)$ for designated streams, where $\alpha \in [0, 1]$. It follows from (5.20) that $\gamma_{\mathrm{c},k} = O\big(P_{\mathrm{t}}^{1-\alpha}\big)$ for all $k \in \mathcal{K}$, and from (5.21) that $\gamma_k = O\big(P_{\mathrm{t}}^{\alpha}\big)$ for all $k \in \mathcal{K} \setminus \bigcup_{m \in \mathcal{M}_{\mathrm{ss}}} \mathcal{G}_m$.

From the SINRs, we observe that $d_{\mathrm{c}} = 1 - \alpha$ and $d_m = \alpha$ for all $m \in \mathcal{M}_{\mathrm{ms}}$. Now, by setting $\alpha = \frac{1}{1+M-M_{\mathrm{ms}}}$, the DoF in (5.28) is achieved for the designated streams, and hence all groups in $\mathcal{M}_{\mathrm{ms}}$. Moreover, we have $d_{\mathrm{c}} = 1 - \frac{1}{1+M-M_{\mathrm{ms}}} = \frac{M-M_{\mathrm{ms}}}{1+M-M_{\mathrm{ms}}}$. Dividing this equally amongst the groups served by the common stream, each gets the DoF in (5.28) as $|\mathcal{M}_{\mathrm{ss}}| = M - M_{\mathrm{ms}}$, which completes the proof. □

Note that the inequality in (5.28) indicates that the MMF-DoF given by $\frac{1}{1+M-M_{\text{ms}}}$ is achievable as shown in the proof, yet not necessarily optimal. Showing that this MMF-DoF cannot be exceeded by any feasible precoding scheme turns the inequality into an equality (converse). This is left for the following section when we discuss the RS scheme.

Now, the next question that comes to mind is: how does the MMF-DoF achieved through the mixed scheme in Proposition 5.1 compare to the one achieved through the conventional scheme in Theorem 5.1? To gain some insight into the answer, let us consider the case where $N_{\text{t}} = 1 + K - G_1 - G_M$. It follows from (5.13) that $d^* = 0$, as $N_{\text{t}} < N_M$. Note that the same number of antennas can be written as $N_{\text{t}} = 1 + \sum_{j=2}^{M-1} G_m$, from which we get $M_{\text{RS}} = M - 1$. Hence, it follows from (5.28) that $d^*_{\text{mix}} \geq 0.5$. It is clear from this example that the mixed scheme maintains 0.5 under a smaller number of transmitting antennas. Moreover, since single-stream transmission is a special case of the mixed scheme, then $d^*_{\text{mix}} \geq \frac{1}{M} > 0$ is guaranteed. More on the DoF gains is given in Section 5.5.1.

5.4.3. Limitations

In the next section, we show that the MMF-DoF in (5.28) is optimum for the mixed scheme in problem (5.23), implying the optimality of the proposed group assignment strategy in a DoF sense. However, it is not yet clear if such fixed assignment strategy is preferred over the entire range of SNRs. One way to obtain the optimum solution of (5.23), and hence the optimum assignment strategy, over the entire range of SNRs is to solve it for all possible $\mathcal{M}_{\text{ms}} \subseteq \mathcal{M}$. In the following section, we show that this is not necessary as the mixed scheme is nothing more than a special case of the RS scheme.

5.5. Rate-Splitting for Multi-Group Multicasting

The generalized RS strategy proposed for MU-MISO in Section 4.3.1 is extended to the multigroup multicasting scenario in this section. Users in the MU-MISO case are replaced by groups here, where each group message is split into a common part and a designated part. For example, $W_m = \{W_{m0}, W_{m1}\}$, with W_{m0} and W_{m1} being the common and designated parts respectively. Common parts are packed into a super common message $W_{\text{c}} \triangleq \{W_{10}, \ldots, W_{M0}\}$, and the resulting common and designated messages are encoded into independent data stream, in a manner similar to (4.7). All streams are then linearly precoded and transmitted through the antenna array.

Similar to the RS signals seen throughout this thesis, the transmitted signal writes as a superposition of the precoded common stream and designated streams such that

$$\mathbf{x} = \mathbf{p}_{\text{c}} s_{\text{c}} + \sum_{m=1}^{M} \mathbf{p}_m s_m. \tag{5.30}$$

The signal received by the kth user is given by

$$y_k = \overbrace{\mathbf{h}_k^H \mathbf{p}_{\mathrm{c}} s_{\mathrm{c}}}^{\text{common signal}} + \overbrace{\mathbf{h}_k^H \mathbf{p}_\mu s_\mu}^{\text{desired designated signal}} + \overbrace{\mathbf{h}_k^H \sum_{m \neq \mu} \mathbf{p}_m s_m}^{\text{interference}} + n_k \tag{5.31}$$

from which the corresponding average received power writes as

$$T_{\mathrm{c},k} = \overbrace{|\mathbf{h}_k^H \mathbf{p}_{\mathrm{c}}|^2}^{S_{\mathrm{c},k}} + \underbrace{\overbrace{|\mathbf{h}_k^H \mathbf{p}_\mu|^2}^{S_k} + \overbrace{\sum_{m \neq \mu} |\mathbf{h}_k^H \mathbf{p}_m|^2 + \sigma_{\mathrm{n}}^2}^{I_k}}_{I_{\mathrm{c},k} = T_k}. \tag{5.32}$$

By treating all designated signals as noise, the SINR of the common stream at the kth user is given by $\gamma_{\mathrm{c},k} \triangleq S_{\mathrm{c},k} I_{\mathrm{c},k}^{-1}$, from which the corresponding rate is given by $R_{\mathrm{c},k} = \log_2(1+\gamma_{\mathrm{c},k})$. The rate of the common message is hence given by $R_{\mathrm{c}} = \min_{\mathcal{K}} R_{\mathrm{c},k}$, ensuring its decodability by all users in the system. After decoding the common message, it is removed from the received signal through SIC. Hence, the SINR and rate of the designated stream are similar to the ones given in Section 5.2.

As in Section 4.3.1, the common rate writes as a sum of M portions: $R_{\mathrm{c}} = \sum_{m=1}^{M} C_m$, where C_m is the portion allocated to the mth group, corresponding to the rate at which W_{m0} is transmitted. It follows that the mth group-rate is given by

$$C_m + \min_{i \in \mathcal{G}_m} R_i \tag{5.33}$$

consisting of a common rate portion plus a designated rate. The MMF problem for multi-group multicasting is formulated in terms of RS as follows

$$\mathcal{R}_{\mathrm{RS}}(P_{\mathrm{t}}) : \begin{cases} \max\limits_{\mathbf{c},\mathbf{P}} & \min\limits_{m \in \mathcal{M}} \left(C_m + \min\limits_{i \in \mathcal{G}_m} R_i \right) \\ \text{s.t.} & R_{\mathrm{c},k} \geq \sum_{m=1}^{M} C_m, \forall k \in \mathcal{K} \\ & C_m \geq 0, \forall m \in \mathcal{M} \\ & \|\mathbf{p}_{\mathrm{c}}\|^2 + \sum\limits_{m=1}^{M} \|\mathbf{p}_m\|^2 \leq P_{\mathrm{t}} \end{cases} \tag{5.34}$$

where $\mathbf{c} \triangleq [C_1, \ldots, C_M]^T$. The constraints in (5.34) can be explained as in Section 4.3.1.

The performances of the conventional multicast beamforming problem in (5.7), the mixed problem in (5.23) and the RS problem in (5.34) are related such that

$$\mathcal{R}(P_{\mathrm{t}}) \leq \mathcal{R}_{\mathrm{mix}}(P_{\mathrm{t}}) \leq \mathcal{R}_{\mathrm{RS}}(P_{\mathrm{t}}). \tag{5.35}$$

The left-hand side inequality in (5.35) follows by noting that solving (5.7) is equivalent to solving (5.23) under a restricted domain, where $\mathcal{M}_{\mathrm{ms}} = \mathcal{M}$. On the other hand, the

right-hand side inequality in (5.35) follows from the fact that solving (5.23) is equivalent to solving a restricted version of (5.34), through suppressing the common parts of groups in $\mathcal{M}_{\text{ms}}$ and the designated parts of groups in $\mathcal{M}_{\text{ss}}$, i.e. by setting $C_m = 0$ for all $m \in \mathcal{M}_{\text{ms}}$, and $\|\mathbf{p}_m\| = 0$ for all $m \in \mathcal{M}_{\text{ss}}$. Next, the relationship between the performances of the different schemes is characterize through DoF analysis.

5.5.1. DoF Analysis

The DoF performance of the RS scheme in (5.34) is characterized in the following result.

Theorem 5.2. *The optimum MMF-DoF achieved by solving* (5.34) *is given by*

$$d^*_{\text{RS}} \triangleq \lim_{P_{\text{t}}\to\infty} \frac{\mathcal{R}_{\text{RS}}(P_{\text{t}})}{\log_2(P_{\text{t}})} = \frac{1}{1 + M - M_{\text{ms}}} \tag{5.36}$$

where M_{ms} *is given in* (5.26).

The achievability of (5.36) follows from Proposition 5.1 and (5.35). In particular, since this MMF-DoF is achievable by the mixed scheme, then it is achievable by the RS scheme, as the former's solution is feasible for the latter. The converse on the other hand is relegated to Appendix C.2. Theorem 5.2 implies that the achievable MMF-DoF in Proposition 5.1 is optimum for the mixed scheme, as we have $d^*_{\text{RS}} \geq d^*_{\text{mix}}$, which follows from (5.35).

In Section 5.4.2, we gained some insight into how the MMF-DoF achieved through the mixed scheme and RS, given in (5.28) and (5.36), compares to the one achieved by the conventional scheme given in (5.13). This is revisited here in more detail. To see the full picture, we substitute the expression of M_{ms} in (5.36) into the MMF-DoF in (5.36) yielding

$$d^*_{\text{RS}} = \begin{cases} 1, & 1 + \sum_{j=2}^{M} G_j \leq N_{\text{t}} \\ \frac{1}{2}, & 1 + \sum_{j=2}^{M-1} G_j \leq N_{\text{t}} < 1 + \sum_{j=2}^{M} G_j \\ \frac{1}{3}, & 1 + \sum_{j=2}^{M-2} G_j \leq N_{\text{t}} < 1 + \sum_{j=2}^{M-1} G_j \\ \vdots & \vdots \\ \frac{1}{M-1}, & 1 + G_2 \leq N_{\text{t}} < 1 + G_2 + G_3 \\ \frac{1}{M}, & 1 \leq N_{\text{t}} < 1 + G_2 \end{cases} \tag{5.37}$$

Now, it is easy to compare d^*_{RS} in (5.37) with d^* in (5.13). It can be seen that the condition under which the MMF-DoF drops below 1 is the same in (5.13) and (5.37), i.e. $N_{\text{t}} < N_1$, which follows directly from Lemma 5.1. However, what is interesting is how the two performances differ once such condition is violated.

As observed in Section 5.4.2, the range of N_{t} under which a MMF-DoF of 0.5 is maintained is wider in (5.37) than it is in (5.13). Moreover, once the condition that guarantees $d^*_{\text{RS}} = 0.5$ is violated, d^*_{RS} drops to $1/3$ and so forth until it reaches a lower-bound of $1/M$. The RS and the mixed schemes exploit all the interference-free multi-stream beamforming capabilities available, and compensate for the rest through the common stream. This

is carried out until N_t cannot support any form of interference-free multi-stream beamforming anymore. In this case, all groups share the inherent DoF of 1 achieved through interference-free single-stream transmission, with $1/M$ for each group.

On the other hand, the conventional scheme that relies solely on multi-stream beamforming looses all DoF capabilities once N_t drops below N_M. In this case, the number of antennas is not sufficient to align all inter-group interference in the path of one group. This creates mutual interference between at least two groups, in which one group's gain is the other's loss. As a result, the MMF-DoF collapses to zero.

5.5.2. Precoder Optimization

The RS problem in (5.34) is solved by invoking the WMMSE algorithm. Since the WMMSE method has already been used multiple times in this thesis, it is only described briefly here to avoid repetition. First, we reformulate (5.34) into an equivalent smooth problem by introducing auxiliary variables to eliminate the point-wise minimizations. This yields

$$\mathcal{R}_{\mathrm{RS}}(P_\mathrm{t}): \begin{cases} \max\limits_{r,\mathbf{r},\mathbf{c},\mathbf{P}} \quad r \\ \text{s.t.} \quad C_m + r_m \geq r, \forall m \in \mathcal{M} \\ \qquad R_i \geq r_m, \forall i \in \mathcal{G}_m, \forall m \in \mathcal{M} \\ \qquad R_{\mathrm{c},k} \geq \sum_{m=1}^{M} C_m, \forall k \in \mathcal{K} \\ \qquad C_m \geq 0, \forall m \in \mathcal{M} \\ \qquad \|\mathbf{p}_\mathrm{c}\|^2 + \sum\limits_{m=1}^{M} \|\mathbf{p}_m\|^2 \leq P_\mathrm{t} \end{cases} \tag{5.38}$$

where at optimality, $\mathbf{r} \triangleq [r_1, \ldots, r_M]^T$ correspond to the rates of the designated streams, and r is the MMF group-rate. Then, the equivalent WMSE problem is formulated as

$$\bar{\mathcal{R}}_{\mathrm{RS}}(P_\mathrm{t}): \begin{cases} \max\limits_{r,\mathbf{r},\mathbf{c},\mathbf{P},\mathbf{g},\mathbf{u}} \quad r \\ \text{s.t.} \quad C_m + r_m \geq r, \forall m \in \mathcal{M} \\ \qquad 1 - \xi_i \geq r_m, \forall i \in \mathcal{G}_m, \forall m \in \mathcal{M} \\ \qquad 1 - \xi_{\mathrm{c},k} \geq \sum_{m=1}^{M} C_m, \forall k \in \mathcal{K} \\ \qquad C_m \geq 0, \forall m \in \mathcal{M} \\ \qquad \|\mathbf{p}_\mathrm{c}\|^2 + \sum\limits_{m=1}^{M} \|\mathbf{p}_m\|^2 \leq P_\mathrm{t} \end{cases} \tag{5.39}$$

where ξ_i and $\xi_{\mathrm{c},k}$ are the WMSEs that correspond to R_i and $R_{\mathrm{c},k}$ respectively. $\mathbf{g}$ and $\mathbf{u}$ are the sets of equalizers and weights, with a pair of equalizers and a pair of weights for each user. The problem in (5.39) is solved using the AO principle in Section 3.1, where the convergence and KKT optimality results also apply.

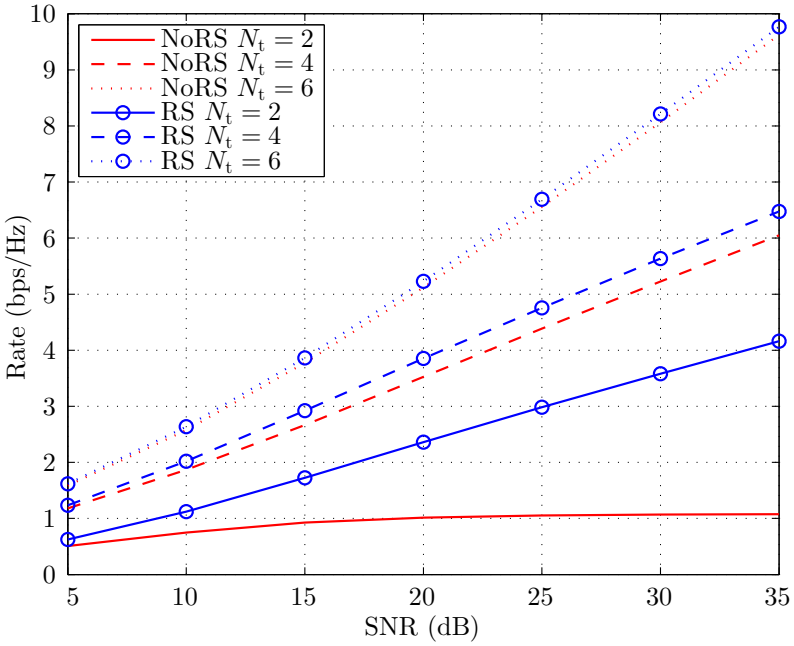


Figure 5.1.: The MMF rate for NoRS and RS under a varied number of antennas: $N_t = 2, 4$ and 6, in a setup with $K = 6$, $M = 3$, $G_1 = 1$, $G_2 = 2$ and $G_3 = 3$.

5.6. Simulation Results and Analysis

In this section, RS transmission for multigroup multicasting is evaluated through simulations. We consider i.i.d channels with entries drawn from $\mathcal{CN}(0, 1)$. All rate results are averaged over 100 channel realizations. RS transmission is compared to conventional beamforming (NoRS). Results for RS are obtained by solving (5.34) using the WMMSE approach, while results for NoRS are obtained by solving (5.7) using the SDR method in [90]. Note that for NoRS we present optimistic rate performances by plotting the SDR upper-bound without performing the randomization step in [90]. On the other hand, the RS results represent the actual achievable performance.

In Figure 5.1, we consider a system with $K = 6$ users divided into $M = 3$ groups with sizes $G_1 = 1$, $G_2 = 2$ and $G_3 = 3$. The number of transmitting antennas is varied such that $N_t = 2, 4$ and 6. For $N_t = 6$, it follows from Theorem 5.1 and Theorem 5.2 that both NoRS and RS achieve the full DoF. This is clear from Figure 5.1, where the two performances are almost identical. For $N_t = 4$, both schemes achieve a MMF-DoF of 0.5. However, RS gives a marginal improvement in the rate performance. This is due to the fact that the NoRS scheme is restricted to the strategy described in Section (5.3.2), while the RS scheme offers more flexibility, particularly when it comes to the largest group, through the common message and SIC. $N_t = 2$ is insufficient to achieve a non-zero MMF-DoF through beamforming, hence the NoRS rate saturates. The RS rate on the other hand keeps growing with a DoF of 1/3. The RS scheme is most beneficial in this overloaded regime, achieving

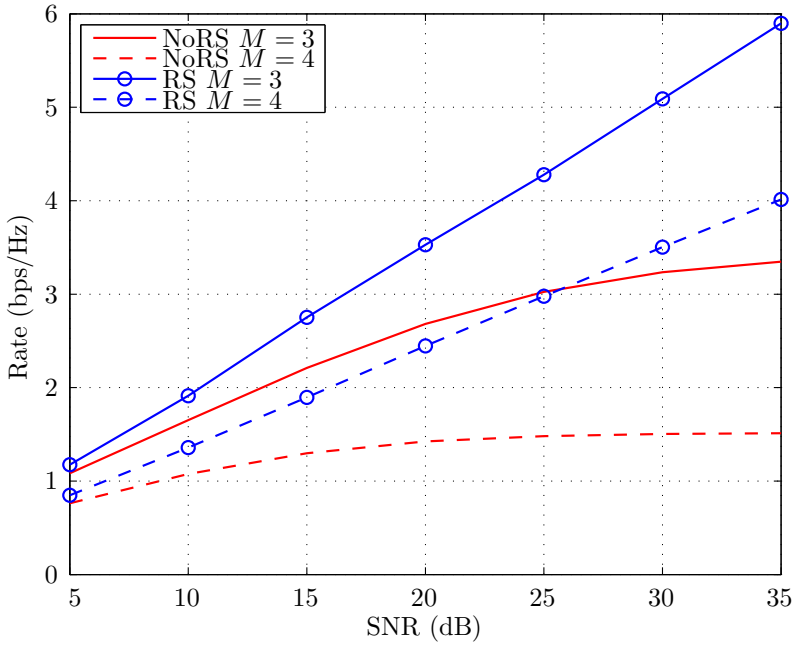


Figure 5.2.: The MMF rate for NoRS and RS in a setup with $N_{\mathrm{t}} = 4$ under a varied number groups: $M = 3$ and 4. Groups have equal sizes with 2 users per group.

significant gains over NoRS.

Next, we focus on the overloaded regime. In Figure 5.2, we consider a fixed number of antennas $N_{\mathrm{t}} = 4$ and a varied number of groups. We consider $M = 3$ and 4 groups, with 2 users per group for both cases. From Theorem 5.1, it can be seen that for both cases, we have zero MMF-DoF for NoRS. This is reflected in the saturating performances in Figure 5.2. Theorem 5.1 on the other hand suggests that RS achieves a MMF-DoF of 0.5 for $M = 3$, and 1/3 for $M = 4$. This is confirmed in Figure 5.2, where it is also demonstrated that RS achieves significant gains for the whole range of SNRs.

5.7. Summary and Conclusion

In this chapter, we considered the problem of achieving max-min fairness in multigroup multicasting systems. We derived the performance limits of the conventional beamforming scheme in terms of the MMF-DoF. The MMF-DoF was expressed as a function of the system parameters, including the number of transmitting antennas, the number of groups, and the distribution of users across groups. The interference-limited regime under which the MMF-DoF collapses to zero was identified. As we learned from the previous chapters, the collapsing MMF-DoF yields a saturating rate performance.

We proposed a strategy based on single-stream/multi-stream beamforming and interference cancellation. This was shown to improve the MMF-DoF performance, and avoid the saturating rate performance. It was then shown that this strategy is in fact a special case

of the RS strategy considered in the previous two chapters for MU-MISO systems with partial CSIT. Hence, the problem was reformulated in terms of RS and the precoders were optimized using the WMMSE approach. Simulations showed that RS provides tremendous gains over conventional beamforming techniques, particularly in interference-limited overloaded scenarios.

6. Conclusion

This thesis addressed a number of design and optimization problems in multiple antenna broadcasting scenarios with CSIT uncertainty in the light of the RS transmission strategy. A two-fold approach was taken for each of the considered scenarios such that:

1. The benefits of applying RS was analytically demonstrated through deriving asymptotic performance limits in terms of the DoF achieved by the optimum design.

2. An algorithm that realizes such high SNR benefits, while guaranteeing an optimized performance across the entire range of SNRs, was developed.

We started by considering the problem of ergodic sum-rate maximization in MU-MISO systems under statistical CSIT uncertainty in Chapter 3. To address this problem, we developed an average WMMSE algorithm based on SAA. The gains achieved by employing RS were demonstrated through simulations for a variety of system parameters. Moreover, we solved the weighted sum-rate maximization problem to obtain the boundary of the ergodic rate region achieved through RS. This highlighted the gains of applying RS to realize any tradeoff point in the achievable region.

In Chapter 4, we addressed the problem of achieving max-min fairness in MU-MISO systems under bounded CSIT uncertainties. We started by characterising the MMF-DoF (or symmetric-DoF) achieved through NoRS and RS designs in the general case where users experience different CSIT qualities. Then we developed a robust WMMSE algorithm based on the cutting-set method. The convergence of the developed algorithm to the set of KKT points of the semi-infinite rate problem was established. Simulation results showed the significant gains achieved by employing RS in such scenarios. The developed methods were extended to address the QoS constrained problem. It was shown that RS resolves the feasibility issue and reduces the transmission power.

In Chapter 5, the RS strategy was extended to the multigroup multicasting setup. We started by characterizing the DoF performance of the conventional beamforming strategy. We identified the antenna regime under which the symmetric-DoF collapses to zero, i.e. the overloaded regime. We then showed through DoF analysis that RS is superior to the conventional beamforming strategy. The precoder design problem was solved using a WMMSE algorithm, and significant performance gains achieved through RS were demonstrated through simulations.

6.1. Future Work

In conclusion of this these, some future research directions are proposed.

- **Rate-Splitting for Interfering Networks:** The work in this thesis was restricted to systems with only one BS[1]. RS however is not restricted to this model, and can be applied to networks with multiple interfering BSs. The design and optimization in such setups brings a whole new set of challenges. This is mainly due to the distributed nature of the transmitters, hence requiring distributed algorithms that operate under little or no cooperation. Prior work on distributed algorithms for multi-cell transmission can be leveraged in this context [8, 45]. In situations where RS is carried out for all users, the common massage cannot be constructed as described in Section 4.3.1. This will likely require a *hierarchy* of common messages. Progress towards the DoF analysis of such systems has been made in [99].

- **Non-linear Precoding:** Preprocessing techniques in this thesis were restricted to linear precoding. A question that comes to mind in this context is: how would a RS design based on non-linear precoding perform? While this is not expected to enhance the DoF performance, it may nontrivially benefit the rate performance at finite SNRs. Such non-linear RS design can be built upon existing works on robust non-linear precoder optimization [75].

- **Application to Information and Power Broadcasting Systems:** It was shown in Section 4.8 that RS transmission is more power-efficient compared to NoRS transmission. This leaves us wondering whether such power efficiency can be leveraged to enhance the performance of broadcasting systems with Simultaneous Wireless Information and Power Transfer (SWIPT) and imperfect CSIT [100]. It is not clear yet if the common stream can act as the energy signal in such setups, or if a dedicated energy signal is required on top.

- **Physical-Layer Secrecy:** We have seen that as the CSIT quality decreases, the RS transmission becomes increasingly dependent on the common message. Since the common message is broadcasted to all users in the system, this raises some questions about operation under physical-layer secrecy constraints. Addressing this requires an investigation of the secure DoF under partial CSIT, before moving into the design and optimization problems.

- **Overloaded Systems:** In Chapter 5 we considered the overloaded regime in multi-group multicasting systems. The overloaded regime; however, is not restricted to multigroup multicasting and can also occur in MU-MISO transmission where each

[1]The results and schemes also apply to systems with multiple fully-cooperating BSs. This may require per-transmitter power constraints that can be easily incorporated into problem formulations.

receiver requests a private message. Such scenarios have not been thoroughly addressed in literature. Imagine a system with $K > N_{\mathrm{t}}$, and heterogeneous CSIT qualities across users. Maximizing the sum-DoF necessitates scheduling the K users with the highest CSIT qualities. If we assume that the rest of the users request low rates, for example low-power sensors, then it may be sufficient to serve them through the common message without influencing the achievable sum-DoF. Moreover, their messages could be transmitted in a non-orthogonal manner with multilayered SIC to further enhance their achievable rates, drawing a connection with Non-Orthogonal Multiple Access (NOMA) techniques [101].

- **Relation to Coded Caching:** In a recent work [102], the relationship between coded caching [103] and the MISO-BC with imperfect CSIT was established. This work demonstrates that coded caching can be used to relax the constraints on the required CSIT quality. This follows from understanding the fundamental role of CSIT: that is eliminating multiuser interference, and the nature of coded caching: that is creating multicast transmission opportunities which are robust with respect to CSIT requirements. The achievability scheme in [102] is based on the RS strategy. In realistic scenarios, the number of subscribed users may exceed N_{t} as pointed out above. A naive scheme would be to perform the transmission in [102] over orthogonal time slots, by scheduling $K = N_{\mathrm{t}}$ users at a time. However, since coded caching relaxes CSIT requirements and reduces the multiuser interference, it may be possible to develop a scheme that serves an overloaded setup with $K > N_{\mathrm{t}}$ users in a non-orthogonal manner in less time. This needs further research and investigation.

A. Proofs for Chapter 3

A.1. Proof of Theorem 3.1

Since $\mathbf{P}$ is designed separately for each $\widehat{\mathbf{H}}$, we focus on a precoding scheme defined for a given $\widehat{\mathbf{H}}$ as $\{\mathbf{P}(P_{\mathrm{t}})\}_{P_{\mathrm{t}}}$. We recall that the powers allocated to the common stream and the kth private stream are given by $q_{\mathrm{c}} \triangleq \|\mathbf{p}_{\mathrm{c}}\|^2$ and $q_k \triangleq \|\mathbf{p}_k\|^2$ respectively. We assume that such powers scale with increased P_{t} as $O(P_{\mathrm{t}}^{a_{\mathrm{c}}})$ and $O(P_{\mathrm{t}}^{a_k})$ respectively, where $a_{\mathrm{c}}, a_k \in [0,1]$ are power scaling factors. The streams interfering with the kth user are dominated by a power scaling factor of $\bar{a}_k \triangleq \max_j\{a_j\}_{j\neq k}$. We define the kth user's conditional average DoFs as

$$\bar{d}_{\mathrm{c},k} \triangleq \lim_{P_{\mathrm{t}}\to\infty} \frac{\bar{R}_{\mathrm{c},k}(P_{\mathrm{t}})}{\log_2(P_{\mathrm{t}})} \quad \text{and} \quad \bar{d}_k \triangleq \lim_{P_{\mathrm{t}}\to\infty} \frac{\bar{R}_k(P_{\mathrm{t}})}{\log_2(P_{\mathrm{t}})}. \tag{A.1}$$

For a given precoding scheme defined over all $\widehat{\mathbf{H}}$, the long-term DoFs are given by

$$d_{\mathrm{c},k} \triangleq \lim_{P_{\mathrm{t}}\to\infty} \frac{\mathrm{E}_{\widehat{\mathrm{H}}}\{\bar{R}_{\mathrm{c},k}(P_{\mathrm{t}})\}}{\log_2(P_{\mathrm{t}})} = \mathrm{E}_{\widehat{\mathrm{H}}}\{\bar{d}_{\mathrm{c},k}\} \quad \text{and} \quad d_k \triangleq \lim_{P_{\mathrm{t}}\to\infty} \frac{\mathrm{E}_{\widehat{\mathrm{H}}}\{\bar{R}_k(P_{\mathrm{t}})\}}{\log_2(P_{\mathrm{t}})} = \mathrm{E}_{\widehat{\mathrm{H}}}\{\bar{d}_k\} \tag{A.2}$$

which follows from the bounded convergence theorem. The same definitions extend to NoRS precoding schemes while discarding the common power. Without loss of generality, $\sigma_{\mathrm{n}}^2 = 1$ is assumed throughout the proof.

Proof of (3.14)

For an arbitrary precoding scheme, let us find an upper-bound for

$$\bar{R}_k = \mathrm{E}_{\mathrm{H}|\widehat{\mathrm{H}}}\{\log_2(T_k) - \log_2(I_k) \mid \widehat{\mathbf{H}}\} \tag{A.3}$$

by upper-bounding and lower-bounding the first and second right-hand-side terms respectively. From Jensen's inequality, we write

$$\mathrm{E}_{\mathrm{H}|\widehat{\mathrm{H}}}\{\log_2(T_k) \mid \widehat{\mathbf{H}}\} \leq \log_2(\bar{T}_k) \tag{A.4}$$

where $\bar{T}_k$ is defined in (3.35b). From the Cauchy-Schwarz inequality and the isotropic property of the CSIT errors, we have

$$\bar{T}_k \leq \left(\|\widehat{\mathbf{h}}_k\|^2 + \sigma_{\mathrm{e}}^2\right) \sum_{i=1}^{K} q_i + 1 \tag{A.5}$$

where $\sigma_{\mathrm{e}}^2 = N_t^{-1}\sigma_{\mathrm{e},k}^2$. Since the actual channel state $\mathbf{H}$ does not depend on SNR, we have $\|\mathbf{h}_k\|^2, \|\widehat{\mathbf{h}}_k\|^2 = O(1)$. It follows that $\bar{T}_k \leq O\big(P_{\mathrm{t}}^{\max\{a_k,\bar{a}_k\}}\big)$, from which we write

$$\mathrm{E}_{\mathrm{H}|\widehat{\mathrm{H}}}\big\{\log_2(T_k) \mid \widehat{\mathbf{H}}\big\} \leq \max\{a_k,\bar{a}_k\}\log_2(P_{\mathrm{t}}) + O(1). \tag{A.6}$$

Next, from the isotropic property and [2, Lemma 1], we write

$$\mathrm{E}_{\mathrm{H}|\widehat{\mathrm{H}}}\big\{\log_2(I_k) \mid \widehat{\mathbf{H}}\big\} \geq \log_2(2^{\kappa}\sigma_{\mathrm{e}}^2\lambda_1 + 1) + O(1) \tag{A.7}$$

where $\kappa \triangleq \mathrm{E}_{\mathrm{H}|\widehat{\mathrm{H}}}\left\{\log_2\left(\frac{|\mathbf{e}_1^T\widetilde{\mathbf{h}}_k|^2}{\sigma_{\mathrm{e}}^2}\right) \mid \widehat{\mathbf{H}}\right\}$ is bounded [2], and λ_1 is the dominant eigenvalue of $\sum_{i\neq k}\mathbf{p}_i\mathbf{p}_i^H$. Since the maximum is lower-bounded by the average, we write

$$\lambda_1 \geq N_{\mathrm{t}}^{-1}\sum_{i\neq k} q_i = O\big(P_{\mathrm{t}}^{\bar{a}_k}\big) \tag{A.8}$$

from which we obtain

$$\mathrm{E}_{\mathrm{H}|\widehat{\mathrm{H}}}\big\{\log_2(I_k) \mid \widehat{\mathbf{H}}\big\} \geq (\bar{a}_k - \alpha)^+\log_2(P_{\mathrm{t}}) + O(1). \tag{A.9}$$

Form (A.1), (A.3), (A.6) and (A.9), we have

$$\bar{d}_k \leq \max\{a_k,\bar{a}_k\} - \max\{\bar{a}_k - \alpha, 0\}. \tag{A.10}$$

When $a_k \geq \bar{a}_k$, the upper-bound is given by $\min\{\alpha + a_k - \bar{a}_k, a_k\}$, otherwise we have 0. Combining the two cases, we write $\bar{d}_k \leq \min\left\{(\alpha + a_k - \bar{a}_k)^+, a_k\right\}$, from which the sum DoF is upper-bounded by

$$\sum_{k=1}^{K}\bar{d}_k \leq \sum_{k=1}^{K}\min\left\{(\alpha + a_k - \bar{a}_k)^+, a_k\right\}. \tag{A.11}$$

To obtain the maximum upper-bound for (A.11), we define $\mathcal{J} \subseteq \mathcal{K}$ as the subset composed of all users with non-zero DoF. Due to the symmetry in the CSIT qualities, it is sufficient to assume, without loss of generality, that $\mathcal{J} \triangleq \{1,\ldots,J\}$ with $J \triangleq |\mathcal{J}|$. We define $\bar{d}(J) \triangleq \sum_{j=1}^{J}\bar{d}_j$. It is evident that $\bar{d}(1) \leq 1$. On the other hand, we have $\bar{d}(J) \leq J\alpha$ for $J > 1$. This is shown from

$$\sum_{k=1}^{J}\bar{d}_k \leq J\alpha + \sum_{k=1}^{J}(a_k - \bar{a}_k) \tag{A.12a}$$

$$\leq J\alpha + (a_J - a_1) + \sum_{k=1}^{J-1}(a_k - a_{k+1}) \tag{A.12b}$$

where (A.12a) follows from $\min\{x,y\} \leq x, y$, and the positivity of DoF assumption. (A.12b) follows from $\bar{a}_k \geq a_j$, $\forall j \neq k$, and is equal to $J\alpha$. We conclude that the maximum upper-

bound is obtained when $J = K$ for $\alpha \geq K^{-1}$, and $J = 1$ otherwise. Hence, we obtain

$$\sum_{k=1}^{K} \bar{d}_k \leq \max\{1, K\alpha\}. \tag{A.13}$$

From the bounded convergence theorem, it follows that $\lim_{P_\mathrm{t}\to\infty} \frac{\mathrm{E}_{\widehat{\mathrm{H}}}\{\mathcal{R}(P_\mathrm{t})\}}{\log_2(P_\mathrm{t})}$ is also upper-bounded by the right-hand side quantity in (A.13).

This part of the proof is completed by showing that the upper-bound can be achieved by a feasible precoding scheme. It is easy to show that a DoF of 1 is achieved by single-user transmission, i.e. transmitting to one user only while switching off all other users. On the other hand, $K\alpha$ is achieved using ZF-BF as shown in Section 2.5.1. Note that $\max\{1, K\alpha\}$ is achieved by selecting between the two modes depending on α.

Proof of (3.15)

From the definitions of the ARs, we write

$$\bar{R}_\mathrm{c} + \bar{R}_1 \leq \bar{R}_{\mathrm{c},1} + \bar{R}_1 = \mathrm{E}_{\mathrm{H}|\widehat{\mathrm{H}}}\{\log_2(T_{\mathrm{c},1}) - \log_2(I_1) \mid \widehat{\mathbf{H}}\}. \tag{A.14}$$

Following the same approach used in the previous part, we obtain

$$\mathrm{E}_{\mathrm{H}|\widehat{\mathrm{H}}}\{\log_2(T_{\mathrm{c},1}) \mid \widehat{\mathbf{H}}\} \leq \log_2(P_\mathrm{t}) + O(1) \tag{A.15}$$

and

$$\mathrm{E}_{\mathrm{H}|\widehat{\mathrm{H}}}\{\log_2(I_1) \mid \widehat{\mathbf{H}}\} \geq (\bar{a}_1 - \alpha)^+ \log_2(P_\mathrm{t}) + O(1) \tag{A.16}$$

from which we write

$$\bar{d}_{\mathrm{c},1} + \bar{d}_1 \leq \min\{1 + \alpha - \bar{a}_1, 1\}. \tag{A.17}$$

Next, we define $\bar{d}_\mathrm{RS}(J) \triangleq \bar{d}_\mathrm{c} + \sum_{k=1}^{J} \bar{d}_k$ for J users with positive DoFs. From (A.17), we have $\bar{d}_\mathrm{RS}(1) \leq 1$. For $J = 2$, we have $\bar{d}_\mathrm{RS}(2) \leq 1 + \alpha - \bar{a}_1 + a_2 \leq 1 + \alpha$, obtained from (A.17), $\bar{d}_2 \leq a_2$, and $\bar{a}_1 \geq a_2$. For $J > 2$, we have $\bar{d}_\mathrm{RS}(J) \leq 1 + (J-1)\alpha$ obtained by combining (A.17) and (A.12). Hence, the maximum upper-bound is $1 + (K-1)\alpha$ obtained when $J = K$ regardless of α. It follows from the bounded convergence theorem that this also acts as an upper-bound for $\lim_{P_\mathrm{t}\to\infty} \frac{\mathrm{E}_{\widehat{\mathrm{H}}}\{\mathcal{R}_\mathrm{RS}(P_\mathrm{t})\}}{\log_2(P_\mathrm{t})}$.

Finally, the achievability for this part is given in Section 2.5.3. □

B. Proofs for Chapter 4

B.1. Important Lemmas for the proof of Theorem 4.1

The following lemmas are instrumental to the proof of Theorem 4.1.

Lemma B.1. [45, Lemma 1] *Given the ball uncertainty model and for any* $\mathbf{p}$*, we have*

$$\max_{\mathbf{h}_k \in \mathbb{H}_k} |\mathbf{h}_k^H \mathbf{p}| = |\widehat{\mathbf{h}}_k^H \mathbf{p}| + \delta_k \|\mathbf{p}\|$$
$$\min_{\mathbf{h}_k \in \mathbb{H}_k} |\mathbf{h}_k^H \mathbf{p}| = \big(|\widehat{\mathbf{h}}_k^H \mathbf{p}| - \delta_k \|\mathbf{p}\|\big)^+.$$

Lemma B.2. *There exists a feasible RS precoding scheme that achieves the DoF*

$$\hat{d}_{\mathrm{c}} = 1 - \bar{a} \quad \textit{and} \quad \hat{d}_k = \min\{(\alpha_k + a_k - \bar{a}_k)^+, a_k\} \tag{B.1}$$

for all $a_k \in [0,1]$*,* $\bar{a} \triangleq \max_j \{a_j\}_{j=1}^K$ *and* $\bar{a}_k \triangleq \max_j \{a_j\}_{j \neq k}$*.*

Proof. For the private precoder, consider the ZF-BF design in (2.24) based on the imperfect estimate $\widehat{\mathbf{H}}$. Recall that this is given by $\mathbf{P}_{\mathrm{p}} = \big(\widehat{\mathbf{H}}^H\big)^\dagger \mathbf{B}\ \mathrm{diag}(\sqrt{q_1}, \ldots, \sqrt{q_K})$, where $\mathbf{B} \triangleq \mathrm{diag}\big(\sqrt{1/b_1}, \ldots, \sqrt{1/b_K}\big)$, and $b_1, \ldots, b_K$ are constants that normalize the columns of $\big(\widehat{\mathbf{H}}^H\big)^\dagger$. The existence of such solution is guaranteed by Assumption 4.1. The common precoder is given as $\mathbf{p}_{\mathrm{c}} = \sqrt{q_{\mathrm{c}}}\mathbf{e}_1$, where a_{c} is set to 1. We define the worst-case SINRs[1] as

$$\bar{\gamma}_{\mathrm{c},k} \triangleq \min_{\mathbf{h}_k \in \mathbb{H}_k} \gamma_{\mathrm{c},k}\big(\mathbf{h}_k\big) \quad \text{and} \quad \bar{\gamma}_k \triangleq \min_{\mathbf{h}_k \in \mathbb{H}_k} \gamma_k\big(\mathbf{h}_k\big). \tag{B.2}$$

By applying the described scheme, $\bar{\gamma}_k$ is lower-bounded as

$$\bar{\gamma}_k \geq \frac{q_k(\sqrt{1/b_k} - \delta_k)^2}{\sum_{i \neq k} \big|\widetilde{\mathbf{h}}_k^H \mathbf{p}_i\big|^2 + \sigma_{\mathrm{n}}^2} \geq \frac{q_k(\sqrt{1/b_k} - \delta_k)^2}{\delta_k^2 \sum_{i \neq k} q_i + \sigma_{\mathrm{n}}^2}. \tag{B.3}$$

The left inequality in (B.3) follows from Lemma B.1 and the assumption that $|\widehat{\mathbf{h}}_k^H \mathbf{p}_k| > \delta_k q_k$, i.e. small error [104]. The right inequality is obtained from applying the Cauchy-Schwarz inequality and $\|\widetilde{\mathbf{h}}_k\|^2 \leq \delta_k^2$ to the denominator. The numerator scales as $O\big(P_{\mathrm{t}}^{a_k}\big)$, while the denominator scales as $O\big(P_{\mathrm{t}}^{(\bar{a}_k - \alpha_k)^+}\big)$. It follows that $\bar{d}_k \geq \min\{(\alpha_k + a_k - \bar{a}_k)^+, a_k\}$.

[1] Worst-case channels are equivalently obtained using the rates or SINRs

For $\bar{\gamma}_{\mathrm{c},k}$, we write

$$\bar{\gamma}_{\mathrm{c},k} \geq \frac{q_{\mathrm{c}}|h_{k,1}|^2}{\|\mathbf{h}_k\|^2 \sum_{k=1}^{K} q_k + \sigma_{\mathrm{n}}^2} = O\big(P_{\mathrm{t}}^{(1-\bar{a})}\big) \tag{B.4}$$

where the Cauchy-Schwarz inequality is applied to the denominator, and both $|h_{k,1}|^2$ and $\|\mathbf{h}_k\|^2$ scale as $O(1)$ from Assumption 4.1. It follows that $\bar{d}_c \geq 1 - \bar{a}$. □

B.2. Proof of Theorem 4.1

To characterize the optimum DoF performance, we define the optimum precoding schemes for (4.9) and (4.5) as $\big\{\mathbf{P}^*(P_{\mathrm{t}})\big\}_{P_{\mathrm{t}}}$ and $\big\{\mathbf{P}_{\mathrm{p}}^*(P_{\mathrm{t}})\big\}_{P_{\mathrm{t}}}$ respectively, where the corresponding powers and exponents are denoted by q_{c}^*, q_k^*, a_{c}^* and a_k^*.

Proof of (4.11)

We start by showing that for any given precoding scheme with a given power allocation, the achievable private DoF in Lemma B.2 cannot be exceeded, i.e.

$$\bar{d}_k \leq \min\left\{(\alpha_k + a_k - \bar{a}_k)^+, a_k\right\}. \tag{B.5}$$

The worst-case SINR is upper-bounded as $\bar{\gamma}_k \leq \gamma_k(\mathbf{h}_k)$, where $\mathbf{h}_k \in \mathbb{H}_k$. $\mathbf{h}_k$ is selected such that the lth user's interference term is maximized in accordance with Lemma B.1, i.e. $\big|\widehat{\mathbf{h}}_k^H \mathbf{p}_l + \widetilde{\mathbf{h}}_k^H \mathbf{p}_l\big| = \big|\widehat{\mathbf{h}}_k^H \mathbf{p}_l\big| + \delta_k \|\mathbf{p}_l\|$, where l is chosen such that $a_l = \bar{a}_k \triangleq \max\left\{a_j\right\}_{j\neq k}$. As a result, we obtain the upper-bound

$$\begin{aligned}\bar{\gamma}_k &\leq \frac{\big|\mathbf{h}_k^H \mathbf{p}_k\big|^2}{\big(|\widehat{\mathbf{h}}_k^H \mathbf{p}_l| + \delta_k \|\mathbf{p}_l\|\big)^2 + \sum\limits_{i\neq k,l} \big|\widehat{\mathbf{h}}_k^H \mathbf{p}_i + \widetilde{\mathbf{h}}_k^H \mathbf{p}_i\big|^2 + \sigma_{\mathrm{n}}^2} \\ &\leq \frac{\|\mathbf{h}_k\|^2 q_k}{\delta_k^2 q_l + \sigma_{\mathrm{n}}^2}. \end{aligned} \tag{B.6}$$

where (B.6) follows from applying the Cauchy-Schwarz inequality and discarding non-negative interference terms. From Assumption 4.1, it is evident that (B.6) scales as the lower-bound in (B.3), from which (B.5) directly follows.

The optimum DoF satisfies $\bar{d}^* \leq \bar{d}_k^*, \forall k \in \mathcal{K}$, where $\bar{d}_k^*$ is the kth user's DoF at optimality. From (B.5), we write

$$\bar{d}^* \leq \min_k \left\{ \min\{\alpha_k + a_k^* - \bar{a}_k^*, a_k^*\} \right\}_{k=1}^{K} \tag{B.7}$$

where $(\cdot)^+$ is omitted by assuming that $(\alpha_k + a_k^* - \bar{a}_k^*) \geq 0$. This assumption is valid as $(\alpha_k + a_k^* - \bar{a}_k^*) < 0$ yields $\bar{d}^* = 0$, which is maintained if a_k^* is increased to $\bar{a}_k^* - \alpha_k$. On the other hand, $\bar{d}^* > 0$ is only obtained when $(\alpha_k + a_k^* - \bar{a}_k^*) > 0$. (B.7) is further upper-bounded

as

$$\bar{d}^* \leq \frac{\min\{\alpha_1+a_1^*-\bar{a}_1^*, a_1^*\} + \min\{\alpha_2+a_2^*-\bar{a}_2^*, a_2^*\}}{2} \tag{B.8}$$

$$\leq \frac{\alpha_1 + a_1^* - \bar{a}_1^* + \alpha_2 + a_2^* - \bar{a}_2^*}{2} \tag{B.9}$$

$$\leq \frac{\alpha_1 + \alpha_2}{2}. \tag{B.10}$$

(B.8) follows from the fact that $\bar{d}^*$ is upper-bounded by the average of any two DoFs, and (B.9) is obtained by noting that the point-wise minimum is upper-bounded by any element in the set. (B.10) follows from $a_j^* \leq \bar{a}_k^*$, $\forall j \neq k$. From Lemma B.2, allocating the private powers such that $a_1 = \alpha_2$ and $a_2, \ldots, a_K = \frac{\alpha_1+\alpha_2}{2}$, we achieve $\bar{d}_k \geq \frac{\alpha_1+\alpha_2}{2}$, $\forall k \in \mathcal{K}$.

Proof of (4.12)

We start this part by showing that

$$\bar{d}_{\mathrm{c}} + \bar{d}_k \leq \min\{1 + \alpha_k - \bar{a}_k, 1\}. \tag{B.11}$$

This result follows from

$$\bar{R}_{\mathrm{c}} + \bar{R}_k \leq \bar{R}_{\mathrm{c},k} + \bar{R}_k$$

$$\leq R_{\mathrm{c},k}(\mathbf{h}_k) + R_k(\mathbf{h}_k) \tag{B.12}$$

$$= \log_2\left(T_{\mathrm{c},k}(\mathbf{h}_k)\right) - \log_2\left(I_k(\mathbf{h}_k)\right) \tag{B.13}$$

$$= \log_2(P_{\mathrm{t}}) - (\bar{a}_k - \alpha_k)^+ \log_2(P_{\mathrm{t}}) + O(1) \tag{B.14}$$

where (B.12) is obtained using the same $\mathbf{h}_k$ employed in (B.6), (B.13) follows from the rate definitions, and (B.14) is obtained using means of previous analysis.

The optimum DoF satisfies $\bar{d}_{\mathrm{RS}}^* \leq \bar{c}_k^* + \bar{d}_k^*$, $\forall k \in \mathcal{K}$, where $\sum_{k=1}^K \bar{c}_k^* = \bar{d}_{\mathrm{c}}^*$ and $\bar{c}_k^* \geq 0$. An upper-bound is obtained by taking the average of any number of user DoFs. To obtain a tighter upper-bound, we optimize over the number of averaged users such that

$$\bar{d}_{\mathrm{RS}}^* \leq \min_{J \in \mathcal{K}} \frac{\sum_{k=1}^J \left(\bar{c}_k^* + \bar{d}_k^*\right)}{J}$$

$$\leq \min_{J \in \mathcal{K}} \frac{\bar{d}_{\mathrm{c}}^* + \sum_{k=1}^J \bar{d}_k^*}{J} \tag{B.15}$$

where (B.15) follows from $\sum_{k=1}^J \bar{c}_k^* \leq \bar{d}_{\mathrm{c}}^*$. The argument used to omit $(\cdot)^+$ in (B.7) cannot be directly applied for (B.15). Alternatively, we start by assuming that $(\alpha_k + a_k^* - \bar{a}_k^*) \geq$

0, $\forall k \in \{1, \ldots, J\}$, for a given J. For the case where J is an odd number, we write

$$\begin{aligned}\frac{\bar{d}_{\mathrm{c}}^* + \sum_{k=1}^{J} \bar{d}_k^*}{J} &= \frac{\bar{d}_{\mathrm{c}}^* + \bar{d}_J^* + \sum_{k=1}^{J-1} \bar{d}_k^*}{J} \\ &\leq \frac{1 + \sum_{k=1}^{J-1} (\alpha_k + a_k^* - \bar{a}_k^*)}{J} \qquad \text{(B.16)} \\ &\leq \frac{1 + \sum_{k=1}^{J-1} \alpha_k}{J}. \qquad \text{(B.17)}\end{aligned}$$

(B.16) follows from (B.11) and (B.5), where the elements 1 and $(\alpha_k + a_k^* - \bar{a}_k^*)$ are picked to upper-bound $\bar{d}_{\mathrm{c}}^* + \bar{d}_J^*$ and $\bar{d}_k^*$ respectively. (B.17) is obtained by writing the sum in (B.16) as a sum of pairs, and using the approach in (B.10). For the case where J is an even number, we write

$$\begin{aligned}\frac{\bar{d}_{\mathrm{c}}^* + \sum_{j=1}^{J} \bar{d}_j^*}{J} &= \frac{\bar{d}_{\mathrm{c}}^* + \bar{d}_1^* + \bar{d}_J^* + \sum_{k=2}^{J-1} \bar{d}_k^*}{J} \\ &\leq \frac{1 + \alpha_1 - \bar{a}_1^* + a_J^* + \sum_{k=2}^{J-1} (\alpha_k + a_k^* - \bar{a}_k^*)}{J} \qquad \text{(B.18)} \\ &\leq \frac{1 + \sum_{k=1}^{J-1} \alpha_k}{J}. \qquad \text{(B.19)}\end{aligned}$$

In (B.18), $1 + \alpha_1 - \bar{a}_1^*$ and a_J^* are chosen to upper-bound $\bar{d}_{\mathrm{c}}^* + \bar{d}_1^*$ and $\bar{d}_J^*$ respectively. (B.19) is obtained from $\bar{a}_1^* \geq a_J^*$ and the approach in (B.17). If we assume that a given $(\alpha_k + a_k^* - \bar{a}_k^*) < 0$ for a subset of $\{1, \ldots, J\}$, and hence $\bar{d}_j^* = 0$, we cannot exceed (B.17) and (B.19). Combining this with (B.15), we obtain

$$\bar{d}_{\mathrm{RS}}^* \leq \min_{J \in \{2, \ldots, K\}} \frac{1 + \sum_{j=1}^{J-1} \alpha_j}{J}. \qquad \text{(B.20)}$$

where $J = 1$ has been omitted. Next, we show that this upper-bound is achievable by a feasible precoding scheme. From Lemma B.2, allocating the powers such that $a_k = \bar{a}$ for all k, we achieve DoFs $\bar{d}_k$ and $\bar{d}_{\mathrm{c}}$ of $\min\{\alpha_k, \bar{a}\}$ and $1 - \bar{a}$ respectively. We show that there exists $\bar{a} \in [0, 1]$ and feasible $\{\bar{c}_k\}_{k=1}^K$ such that $\bar{c}_k + \min\{\alpha_k, \bar{a}\}$ achieves the upper-bound in (B.20).

For a given J, the corresponding upper-bound $\frac{1 + \sum_{j=1}^{J-1} \alpha_j}{J}$ is denoted by $\bar{d}_{\mathrm{RS}}^{\mathrm{UB}}(J)$. Let J^* be the argument of the minimization in (B.20), i.e. $\bar{d}_{\mathrm{RS}}^{\mathrm{UB}}(J^*) \leq \bar{d}_{\mathrm{RS}}^{\mathrm{UB}}(J)$. If $J^* < K$,

$$\alpha_{J^*-1} \leq \bar{d}_{\mathrm{RS}}^{\mathrm{UB}}(J^*) \leq \alpha_{J^*} \qquad \text{(B.21)}$$

which is shown in the following. First, we note that

$$\begin{aligned}\bar{d}_{\mathrm{RS}}^{\mathrm{UB}}(J+1) &= \frac{J \bar{d}_{\mathrm{RS}}^{\mathrm{UB}}(J) + \alpha_J}{J+1} \qquad \text{(B.22)} \\ \bar{d}_{\mathrm{RS}}^{\mathrm{UB}}(J-1) &= \frac{J \bar{d}_{\mathrm{RS}}^{\mathrm{UB}}(J) - \alpha_{J-1}}{J-1}. \qquad \text{(B.23)}\end{aligned}$$

Since $\bar{d}_{\mathrm{RS}}^{\mathrm{UB}}(J^*) \leq \bar{d}_{\mathrm{RS}}^{\mathrm{UB}}(J^*+1)$ and $\bar{d}_{\mathrm{RS}}^{\mathrm{UB}}(J^*) \leq \bar{d}_{\mathrm{RS}}^{\mathrm{UB}}(J^*-1)$, the right and left inequalities in (B.21) follow from (B.22) and (B.23) respectively, as the average increases by including α_J in (B.22) and excluding α_{J-1} in (B.23). For $J^* = K$, the right inequality in (B.21) does not necessarily hold, but the left inequality always holds. Hence, we have one of the two following cases.

- **Case $J^* < K$, or $J^* = K$ and (B.21) holds:** For this case, we set $\bar{a} = \bar{d}_{\mathrm{RS}}^{\mathrm{UB}}(J^*)$. We obtain DoFs of $\bar{d}_k = \alpha_k,\ \forall k < J^*$, $\bar{d}_k = \bar{d}_{\mathrm{RS}}^{\mathrm{UB}}(J^*),\ \forall k \geq J^*$, and $\bar{d}_{\mathrm{c}} = 1 - \bar{d}_{\mathrm{RS}}^{\mathrm{UB}}(J^*)$. The common DoF is split such that $\bar{c}_k = \bar{d}_{\mathrm{RS}}^{\mathrm{UB}}(J^*) - \alpha_k,\ \forall k < J^*$, and $\bar{c}_k = 0,\ \forall k \geq J^*$. The left inequality in (B.21) guarantees that $\bar{c}_k \geq 0$, while we can see that $\sum_{k=1}^{J^*-1} \bar{c}_k = (J^*-1)\bar{d}_{\mathrm{RS}}^{\mathrm{UB}}(J^*) - \sum_{k=1}^{J^*-1} \alpha_k = \bar{d}_{\mathrm{c}}$.
- **Case $J^* = K$ and $\bar{d}_{\mathrm{RS}}^{\mathrm{UB}}(K) \geq \alpha_k,\ \forall k \in \mathcal{K}$:** We set $\bar{a} = \alpha_K$ obtaining DoFs of $\bar{d}_k = \alpha_k$ and $\bar{d}_{\mathrm{c}} = 1 - \alpha_K$. The common DoF is split as: $\bar{c}_k = \bar{d}_{\mathrm{RS}}^{\mathrm{UB}}(K) - \alpha_K$, which are non-negative and satisfy $\sum_{k=1}^{K} \bar{c}_k = \bar{d}_{\mathrm{c}}$.

This completes the proof. □

B.3. Proof of Proposition 4.1

Consider the semi-infinite optimization problem

$$\begin{aligned} \min_{\mathbf{x}} \quad & f_0(\mathbf{x}) \\ \text{s.t.} \quad & f_m(\mathbf{x}, t) \leq 0,\ \forall t \in \mathcal{T}_m, m \in \mathcal{M} \end{aligned} \tag{B.24}$$

where $\mathcal{M} \triangleq \{1, \ldots, M\}$, and $\mathcal{T}_1, \ldots, \mathcal{T}_M$ are compact infinite index sets (or uncertainty regions) [68,69][2]. The cutting-set algorithm solves (B.24) by solving a sequence of sampled problems. The ith sampled problem is given by

$$\begin{aligned} \min_{\mathbf{x}} \quad & f_0(\mathbf{x}) \\ \text{s.t.} \quad & f_m(\mathbf{x}, t) \leq 0,\ \forall t \in \mathcal{T}_m^{(i)}, m \in \mathcal{M} \end{aligned} \tag{B.25}$$

where $\mathcal{T}_m^{(i)} \subset \mathcal{T}_m$ is a finite subset. Let $\mathcal{F}^{(i)}$ be the feasible set of the ith problem, and $\bar{\mathbf{x}}^{(i)} \in \mathcal{F}^{(i)}$ be a feasible solution (not necessarily optimum). We assume that $\mathcal{F}^{(1)}$ is compact, $f_0(\cdot)$ and $f_1(\cdot, t), \ldots, f_M(\cdot, t)$ are continuously differentiable in $\mathbf{x} \in \mathcal{F}^{(1)}$, and the pessimization step is exact. Under such assumptions, it follows from [10, Section 5.2] that the iterates generated by the cutting-set algorithm converge to a feasible point of problem (B.24). In particular, we have

$$f_m(\bar{\mathbf{x}}, t) \leq 0, \forall t \in \mathcal{T}_m, m \in \mathcal{M} \tag{B.26}$$

[2] $\mathbf{x}$ is the optimization variable here and should not be confused with the transmit signal.

where $\bar{\mathbf{x}}$ is a limit point of the algorithm. Note that while global optimality of the optimization step is assumed in [10], it is not necessary for convergence and the feasibility of its limit point. This is also shown in the proof of [105, Theorem 2.1]. Next, we show that if $\bar{\mathbf{x}}^{(i)}$ is a KKT point of (B.25) for all i, then $\bar{\mathbf{x}}$ is a KKT point of (B.24).

The Lagrangian of (B.25) is given by

$$L(\mathbf{x}, \boldsymbol{\lambda}^{(i)}) = f_0(\mathbf{x}) + \sum_{m=1}^{M} \sum_{t \in \mathcal{T}_m^{(i)}} \lambda_{m,t}^{(i)} f_m(\mathbf{x}, t) \tag{B.27}$$

where $\boldsymbol{\lambda}^{(i)} \triangleq \left\{\lambda_{m,t}^{(i)} \mid t \in \mathcal{T}_m^{(i)}, m \in \mathcal{M}\right\}$ is the associated set of non-negative multipliers. We define the discrete measures $\mu_1^{(i)}, \ldots, \mu_M^{(i)}$ on $\mathcal{T}_1, \ldots, \mathcal{T}_M$ respectively such that

$$\mu_m^{(i)}(t) = \begin{cases} \lambda_{m,t}^{(i)}, & \forall t \in \mathcal{T}_m^{(i)} \\ 0, & \forall t \in \mathcal{T}_m \setminus \mathcal{T}_m^{(i)}. \end{cases} \tag{B.28}$$

It follows that the Lagrangian in (B.27) can be expressed as

$$L(\mathbf{x}, \boldsymbol{\mu}^{(i)}) = f_0(\mathbf{x}) + \sum_{m=1}^{M} \int_{t \in \mathcal{T}_m} f_m(\mathbf{x}, t) \mathrm{d}\mu_m^{(i)}(t) \tag{B.29}$$

where $\boldsymbol{\mu}^{(i)} \triangleq \left\{\mu_m^{(i)} \mid m \in \mathcal{M}\right\}$. Let $(\bar{\mathbf{x}}^{(i)}, \bar{\boldsymbol{\mu}}^{(i)})$ denote the KKT solution of problem (B.25) obtained in the i iteration and suppose that some regularity condition holds[3]. The corresponding KKT optimality conditions are given by

$$\nabla_{\mathbf{x}} L(\bar{\mathbf{x}}^{(i)}, \bar{\boldsymbol{\mu}}^{(i)}) = \mathbf{0} \tag{B.30a}$$

$$f_m(\bar{\mathbf{x}}^{(i)}, t) \leq 0, \forall t \in \mathcal{T}_m^{(i)}, m \in \mathcal{M} \tag{B.30b}$$

$$\bar{\mu}_m^{(i)} \geq 0, \forall m \in \mathcal{M} \tag{B.30c}$$

$$\int_{t \in \mathcal{T}_m} f_m(\bar{\mathbf{x}}^{(i)}, t) \mathrm{d}\bar{\mu}_m^{(i)}(t) = 0, \forall m \in \mathcal{M} \tag{B.30d}$$

where $\bar{\mu}_m^{(i)} \geq 0$ means that the measure is non-negative.

The sequence $\left\{\bar{\mathbf{x}}^{(i)}\right\}_{i=1}^{\infty}$ lies in the compact set $\mathcal{F}^{(1)}$, as $\mathcal{F}^{(1)} \supseteq \mathcal{F}^{(i)}$ for all i. Hence, there exists a subsequence $\left\{\bar{\mathbf{x}}^{(i_r)}\right\}_{r=1}^{\infty}$ converging to $\bar{\mathbf{x}}$. The regularity condition implies that at each $\bar{\mathbf{x}}^{(i)}$, the set of KKT multipliers that satisfy (B.30) is bounded [106]. Therefore, it is assumed without loss of generality that the subsequence $\left\{\bar{\boldsymbol{\mu}}^{(i_r)}\right\}_{r=1}^{\infty}$ converges weakly to the accumulation point $\bar{\boldsymbol{\mu}}$. Combining these observations with the continuity of the objective and constraint functions and their gradients, it can be shown that $\|\nabla_{\mathbf{x}} L(\bar{\mathbf{x}}, \bar{\boldsymbol{\mu}})\| = 0$ and $\int_{t \in \mathcal{T}_m} f_m(\bar{\mathbf{x}}, t) \mathrm{d}\bar{\mu}_m(t) = 0$ using the same steps given in the proof of [105, Theorem 2.1]. It

[3] In particular, it is assumed that the Mangasarian-Fromovitz Constraint Qualification (MFCQ) holds at stationary points of (B.24) and (B.25).

follows that the solution $(\bar{\mathbf{x}}, \bar{\boldsymbol{\mu}})$ satisfies

$$\nabla_{\mathbf{x}} L(\bar{\mathbf{x}}, \bar{\boldsymbol{\mu}}) = \mathbf{0} \tag{B.31a}$$

$$\bar{\mu}_m \geq 0, \forall m \in \mathcal{M} \tag{B.31b}$$

$$\int_{t \in \mathcal{T}_m} f_m(\bar{\mathbf{x}}, t) \mathrm{d}\bar{\mu}_m(t) = 0, \forall m \in \mathcal{M} \tag{B.31c}$$

where (B.31b) follows from (B.30c). Combining (B.31) with (B.26) implies that $(\bar{\mathbf{x}}, \bar{\boldsymbol{\mu}})$ satisfies the KKT conditions of problem (B.24), and $\bar{\mathbf{x}}$ is a KKT point[4]. Since the iterates lie in a compact set, the result holds for any sequence of iterates generated by the algorithm.

Now, we observe that problems (4.9) and (4.28) are instances of problems (B.24) and (B.25) respectively, with continuously differentiable objective and constraint functions [8, 51]. Also, $\mathbf{P}$ lies in the compact set given by $\left\{\mathbf{P} \mid \mathrm{tr}\left(\mathbf{P}\mathbf{P}^H\right) \leq P_{\mathrm{t}}\right\}$. The same holds for the rate variables $\bar{R}_{\mathrm{t}}$ and $\bar{\mathbf{c}}$, which belong to compact rate regions. Hence, the feasible sets for (4.9) and (4.28) are compact, which completes the proof. □

B.4. Solving (4.36) to Global Optimality

The auxiliary problems in Lemma 4.2 are rewritten as

$$\mathcal{D}_k(\lambda_k):$$
$$\max_{\mathbf{h}_k \in \mathbb{H}_k} \mathbf{h}_k^H \underbrace{\left((1-\lambda_k)\bar{\mathbf{Q}}_k - \lambda_k \mathbf{Q}_k\right)}_{\mathbf{A}_k(\lambda_k)} \mathbf{h}_k + (1-\lambda_k)\sigma_{\mathrm{n}}^2 \tag{B.32}$$

$$\mathcal{D}_{\mathrm{c},k}(\lambda_{\mathrm{c},k}):$$
$$\max_{\mathbf{h}_{\mathrm{c},k} \in \mathbb{H}_k} \mathbf{h}_{\mathrm{c},k}^H \underbrace{\left((1-\lambda_{\mathrm{c},k})\mathbf{Q}_{\mathrm{p}} - \lambda_{\mathrm{c},k}\mathbf{Q}_{\mathrm{c}}\right)}_{\mathbf{A}_{\mathrm{c},k}(\lambda_{\mathrm{c},k})} \mathbf{h}_{\mathrm{c},k} + (1-\lambda_{\mathrm{c},k})\sigma_{\mathrm{n}}^2 \tag{B.33}$$

where $\mathbf{Q}_k \triangleq \mathbf{p}_k \mathbf{p}_k^H$, $\mathbf{Q}_{\mathrm{c}} \triangleq \mathbf{p}_{\mathrm{c}} \mathbf{p}_{\mathrm{c}}^H$, $\mathbf{Q}_{\mathrm{p}} \triangleq \sum_{k=1}^K \mathbf{Q}_k$, and $\bar{\mathbf{Q}}_k \triangleq \mathbf{Q}_{\mathrm{p}} - \mathbf{Q}_k$. This follows from substituting the receive power and interference expressions in (2.37). For given parameters, (B.32) and (B.33) are QCQPs, where $\mathbf{A}_k(\lambda_k)$ and $\mathbf{A}_{\mathrm{c},k}(\lambda_{\mathrm{c},k})$ are symmetric and possibly indefinite[5]. Hence, (B.32) and (B.33) are non-convex optimization problems in general. For this reason, we resort to relaxation.

First, we introduce the matrix variables $\mathbf{X}_k = \mathbf{h}_k \mathbf{h}_k^H$ and $\mathbf{X}_{\mathrm{c},k} = \mathbf{h}_{\mathrm{c},k} \mathbf{h}_{\mathrm{c},k}^H$ from which the quadratic terms in (B.32) and (B.33) are eliminated by writing $\mathbf{h}_k^H \mathbf{A}_k \mathbf{h}_k = \mathrm{tr}\left(\mathbf{X}_k \mathbf{A}_k\right)$ and $\mathbf{h}_{\mathrm{c},k}^H \mathbf{A}_{\mathrm{c},k} \mathbf{h}_{\mathrm{c},k} = \mathrm{tr}\left(\mathbf{X}_{\mathrm{c},k} \mathbf{A}_{\mathrm{c},k}\right)$. Next, the equalities associated with the introduced matrices are relaxed into inequalities such that $\mathbf{X}_k \succeq \mathbf{h}_k \mathbf{h}_k^H$ and $\mathbf{X}_{\mathrm{c},k} \succeq \mathbf{h}_{\mathrm{c},k} \mathbf{h}_{\mathrm{c},k}^H$. The resulting

[4] Note that under the aforementioned assumptions, the semi-infinite problem in (B.24) has a finite number of active constraints at KKT points. Hence, the measures $\bar{\mu}_1, \ldots, \bar{\mu}_M$ have finite supports [68, 69].

[5] Updating the parameters using Dinkelbach's algorithm, we have $\lambda_k, \lambda_{\mathrm{c},k} \in [0, 1]$. $\mathbf{A}_k(0), \mathbf{A}_{\mathrm{c},k}(0) \succeq 0$, while $\mathbf{A}_k(1), \mathbf{A}_{\mathrm{c},k}(1) \preceq 0$. Otherwise, they are generally indefinite.

relaxed problems are formulated as

$$\mathcal{D}_k^{\mathrm{r}}(\lambda_k): \begin{cases} \max\limits_{\mathbf{X}_k, \mathbf{h}_k} & \mathrm{tr}\big(\mathbf{X}_k \mathbf{A}_k(\lambda_k)\big) + (1-\lambda_k)\sigma_{\mathrm{n}}^2 \\ \text{s.t.} & \mathrm{tr}(\mathbf{X}_k) - 2\Re(\mathbf{h}_k^H \widehat{\mathbf{h}}_k) + \widehat{\mathbf{h}}_k^H \widehat{\mathbf{h}}_k \leq \delta_k^2 \\ & \begin{bmatrix} \mathbf{X}_k & \mathbf{h}_k \\ \mathbf{h}_k^H & 1 \end{bmatrix} \succeq 0 \end{cases} \tag{B.34}$$

$$\mathcal{D}_{\mathrm{c},k}^{\mathrm{r}}(\lambda_{\mathrm{c},k}): \begin{cases} \max\limits_{\mathbf{X}_{\mathrm{c},k}, \mathbf{h}_{\mathrm{c},k}} & \mathrm{tr}\big(\mathbf{X}_{\mathrm{c},k} \mathbf{A}_{\mathrm{c},k}(\lambda_{\mathrm{c},k})\big) + (1-\lambda_{\mathrm{c},k})\sigma_{\mathrm{n}}^2 \\ \text{s.t.} & \mathrm{tr}(\mathbf{X}_{\mathrm{c},k}) - 2\Re(\mathbf{h}_{\mathrm{c},k}^H \widehat{\mathbf{h}}_k) + \widehat{\mathbf{h}}_k^H \widehat{\mathbf{h}}_k \leq \delta_k^2 \\ & \begin{bmatrix} \mathbf{X}_{\mathrm{c},k} & \mathbf{h}_{\mathrm{c},k} \\ \mathbf{h}_{\mathrm{c},k}^H & 1 \end{bmatrix} \succeq 0 \end{cases} \tag{B.35}$$

where the relaxed inequalities are rewritten using the Schur Complement. (B.34) and (B.35) are SDPs and can be efficiently solved. Due to the relaxations, the feasible sets in (B.34) and (B.35) contain their counterparts in (B.32) and (B.33). It follows that $\mathcal{D}_k^{\mathrm{r}}(\lambda_k) \geq \mathcal{D}_k(\lambda_k)$ and $\mathcal{D}_{\mathrm{c},k}^{\mathrm{r}}(\lambda_{\mathrm{c},k}) \geq \mathcal{D}_{\mathrm{c},k}(\lambda_{\mathrm{c},k})$. Before proceeding to the next result, we denote the optimum solutions of (B.34) and (B.35) as $\big(\mathbf{X}_k^\circ, \mathbf{h}_k^\circ\big)$ and $\big(\mathbf{X}_{\mathrm{c},k}^\circ, \mathbf{h}_{\mathrm{c},k}^\circ\big)$ respectively.

Lemma B.3. *The relaxations in* (B.34) *and* (B.35) *are tight at optimality, i.e.* $\mathbf{X}_k^\circ = \mathbf{h}_k^\circ \mathbf{h}_k^{\circ H}$ *and* $\mathbf{X}_{\mathrm{c},k}^\circ = \mathbf{h}_{\mathrm{c},k}^\circ \mathbf{h}_{\mathrm{c},k}^{\circ H}$. *As a result,* $\mathbf{h}_k^\circ$ *and* $\mathbf{h}_{\mathrm{c},k}^\circ$ *are optimum solutions for* (B.32) *and* (B.33) *respectively. Finally, we have* $\mathcal{D}_k^{\mathrm{r}}(\lambda_k) = \mathcal{D}_k(\lambda_k)$ *and* $\mathcal{D}_{\mathrm{c},k}^{\mathrm{r}}(\lambda_{\mathrm{c},k}) = \mathcal{D}_{\mathrm{c},k}(\lambda_{\mathrm{c},k})$.

Lemma B.3 follows directly from [52, Appendix B.1], by noting that (B.32) and (B.33) are QCQPs, with a single constraint each, that satisfy Slater's condition.

C. Proofs for Chapter 5

Before going into the proofs of Theorem 5.1 and Theorem 5.2, we start with some important definitions. For the conventional transmission in Section 5.2, recall that a precoding scheme is denoted by $\{\mathbf{P}_{\mathrm{p}}(P_{\mathrm{t}})\}_{P_{\mathrm{t}}}$. The associated powers $q_1, \ldots, q_M$ scale as $O(P_{\mathrm{t}}^{a_1}), \ldots, O(P_{\mathrm{t}}^{a_M})$ respectively, where $a_1, \ldots, a_M \in [0,1]$ are the corresponding scaling factors. For any set of scaling factors, let $\bar{a} \triangleq \max_{m\in\mathcal{M}} a_m$ be the maximum scaling.

A pecoding scheme has certain inter-group interference nulling capabilities, which depend on the system setup: N_{t}, M and the sizes of groups, and the design itself. Let $\mathcal{I}_m \subset \mathcal{M}$ be the set of groups with precoding vectors that interfere with the mth group, i.e. the precoding vector of each group in $\mathcal{I}_m$ causes non-zero interference[1] to at least one user in $\mathcal{G}_m$. Moreover, we denote the exponent of the dominant interference by $\bar{a}_m \triangleq \max_{j\in\mathcal{I}_m} a_j$. Hence, there exists at least one SINR scaling as $\gamma_i = O(P_{\mathrm{t}}^{a_m-\bar{a}_m})$ such that $i \in \mathcal{G}_m$. Recalling the DoF definitions in Section 5.3.1, we can write

$$d_m \leq \left(a_m - \bar{a}_m\right)^+ \tag{C.1}$$

where $(\cdot)$ follows from the fact that the DoF is non-negative. We recall that for a given precoding scheme, the achievable MMF-DoF satisfies $d \leq d_m$ for all $m \in \mathcal{M}$.

For the RS strategy in Section 5.5, a precoding scheme is denoted by $\{\mathbf{P}(P_{\mathrm{t}})\}_{P_{\mathrm{t}}}$. The power allocation and power scaling factors for the designated stream are as defined above. On the other hand, $q_{\mathrm{c}} = O(P_{\mathrm{t}}^{a_{\mathrm{c}}})$ with $a_{\mathrm{c}} \in [0,1]$ is the power allocated to the common stream. Since at each receiver, the common stream is decoded while treating the designated stream as noise, the common DoF is given by

$$d_{\mathrm{c}} \leq \left(a_{\mathrm{c}} - \bar{a}\right)^+ \leq 1 - \bar{a} \tag{C.2}$$

which is limited by the maximum power scaling across all designated streams. The fraction of d_{c} allocated to the mth groups is denoted by c_m, where $\sum_{m=1}^{M} c_m = d_{\mathrm{c}}$. Hence, the mth group-DoF is given by $c_m + d_m$, consisting of a common part and a designated part. Hence, the MMF-DoF for a given precoding scheme satisfies $d_{\mathrm{RS}} \leq c_m + d_m$ for all $m \in \mathcal{M}$.

[1] This interference scales as $|\mathbf{h}_i^H \mathbf{p}_j|^2 = O(P_{\mathrm{t}}^{a_j})$ for some $i \in \mathcal{G}_m$ and $j \in \mathcal{I}_m$.

C.1. Proof of Theorem 5.1

The achievability of the MMF-DoF in (5.13) is shown in Section 5.3.2. Here we show the converse, i.e. $d \leq d^*$ for any feasible precoding scheme. Since $d \leq 1$ is shown in (5.9), we focus on the two other cases.

For $d^* = 0.5$, it is sufficient to show that $d \leq 0.5$ for $N_\mathrm{t} = N_1 - 1$, as further decreasing the number of antennas cannot increase the DoF. Since $N_\mathrm{t} < N_1$, at least one group sees interference from $\mathbf{p}_1$ for any precoding scheme. Let $\mathcal{G}_{m_1}$ be a group that sees such interference, i.e. $1 \in \mathcal{I}_{m_1}$. We may assume that $a_{m_1} > \bar{a}_{m_1}$, as the contrary will limit the MMF-DoF to 0 as seen from (C.1). Next, we write the following set of inequalities

$$\begin{aligned} d &\leq \frac{d_1 + d_{m_1}}{2} && \text{(C.3)} \\ &\leq \frac{a_1 + a_{m_1} - \bar{a}_{m_1}}{2} && \text{(C.4)} \\ &\leq \frac{a_1 + a_{m_1} - a_1}{2} && \text{(C.5)} \\ &\leq 0.5. && \text{(C.6)} \end{aligned}$$

(C.3) follows from the fact that the minimum group-DoF is upper-bounded by the average of any number of group-DoFs. (C.4) follows from (C.1), while (C.5) follows from $1 \in \mathcal{I}_{m_1}$. Finally, (C.6) follows from $a_{m_1} \leq 1$. This completes the converse for this part.

Now, we show that $d \leq 0$ for $N_\mathrm{t} = N_M - 1$. Note that $N_\mathrm{t} < N_m$ for all $m \in \mathcal{M}$. Hence, each beamforming vector causes interference to at least one group it is not intended to. Equivalently, we have $\bigcup_{m \in \mathcal{M}} \mathcal{I}_m = \mathcal{M}$. Therefore, for any power allocation with exponents $a_1, \ldots, a_M$, there exists at least one group that sees interference from $\mathbf{p}_{m_1}$, where $a_{m_1} = \bar{a}$. Let the index of such group be m_2, i.e $m_1 \in \mathcal{I}_{m_2}$. We have

$$d_{m_2} \leq \left(a_{m_2} - \bar{a}\right)^+ = 0. \tag{C.7}$$

Hence, it follows that $d \leq d_{m_2} \leq 0$, which completes the proof. □

C.2. Proof of Theorem 5.2

We recall that for any given N_t, the maximum number of groups that can be served with interference-free beamforming is denoted by M_ms, which is expressed in (5.26). Hence, for any feasible precoding scheme, at least $M_\mathrm{ss} = M - M_\mathrm{ms}$ groups receive non-zero interference from the designated beams. In this proof, we show that $d_\mathrm{RS} \leq \frac{1}{1+M_\mathrm{ss}}$. Before proceeding, we present the following result which plays an important role in this proof.

Lemma C.1. *For any precoding scheme, $\mathbf{p}_1$ interferes with at least M_ss groups. Moreover, each of $\mathbf{p}_2, \ldots, \mathbf{p}_M$ interfere with at least $M_\mathrm{ss} - 1$ groups.*

Proof. We recall that to place $\mathbf{p}_m$ in the null space of all groups in $\mathcal{S}_m \subseteq \mathcal{M} \setminus m$, the

following condition should be satisfied:

$$N_{\mathrm{t}} \geq 1 + \sum_{j \in \mathcal{S}_m} G_j. \tag{C.8}$$

Hence, finding the minimum number of groups $\mathbf{p}_m$ interferes with is equivalent to finding the maximum $|\mathcal{S}_m|$. Following the same approach in (5.26), this is found by packing the maximum number of groups in $\mathcal{M} \setminus m$ into $N_{\mathrm{t}} - 1$, starting with the smallest. To this end, we express N_{t} in terms of M_{ms} and group sizes as

$$N_{\mathrm{t}} = \begin{cases} 1 + \bar{N}, & M_{\mathrm{ms}} = 1 \\ 1 + \sum_{j=2}^{M_{\mathrm{ms}}} G_j + \bar{N}, & M_{\mathrm{ms}} \geq 2 \end{cases} \tag{C.9}$$

where $0 \leq \bar{N} < G_{M_{\mathrm{ms}}+1}$. This follows directly from (5.26).

First, we start with $\mathbf{p}_1$. This can be placed in the null space of at most $M_{\mathrm{ms}} - 1$ groups, i.e. groups $\mathcal{G}_2, \ldots, \mathcal{G}_{M_{\mathrm{ms}}}$. This follows from (C.8) and observing that

$$1 + \sum_{j=2}^{M_{\mathrm{ms}}} G_j \leq N_{\mathrm{t}} < 1 + \sum_{j=2}^{M_{\mathrm{ms}}+1} G_j. \tag{C.10}$$

Hence, $\mathbf{p}_1$ causes interference to the remaining groups after excluding $\mathcal{G}_1$, i.e. M_{ss} groups.

Next, we consider $\mathbf{p}_m$ for all $m \in \{2, \ldots, M_{\mathrm{ms}}\}$. We can write

$$N_{\mathrm{t}} = 1 + \sum_{j=1, j \neq m}^{M_{\mathrm{ms}}} G_j + (G_m - G_1) + \bar{N} \tag{C.11}$$

$$< 1 + \sum_{j=1, j \neq m}^{M_{\mathrm{ms}}} G_j + G_{M_{\mathrm{ms}}+1} + G_{M_{\mathrm{ms}}+2} \tag{C.12}$$

where (C.12) follows from $\bar{N} < G_{M_{\mathrm{ms}}+1}$ and $G_m - G_1 < G_{M_{\mathrm{ms}}+2}$. Hence, in the best case scenario, $\mathbf{p}_m$ is placed in the null space of groups $\mathcal{G}_1, \ldots, \mathcal{G}_{m-1}, \mathcal{G}_{m+1}, \ldots \mathcal{G}_{M_{\mathrm{ms}}+1}$, and causes interference to the remaining $M_{\mathrm{ss}} - 1$ groups (by excluding $\mathcal{G}_m$).

Finally, consider $\mathbf{p}_m$ for all $m \in \{M_{\mathrm{ms}} + 1, \ldots, M\}$. Here, the best scenario occurs when $\bar{N} \geq G_1$, from which we can write

$$1 + \sum_{j=1}^{M_{\mathrm{ms}}} G_j \leq N_{\mathrm{t}} < 1 + \sum_{j=2}^{M_{\mathrm{ms}}+1} G_j. \tag{C.13}$$

It follows that $\mathbf{p}_m$ is placed in the null space of groups $\mathcal{G}_1, \ldots, \mathcal{G}_{M_{\mathrm{ms}}}$, and causes interference to the remaining $M_{\mathrm{ss}} - 1$ groups. □

The minimum group-DoF is upper-bounded by the average of any number of group-DoFs.

Hence, taking the subset $\mathcal{S} \subseteq \mathcal{M}$, we can write

$$d_{\mathrm{RS}} \leq \frac{\sum_{m\in\mathcal{S}} d_m + c_m}{|\mathcal{S}|} \leq \frac{\sum_{m\in\mathcal{S}} d_m + d_{\mathrm{c}}}{|\mathcal{S}|}. \tag{C.14}$$

where the right-hand side inequality follows from the fact that $\sum_{m\in\mathcal{S}} c_m \leq \sum_{m\in\mathcal{M}} c_m = d_{\mathrm{c}}$. Now, we need to find the *right* subset $\mathcal{S}$ which gives us a meaningful upper-bound in closed-form, that applies to any feasible precoding scheme.

Let $\bar{m} \in \mathcal{M}$ be the index of the group with the largest power scaling, i.e. $a_{\bar{m}} = \bar{a}$. Moreover, let $\bar{\mathcal{S}} \subseteq \mathcal{M} \setminus \bar{m}$ be the set of groups that see interference from $\mathbf{p}_{\bar{m}}$. From Lemma C.1, we know that $|\bar{\mathcal{S}}| \geq M_{\mathrm{ss}} - 1$. For the upper-bound, we assume that $|\bar{\mathcal{S}}| = M_{\mathrm{ss}} - 1$, as increasing the number of groups that see interference does not increase the DoF. We also assume that $\bar{m} \neq 1$, as the contrary does not influence the result as we see later. Since $\mathbf{p}_1$ interferes with at least M_{ss} groups (from Lemma C.1), we have at least one group that sees interference from $\mathbf{p}_1$ and is not in $\bar{\mathcal{S}}$. Let the index of such group be m_1. From (C.1), note that $a_1 \geq a_{m_1}$ implies $d_1 + d_{m_1} \leq a_1$, while if $a_1 \leq a_{m_1}$ implies $d_1 + d_{m_1} \leq a_{m_1}$. We assume, without loss of generality, that $a_1 \geq a_{m_1}$, as the contrary does not affect the result as we see next. The set of groups for the upper-bound is taken as $\mathcal{S} = \{1, m_1, \bar{\mathcal{S}}\}$ with $|\mathcal{S}| = M_{\mathrm{ss}} + 1$. Now, we can write

$$d_{\mathrm{RS}} \leq \frac{d_1 + d_{m_1} + \sum_{m\in\bar{\mathcal{S}}} d_m + d_{\mathrm{c}}}{M_{\mathrm{ss}} + 1} \tag{C.15}$$

$$\leq \frac{a_1 + 1 - \bar{a}}{M_{\mathrm{ss}} + 1} \tag{C.16}$$

$$\leq \frac{1}{M_{\mathrm{ss}} + 1} \tag{C.17}$$

where (C.16) follows from the fact that $d_m = 0$ for all $m \in \bar{\mathcal{S}}$ and (C.2), where (C.17) follows from $\bar{a} \geq a_1$. Note that if we assume that $\bar{m} = 1$, then we can also assume that $|\bar{\mathcal{S}}| = M_{\mathrm{ss}}$. As a result, the same upper-bound holds by adding m_1 to $\bar{\mathcal{S}}$ and setting $d_{m_1} = 0$. This completes the proof. □

Bibliography

[1] B. Clerckx and C. Oestges, *MIMO Wireless Networks: Channels, Techniques and Standards for Multi-antenna, Multi-user and Multi-cell Systems.* Academic Press, 2013.

[2] S. Yang, M. Kobayashi, D. Gesbert, and X. Yi, "Degrees of freedom of time correlated MISO broadcast channel with delayed CSIT," *IEEE Transactions on Information Theory*, vol. 59, no. 1, pp. 315–328, 2013.

[3] C. Hao, Y. Wu, and B. Clerckx, "Rate analysis of two-receiver MISO broadcast channel with finite rate feedback: A rate-splitting approach," *IEEE Transactions on Communications*, vol. 63, no. 9, pp. 3232–3246, Sep. 2015.

[4] B. Yuan and S. A. Jafar, "Elevated multiplexing and signal space partitioning in the 2 user MIMO IC with partial CSIT," *arXiv preprint arXiv:1604.00582*, 2016.

[5] E. Björnson, M. Bengtsson, and B. Ottersten, "Optimal multiuser transmit beamforming: A difficult problem with a simple solution structure [lecture notes]," *IEEE Signal Processing Magazine*, vol. 31, no. 4, pp. 142–148, Jul. 2014.

[6] N. Sidiropoulos, T. Davidson, and Z.-Q. Luo, "Transmit beamforming for physical-layer multicasting," *IEEE Transactions on Signal Processing*, vol. 54, no. 6, pp. 2239–2251, Jun. 2006.

[7] S. Christensen, R. Agarwal, E. Carvalho, and J. Cioffi, "Weighted sum-rate maximization using weighted MMSE for MIMO-BC beamforming design," *IEEE Transactions on Wireless Communications*, vol. 7, no. 12, pp. 4792–4799, Dec. 2008.

[8] Q. Shi, M. Razaviyayn, Z.-Q. Luo, and C. He, "An iteratively weighted MMSE approach to distributed sum-utility maximization for a MIMO interfering broadcast channel," *IEEE Transactions on Signal Processing*, vol. 59, no. 9, pp. 4331–4340, Sep. 2011.

[9] A. Shapiro, D. Dentcheva *et al.*, *Lectures on stochastic programming: modeling and theory.* SIAM, 2009, vol. 9.

[10] A. Mutapcic and S. Boyd, "Cutting-set methods for robust convex optimization with pessimizing oracles," *Optimization Methods and Software*, vol. 24, no. 3, pp. 381–406, 2009.

[11] T. M. Cover and J. A. Thomas, *Elements of information theory.* John Wiley & Sons, 2006.

[12] A. El Gamal and Y.-H. Kim, *Network information theory.* Cambridge university press, 2011.

[13] B. Hassibi and B. Hochwald, "How much training is needed in multiple-antenna wireless links?" *IEEE Transactions on Information Theory*, vol. 49, no. 4, pp. 951–963, Apr. 2003.

[14] T. L. Marzetta, "How much training is required for multiuser MIMO?" in *Proc. Asilomar*, Oct. 2006, pp. 359–363.

[15] D. Love, R. Heath, V. Lau, D. Gesbert, B. Rao, and M. Andrews, "An overview of limited feedback in wireless communication systems," *IEEE Journal on Selected Areas in Communications*, vol. 26, no. 8, pp. 1341–1365, Oct. 2008.

[16] G. Caire and K. Kumar, "Information theoretic foundations of adaptive coded modulation," *Proceedings of the IEEE*, vol. 95, no. 12, pp. 2274–2298, Dec. 2007.

[17] D. Tse and P. Viswanath, *Fundamentals of wireless communication.* Cambridge university press, 2005.

[18] G. Caire, N. Jindal, M. Kobayashi, and N. Ravindran, "Multiuser MIMO achievable rates with downlink training and channel state feedback," *IEEE Transactions on Information Theory*, vol. 56, no. 6, pp. 2845–2866, Jun. 2010.

[19] N. Jindal, "MIMO broadcast channels with finite-rate feedback," *IEEE Transactions on Information Theory*, vol. 52, no. 11, pp. 5045–5060, Nov. 2006.

[20] A. G. Davoodi and S. A. Jafar, "Aligned image sets under channel uncertainty: Settling a conjecture by Lapidoth, Shamai and Wigger on the collapse of degrees of freedom under finite precision CSIT," *arXiv preprint arXiv:1403.1541*, 2014.

[21] G. Caire and S. Shamai, "On the achievable throughput of a multiantenna Gaussian broadcast channel," *IEEE Transactions on Information Theory*, vol. 49, no. 7, pp. 1691–1706, Jul. 2003.

[22] A. Lapidoth, S. Shamai, and M. Wigger, "On the capacity of fading MIMO broadcast channels with imperfect transmitter side-information," in *Proc. Allerton Conf. on Commun., Control and Computing*, Sep. 2005.

[23] H. Weingarten, Y. Steinberg, and S. S. Shamai, "The capacity region of the gaussian multiple-input multiple-output broadcast channel," *IEEE Transactions on Information Theory*, vol. 52, no. 9, pp. 3936–3964, Sep. 2006.

[24] C. S. Vaze and M. K. Varanasi, "The degree-of-freedom regions of MIMO broadcast, interference, and cognitive radio channels with no CSIT," *IEEE Transactions on Information Theory*, vol. 58, no. 8, pp. 5354–5374, Aug. 2012.

[25] C. Hao and B. Clerckx, "MISO BC with imperfect and (un)matched CSIT in the frequency domain: DoF region and transmission strategies," in *Proc. IEEE PIMRC*, Sep. 2013, pp. 1–6.

[26] R. Tandon, S. Jafar, S. Shamai Shitz, and H. Poor, "On the synergistic benefits of alternating CSIT for the MISO broadcast channel," *IEEE Transactions on Information Theory*, vol. 59, no. 7, pp. 4106–4128, Jul. 2013.

[27] J. Chen and P. Elia, "Toward the performance versus feedback tradeoff for the two-user MISO broadcast channel," *IEEE Transactions on Information Theory*, vol. 59, no. 12, pp. 8336–8356, Dec. 2013.

[28] L. Zheng and D. Tse, "Diversity and multiplexing: A fundamental tradeoff in multiple-antenna channels," *IEEE Transactions on Information Theory*, vol. 49, no. 5, pp. 1073–1096, May 2003.

[29] A. Paulraj, R. Nabar, and D. Gore, *Introduction to space-time wireless communications*. Cambridge university press, 2003.

[30] U. Erez and S. ten Brink, "A close-to-capacity dirty paper coding scheme," *IEEE Transactions on Information Theory*, vol. 51, no. 10, pp. 3417–3432, Oct. 2005.

[31] Y. Sun, Y. Yang, A. D. Liveris, V. Stankovic, and Z. Xiong, "Near-capacity dirty-paper code design: A source-channel coding approach," *IEEE Transactions on Information Theory*, vol. 55, no. 7, pp. 3013–3031, Jul. 2009.

[32] A. Wiesel, Y. C. Eldar, and S. Shamai, "Zero-forcing precoding and generalized inverses," *IEEE Transactions on Signal Processing*, vol. 56, no. 9, pp. 4409–4418, Sep. 2008.

[33] L. Li and A. J. Goldsmith, "Capacity and optimal resource allocation for fading broadcast channels .I. Ergodic capacity," *IEEE Transactions on Information Theory*, vol. 47, no. 3, pp. 1083–1102, Mar. 2001.

[34] N. Jindal and A. Goldsmith, "Dirty-paper coding versus TDMA for MIMO broadcast channels," *IEEE Transactions on Information Theory*, vol. 51, no. 5, pp. 1783–1794, May 2005.

[35] Y. F. Liu, Y. H. Dai, and Z. Q. Luo, "Coordinated beamforming for MISO interference channel: Complexity analysis and efficient algorithms," *IEEE Transactions on Signal Processing*, vol. 59, no. 3, pp. 1142–1157, Mar. 2011.

[36] E. Björnson and E. Jorswieck, "Optimal resource allocation in coordinated multi-cell systems," *Foundations and Trends in Communications and Information Theory*, vol. 9, no. 2-3, pp. 381–p, 2013.

[37] M. Stojnic, H. Vikalo, and B. Hassibi, "Rate maximization in multi-antenna broadcast channels with linear preprocessing," *IEEE Transactions on Wireless Communications*, vol. 5, no. 9, pp. 2338–2342, Sep. 2006.

[38] S. Shi, M. Schubert, and H. Boche, "Rate optimization for multiuser MIMO systems with linear processing," *IEEE Transactions on Signal Processing*, vol. 56, no. 8, pp. 4020–4030, Aug. 2008.

[39] C. Guthy, W. Utschick, R. Hunger, and M. Joham, "Efficient weighted sum rate maximization with linear precoding," *IEEE Transactions on Signal Processing*, vol. 58, no. 4, pp. 2284–2297, Apr. 2010.

[40] D. P. Bertsekas, *Nonlinear programming*, 2nd ed. Belmont, MA: Athena Scientific, 1999.

[41] E. Biglieri, J. Proakis, and S. Shamai, "Fading channels: information-theoretic and communications aspects," *IEEE Transactions on Information Theory*, vol. 44, no. 6, pp. 2619–2692, Oct. 1998.

[42] L. Li and A. J. Goldsmith, "Capacity and optimal resource allocation for fading broadcast channels .II. Outage capacity," *IEEE Transactions on Information Theory*, vol. 47, no. 3, pp. 1103–1127, Mar. 2001.

[43] W. C. Li, T. H. Chang, C. Lin, and C. Y. Chi, "Coordinated beamforming for multiuser MISO interference channel under rate outage constraints," *IEEE Transactions on Signal Processing*, vol. 61, no. 5, pp. 1087–1103, Mar. 2013.

[44] J. Jose, N. Prasad, M. Khojastepour, and S. Rangarajan, "On robust weighted-sum rate maximization in MIMO interference networks," in *Proc. IEEE ICC 2011*, Jun. 2011, pp. 1–6.

[45] A. Tajer, N. Prasad, and X. Wang, "Robust linear precoder design for multi-cell downlink transmission," *IEEE Transactions on Signal Processing*, vol. 59, no. 1, pp. 235–251, Jan. 2011.

[46] F. Negro, I. Ghauri, and D. T. M. Slock, "Sum Rate maximization in the noisy MIMO interfering broadcast channel with partial CSIT via the expected weighted MSE," in *Proc. ISWCS*, Aug. 2012, pp. 576–580.

[47] M. Bashar, Y. Lejosne, D. Slock, and Y. Yuan-Wu, "MIMO broadcast channels with Gaussian CSIT and application to location based CSIT," in *Proc. ITA*, Feb. 2014, pp. 1–7.

[48] M. Razaviyayn, M. Boroujeni, and Z.-Q. Luo, "A stochastic weighted MMSE approach to sum rate maximization for a MIMO interference channel," in *Proc. IEEE SPAWC*, Jun. 2013, pp. 325–329.

[49] S. Shi, M. Schubert, and H. Boche, "Downlink MMSE transceiver optimization for multiuser MIMO systems: Duality and sum-MSE minimization," *IEEE Transactions on Signal Processing*, vol. 55, no. 11, pp. 5436–5446, Nov. 2007.

[50] R. Hunger, M. Joham, and W. Utschick, "On the MSE-duality of the broadcast channel and the multiple access channel," *IEEE Transactions on Signal Processing*, vol. 57, no. 2, pp. 698–713, Feb. 2009.

[51] M. Razaviyayn, M. Hong, and Z.-Q. Luo, "Linear transceiver design for a MIMO interfering broadcast channel achieving max–min fairness," *Signal Processing*, vol. 93, no. 12, pp. 3327 – 3340, 2013.

[52] S. P. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge university press, 2004.

[53] M. Razaviyayn, M. Hong, and Z.-Q. Luo, "A unified convergence analysis of block successive minimization methods for nonsmooth optimization," *SIAM Journal on Optimization*, vol. 23, no. 2, pp. 1126–1153, 2013.

[54] M. Razaviyayn, "Successive convex approximation: Analysis and applications," Ph.D. dissertation, UNIVERSITY OF MINNESOTA, 2014.

[55] B. E. Fristedt and L. F. Gray, *A modern approach to probability theory*. Springer, 1997.

[56] S. Kim, R. Pasupathy, and S. G. Henderson, "A guide to sample average approximation," in *Handbook of Simulation Optimization*. Springer, 2015, pp. 207–243.

[57] M. Grant, S. Boyd, and Y. Ye, "CVX: MATLAB software for disciplined convex programming [Online]," *Available: http://www.stanford.edu/ boyd/cvx*, 2008.

[58] M. Bengtsson and B. Ottersten, "Optimal and suboptimal transmit beamforming," in *Handbook of Antennas in Wireless Communications*, L. C. Godara, Ed. CRC Press, 2001.

[59] A. Wiesel, Y. Eldar, and S. Shamai, "Linear precoding via conic optimization for fixed MIMO receivers," *IEEE Transactions on Signal Processing*, vol. 54, no. 1, pp. 161–176, Jan. 2006.

[60] M. Schubert and H. Boche, "Solution of the multiuser downlink beamforming problem with individual SINR constraints," *IEEE Transactions on Vehicular Technology*, vol. 53, no. 1, pp. 18–28, Jan. 2004.

[61] ——, "Iterative multiuser uplink and downlink beamforming under SINR constraints," *IEEE Transactions on Signal Processing*, vol. 53, no. 7, pp. 2324–2334, Jul. 2005.

[62] S. Shi and M. Schubert, "MMSE transmit optimization for multi-user multi-antenna systems," in *Proc. IEEE ICASSP*, vol. 3, Mar. 2005, pp. iii/409–iii/412 Vol. 3.

[63] P. Ding, D. Love, and M. Zoltowski, "Multiple antenna broadcast channels with shape feedback and limited feedback," *IEEE Transactions on Signal Processing*, vol. 55, no. 7, pp. 3417–3428, Jul. 2007.

[64] A. Pascual-Iserte, D. Palomar, A. Perez-Neira, and M. Lagunas, "A robust maximin approach for MIMO communications with imperfect channel state information based on convex optimization," *IEEE Transactions on Signal Processing*, vol. 54, no. 1, pp. 346–360, Jan. 2006.

[65] M. Payaró, A. Pascual-Iserte, and M. A. Lagunas, "Robust power allocation designs for multiuser and multiantenna downlink communication systems through convex optimization," *IEEE Journal on Selected Areas in Communications*, vol. 25, no. 7, pp. 1390–1401, Sep. 2007.

[66] K.-Y. Wang, A.-C. So, T.-H. Chang, W.-K. Ma, and C.-Y. Chi, "Outage constrained robust transmit optimization for multiuser MISO downlinks: Tractable approximations by conic optimization," *IEEE Transactions on Signal Processing*, vol. 62, no. 21, pp. 5690–5705, Nov. 2014.

[67] S. Kassam and H. Poor, "Robust techniques for signal processing: A survey," *Proceedings of the IEEE*, vol. 73, no. 3, pp. 433–481, Mar. 1985.

[68] M. López and G. Still, "Semi-infinite programming," *European Journal of Operational Research*, vol. 180, no. 2, pp. 491 – 518, 2007.

[69] A. Shapiro, "Semi-infinite programming, duality, discretization and optimality conditions," *Optimization*, vol. 58, no. 2, pp. 133–161, 2009.

[70] A. Mutapcic, S. Kim, and S. Boyd, "A tractable method for robust downlink beamforming in wireless communications," in *Proc. Asilomar 2007*, Nov. 2007, pp. 1224–1228.

[71] N. Vucic, H. Boche, and S. Shi, "Robust transceiver optimization in downlink multiuser MIMO systems," *IEEE Transactions on Signal Processing*, vol. 57, no. 9, pp. 3576–3587, Sep. 2009.

[72] N. Vucic and H. Boche, "Robust QoS-constrained optimization of downlink multiuser MISO systems," *IEEE Transactions on Signal Processing*, vol. 57, no. 2, pp. 714–725, Feb. 2009.

[73] E. Song, Q. Shi, M. Sanjabi, R.-Y. Sun, and Z.-Q. Luo, "Robust SINR-constrained MISO downlink beamforming: When is semidefinite programming relaxation tight?" *EURASIP Journal on Wireless Communications and Networking*, vol. 2012, no. 1, pp. 1–11, 2012.

[74] M. Shenouda and T. Davidson, "Convex conic formulations of robust downlink precoder designs with quality of service constraints," *IEEE Journal of Selected Topics in Signal Processing*, vol. 1, no. 4, pp. 714–724, Dec. 2007.

[75] ——, "Nonlinear and linear broadcasting with QoS requirements: Tractable approaches for bounded channel uncertainties," *IEEE Transactions on Signal Processing*, vol. 57, no. 5, pp. 1936–1947, 2009.

[76] ——, "On the design of linear transceivers for multiuser systems with channel uncertainty," *IEEE Journal on Selected Areas in Communications*, vol. 26, no. 6, pp. 1015–1024, Aug. 2008.

[77] I. Pólik and T. Terlaky, "A survey of the S-lemma," *SIAM review*, vol. 49, no. 3, pp. 371–418, 2007.

[78] Z.-Q. Luo, W.-K. Ma, A.-C. So, Y. Ye, and S. Zhang, "Semidefinite relaxation of quadratic optimization problems," *IEEE Signal Processing Magazine*, vol. 27, no. 3, pp. 20–34, May 2010.

[79] W.-K. Ma, J. Pan, A. M.-C. So, and T.-H. Chang, "Unraveling the rank-one solution mystery of robust MISO downlink transmit optimization: A verifiable sufficient condition via a new duality result," *arXiv preprint arXiv:1602.01569*, 2016.

[80] S. A. Jafar, "Topological interference management through index coding," *IEEE Transactions on Information Theory*, vol. 60, no. 1, pp. 529–568, Jan. 2014.

[81] R. A. Horn and C. R. Johnson, *Matrix analysis*. Cambridge university press, 2012.

[82] Y. Eldar and N. Merhav, "A competitive minimax approach to robust estimation of random parameters," *IEEE Transactions on Signal Processing*, vol. 52, no. 7, pp. 1931–1946, Jul. 2004.

[83] P. Ubaidulla and A. Chockalingam, "Relay precoder optimization in MIMO-relay networks with imperfect CSI," *IEEE Transactions on Signal Processing*, vol. 59, no. 11, pp. 5473–5484, Nov. 2011.

[84] P. Ubaidulla and S. Aissa, "Robust two-way cognitive relaying: Precoder designs under interference constraints and imperfect CSI," *IEEE Transactions on Wireless Communications*, vol. 13, no. 5, pp. 2478–2489, May 2014.

[85] J. Liu, F. Gao, and Z. Qiu, "Robust transceiver design for downlink multiuser MIMO AF relay systems," *IEEE Transactions on Wireless Communications*, vol. 14, no. 4, pp. 2218–2231, Apr. 2015.

[86] W. Dinkelbach, "On nonlinear fractional programming," *Management Science*, vol. 13, no. 7, pp. 492–498, 1967.

[87] A. Ben-Tal and A. Nemirovski, *Lectures on modern convex optimization: analysis, algorithms, and engineering applications*, ser. MPS-SIAM Series on Optimization. Philadelphia, PA: SIAM, 2001.

[88] N. D. Sidiropoulos and T. N. Davidson, "Broadcasting with channel state information," in *Proc. IEEE Sensor Array and Multichannel Signal Processing Workshop (SAM)*, Jul. 2004, pp. 489–493.

[89] N. Jindal and Z.-Q. Luo, "Capacity limits of multiple antenna multicast," in *Proc. IEEE ISIT*, 2006, pp. 1841–1845.

[90] E. Karipidis, N. Sidiropoulos, and Z.-Q. Luo, "Quality of service and max-min fair transmit beamforming to multiple cochannel multicast groups," *IEEE Transactions on Signal Processing*, vol. 56, no. 3, pp. 1268–1279, Mar. 2008.

[91] E. Karipidis, N. D. Sidiropoulos, and Z. Q. Luo, "Far-field multicast beamforming for uniform linear antenna arrays," *IEEE Transactions on Signal Processing*, vol. 55, no. 10, pp. 4916–4927, Oct. 2007.

[92] N. Bornhorst and M. Pesavento, "An iterative convex approximation approach for transmit beamforming in multi-group multicasting," in *Proc. IEEE SPAWC*, Jun. 2011, pp. 426–430.

[93] A. Schad and M. Pesavento, "Max-min fair transmit beamforming for multi-group multicasting," in *Proc. Int. ITG WSA*, Mar. 2012, pp. 115–118.

[94] D. Christopoulos, S. Chatzinotas, and B. Ottersten, "Weighted fair multicast multigroup beamforming under per-antenna power constraints," *IEEE Transactions on Signal Processing*, vol. 62, no. 19, pp. 5132–5142, Oct. 2014.

[95] ——, "Multicast multigroup beamforming for per-antenna power constrained large-scale arrays," in *Proc. IEEE SPAWC*, Jun. 2015, pp. 271–275.

[96] Y. W. P. Hong, W. C. Li, T. H. Chang, and C. H. Lee, "Coordinated multicasting with opportunistic user selection in multicell wireless systems," *IEEE Transactions on Signal Processing*, vol. 63, no. 13, pp. 3506–3521, Jul. 2015.

[97] D. Christopoulos, S. Chatzinotas, and B. Ottersten, "Sum rate maximizing multigroup multicast beamforming under per-antenna power constraints," in *Proc. IEEE GLOBECOM*, Dec. 2014, pp. 3354–3359.

[98] T. Gou, C. R. C. M. da Silva, J. Lee, and I. Kang, "Partially connected interference networks with no CSIT: Symmetric degrees of freedom and multicast across alignment blocks," *IEEE Communications Letters*, vol. 17, no. 10, pp. 1893–1896, Oct. 2013.

[99] C. Hao and B. Clerckx, "MISO networks with imperfect CSIT: A topological rate-splitting approach," *arXiv preprint arXiv:1602.03768*, 2016.

[100] F. Wang, T. Peng, Y. Huang, and X. Wang, "Robust transceiver optimization for power-splitting based downlink MISO SWIPT systems," *IEEE Signal Processing Letters*, vol. 22, no. 9, pp. 1492–1496, Sep. 2015.

[101] Y. Saito, Y. Kishiyama, A. Benjebbour, T. Nakamura, A. Li, and K. Higuchi, "Non-orthogonal multiple access (NOMA) for cellular future radio access," in *Proc. IEEE VTC*, Jun. 2013, pp. 1–5.

[102] J. Zhang, F. Engelmann, and P. Elia, "Coded caching for reducing CSIT-feedback in wireless communications," in *Proc. Allerton Conference on Communication, Control, and Computing*, Sep. 2015, pp. 1099–1105.

[103] M. A. Maddah-Ali and U. Niesen, "Fundamental limits of caching," *IEEE Transactions on Information Theory*, vol. 60, no. 5, pp. 2856–2867, May 2014.

[104] S. Vorobyov, A. Gershman, and Z.-Q. Luo, "Robust adaptive beamforming using worst-case performance optimization: a solution to the signal mismatch problem," *IEEE Transactions on Signal Processing*, vol. 51, no. 2, pp. 313–324, Feb. 2003.

[105] S.-Y. Wu, D.-H. Li, L. Qi, and G. Zhou, "An iterative method for solving KKT system of the semi-infinite programming," *Optimization Methods and Software*, vol. 20, no. 6, pp. 629–643, 2005.

[106] J. Gauvin, "A necessary and sufficient regularity condition to have bounded multipliers in nonconvex programming," *Mathematical Programming*, vol. 12, no. 1, pp. 136–138, 1977.